Exploring the wheat virome using high-throughput Nanopore sequencing: A metagenomic and phylogenetic analysis

by

Nar Bahadur Ranabhat

B.S., Tribhuvan University, 2006
M.Sc., University of Hohenheim, 2012
M.S., Montana State University, 2017

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Plant Pathology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Abstract

Kansas is one of the top wheat-producing states in the United States. Many wheat viruses have been recognized as common yield-reducing factors. The synergistic impact of several concurrent infecting wheat viruses is responsible for millions of dollars lost in wheat production. Genetic resistance, using virus-resistant cultivars, is one of the key management practices of wheat viruses. The primary threat to durable genetic resistance is the presence of potential new virus variants in the field. The main objective of this study was to explore the field wheat virus population. We used the high-throughput Oxford Nanopore sequencing technique (ONT) to study the wheat virome. A survey was conducted in 2019, 2020, and 2021 in major wheat-growing counties of Kansas, and wheat leaves showing virus-like symptoms were collected. Total RNA was extracted, and cDNA sequencing libraries were made using the PCR-cDNA barcoding kit and loaded into ONT MinION flow cells. Sequencing reads were aligned to cereal virus references. We identified eight wheat viruses belonging to the genera: *Tritimovirus*, *Poacevirus*, *Emaravirus*, *Bromovirus*, *Luteovirus*, *Polerovirus*, *Bymovirus*, and *Furovirus.* We recorded mixed infections of two to five viruses in a single sample. Wheat streak mosaic virus (WSMV) + triticum mosaic virus (TriMV) mixed infection was the most predominant infection (16.7%), followed by WSMV + TriMV + brome mosaic virus (BMV) (11.9%) and WSMV single infection (11.9%). Phylogenetic analysis of the whole genomes of WSMV revealed the wide distribution of isolates into clades and subclades including European isolate. Potential WSMV recombinant isolates were found. BMV was identified for the first time in Kansas wheat. We used genetic and evolutionary approaches to characterize BMV isolates. On average, US BMV isolates showed low divergence. Coding regions of all BMV RNAs were under purifying or negative selection pressure. The whole-genome sequences of multiple isolates of High Plains

wheat mosaic emaravirus and soilborne wheat mosaic virus were characterized. Additionally, virus viability in the inoculum over time was determined, and the relation of viral load and phenotypic symptoms were established, which assists in unbiased disease assessment in wheat virus varietal screening nurseries. Overall, the knowledge of the complexity of host-virus interactions, information of genetic variability, the phylogenetic relationship among isolates, and reports of new isolates of viruses and their co-infections will help in recommendation for sustainable management practices for wheat viruses.

Exploring the wheat virome using high-throughput Nanopore sequencing: A metagenomics and phylogenetic analysis

by

Nar Bahadur Ranabhat

B.S., Tribhuvan University, 2006
M.Sc., University of Hohenheim, 2012
M.S., Montana State University, 2017

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Plant Pathology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Jessica L. S. Rupp

Copyright

Abstract

Kansas is one of the top wheat-producing states in the United States. Many wheat viruses have been recognized as common yield-reducing factors. The synergistic impact of several concurrent infecting wheat viruses is responsible for millions of dollars lost in wheat production. Genetic resistance, using virus-resistant cultivars, is one of the key management practices of wheat viruses. The primary threat to durable genetic resistance is the presence of potential new virus variants in the field. The main objective of this study was to explore the field wheat virus population. We used the high-throughput Oxford Nanopore sequencing technique (ONT) to study the wheat virome. A survey was conducted in 2019, 2020, and 2021 in major wheat-growing counties of Kansas, and wheat leaves showing virus-like symptoms were collected. Total RNA was extracted, and cDNA sequencing libraries were made using the PCR-cDNA barcoding kit and loaded into ONT MinION flow cells. Sequencing reads were aligned to cereal virus references. We identified eight wheat viruses belonging to the genera: *Tritimovirus*, *Poacevirus*, *Emaravirus*, *Bromovirus*, *Luteovirus*, *Polerovirus*, *Bymovirus*, and *Furovirus.* We recorded mixed infections of two to five viruses in a single sample. Wheat streak mosaic virus (WSMV) + triticum mosaic virus (TriMV) mixed infection was the most predominant infection (16.7%), followed by WSMV + TriMV + brome mosaic virus (BMV) (11.9%) and WSMV single infection (11.9%). Phylogenetic analysis of the whole genomes of WSMV revealed the wide distribution of isolates into clades and subclades including European isolate. Potential WSMV recombinant isolates were found. BMV was identified for the first time in Kansas wheat. We used genetic and evolutionary approaches to characterize BMV isolates. On average, US BMV isolates showed low divergence. Coding regions of all BMV RNAs were under purifying or negative selection pressure. The whole-genome sequences of multiple isolates of High Plains

wheat mosaic emaravirus and soilborne wheat mosaic virus were characterized. Additionally, virus viability in the inoculum over time was determined, and the relation of viral load and phenotypic symptoms were established, which assists in unbiased disease assessment in wheat virus varietal screening nurseries. Overall, the knowledge of the complexity of host-virus interactions, information of genetic variability, the phylogenetic relationship among isolates, and reports of new isolates of viruses and their co-infections will help in recommendation for sustainable management practices for wheat viruses.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my advisor, Dr. Jessica Rupp, for guiding and supporting me throughout my Ph.D. program. Her vision of taking biochemistry and genetics courses while working as a research assistant helped me to be a successful graduate student. I appreciate her availability, accessibility, and creating a happy and healthy working environment during the entire time. I am grateful to have you as a mentor and your thoughtful guidance has made this work possible.

I would like to extend my appreciation to my committee members. Thank you, Dr. John Fellers, for your guidance, resources, and for making your lab always available for me to conduct nanopore sequencing and bioinformatics. Your guidance and support made this research possible. I would like to thank Dr. Allan Fritz and Dr. Erick DeWolf for being part of my committee and for their thoughtful discussions and valuable guidance. I would like to extend my sincere thanks to Dr. Myron Bruce for being my molecular Guru. His continuous guidance, support, and inspiration helped me to learn molecular-based research. I appreciate his arduous mentorship.

A special thanks to Dylan Mangel for being such a wonderful lab and office mate as well as for our discussions, laughing, and especially walking together throughout this entire Ph.D. life. I also like to thank all our Applied Wheat Pathology lab members for their support, especially Mark Davis, for his support in the field site projects. I would like to thank agronomist Jeanne Falk Jones, K-State extension agents and Kansas wheat growers who assisted in sample collection. I also thank the K-State Plant Disease Diagnostic Lab, Chandler Day, Christian Webb, and Judy O'Mara, for providing virus-infected wheat samples.

The sincerest gratitude goes to my family; my wife Mana Ranabhat, my daughter Sampada Ranabhat, and my son Vivaan Ranabhat for their tireless support, love, and understanding throughout my entire graduate life.

Dedication

For my late parents, especially for my mom Chhabi Maya Ranabhat who taught me "Climb mountains to have better view".

Chapter 1 - Introduction

General background

Agriculture production has been increasing globally, but the food production system will face significant challenges to meet such unprecedented demand. Food security will be the primary concern to feed the ever-increasing human population. It is projected that the human population will reach over 9 billion in 2050, which is close to the maximum carrying capacity of the Earth's food resources (Alexandratos and Bruinsma, 2012). The current agricultural production needs to double to meet this demand while increasing the sustainability of the environment and society (FAO, 2017). Production of major crops will need to increase by enhancing the production potential of crops while minimizing crop yield losses. The potential yield of major crops is significantly affected by diseases, pests, and other abiotic factors (Mylonas et al., 2020). The global estimated average yield loss of major crops at a global level due to the emerging or re-emerging of pests and plant pathogens is: wheat (*Triticum aestivum* L.) 21.5%, rice (*Oryza sativa* L.) 30.0%, maize (*Zea mays* L.) 22.5%, potato (*Solanum tuberosum* L.) 17.2%, and soybean (*Glycine max* L. merr.) 21.4% (Savary et al., 2019).

Wheat plays a vital role in food security. Wheat belongs to the family Poaceae and shared a common ancestor with rice (*Oryza sativa* L.) approximately 40 to 54 million years BCE, with *Brachypodium* 32 - 39 million BCE (Vogel et al., 2010) and it diverged from barley (*Hordeum vulgare* L) approximately 8 - 9 million years ago (Middleton et al., 2014). The center of origin of cultivated wheat is the Fertile Crescent region in the Middle East. The first wheat domestication occurred around 10,000 years ago during the 'Neolithic Agricultural Revolution' as cultivated emmer (Chantret et al., 2005; Shewry, 2009). Hexaploid wheat (AABBDD, 2n = 6x = 42) is derived from hybridization between tetraploid emmer wheat (*Triticum turgidum* ssp *dicoccoide*s,

AABB) and diploid goatgrass (*Aegilops tauchii,* DD) hybridized approximately 7,000 to 9,000 years ago (Brenchley et al., 2012; Chantret et al., 2005). This hybridization might have occurred few times independently as farmers selected hexaploid wheat for its superior qualities (Shewry, 2009). The AA genomes originated from the progenitor genomes of the earliest cultivated einkorn wheat (*Triticum monococcum*, eg, *T. urartu*), BB genome is believed to be originated from *Sitopsis* section of *Aegilops,* which is in the same section as *Aegilops speltoides* (Brenchley et al., 2012; Feldman et al., 2001).

Wheat is an important cereal crop and basic staple food grown in all food production regions of the world. In 2021, world wheat production was 28.6 billion bushels (USDA, Foreign Agriculture Service). In the USA, it is a principal food grain, and a total of 1.65 billion bushels was produced (USDA Crop Production 2021 summary, National Agricultural Statistics Service). There are three major categories of wheat grains: winter wheat, spring wheat, and durum wheat. Among them, winter wheat production represents the most productive of total wheat production, equal to 1.28 billion bushels (USDA Crop Production 2021 summary, National Agricultural Statistics Service). In the United States, three categories of wheat are further classified into five classes: hard red winter wheat, hard red spring wheat, soft red winter wheat, white (winter and spring) wheat, and durum wheat. Hard red winter wheat is primarily grown in the Great Plains and accounts for 40% of the total wheat production. Kansas is one of the top two wheat-producing states in the United States. It produced 281.25 million bushels in 2020 and 364 million bushels in 2021, ranked second and first leading producer respectively (USDA, National Agricultural Statistics Service, 2020 and 2021).

Several abiotic and biotic factors threaten wheat production. The common abiotic stress factors are variable weather and climatic conditions, prolonged heat and cold, drought, salinity,

and nutrient limitations (Halvin et al., 2005; Kajla et al., 2015). Many biotic agents include fungal, bacterial, and viral pathogens, insects, and nematodes (Kashyap et al., 2020). It is difficult to accurately estimate crop loss caused by biotic factors only. Researchers in Kansas estimated the history of cumulative disease yield loss for wheat from 1976 to 2020 ranged from 0.2% to 22.2% of potential yield lost. In 2020 and 2021, the cumulative wheat disease loss of wheat by disease (excluding nematodes) was estimated to be 10.8% and 16.2% or about 31.8 million bushels and 70.4 million bushels respectively (Hollandbeck et al 2020 and 2021).

Viral pathogens are a common threat to wheat production. These viruses are vectored by aphids (e.g., Barley yellow dwarf virus, Cereal yellow dwarf virus), mite (e.g., viruses of wheat streak mosaic complex), Plasmodiophorid, *Polymyxa graminis* (e.g., soilborne wheat viruses), or leafhopper (wheat dwarf virus) (Ordon et al., 2009). In Kansas, the most common wheat viral diseases of wheat are Barley yellow dwarf, the 'wheat steak mosaic' (WSM) complex, and soilborne viruses. Barley yellow dwarf is a disease caused by the infection of isolates of barley yellow dwarf viruses (BYDV) of Genus *Luteovirus* (e.g., barley yellow dwarf virus MAV, barley yellow dwarf virus PAS, barley yellow dwarf virus PAV), and Genus *Sobemovirus* (e.g., barley yellow dwarf virus GPV and barley yellow dwarf virus SGV) according to efficiency of aphid transmission (Virus Taxonomy 2020 Release). WSM disease can be caused by up to three distinct viral pathogens; wheat streak mosaic virus (WSMV), triticum mosaic virus (TriMV), and High Plains wheat mosaic emaravirus (HPWMoV). All three viruses of the WSM complex are transmitted by an eriophyid mite, the wheat curl mite (*Aceria tosichella* Keifer) within and between the fields (Mahmood et al., 1998; Seifers et al., 2009, 1997). In the Great Plains, WSM causes approximately 2% estimated loss over 5 years average, however, much greater losses in localized fields is common especially during epidemic years (Burrows et al. 2009; Ranabhat, N.

personal observation 2019, Figure 1.1). Annual yield loss due to WSM is variable, in 2017 Kansas wheat producers lost about 5.6% of total yield (19.2 million bushels of wheat) worth $76.8 million (KSwheat 2017, Hollandbeck et al 2017). The typical symptom of WSM is a yellow mosaic or streaking on infected leaves (Rahman et al. 1974). Other symptoms include stunted growth, reduced root biomass, low water use efficiency, low seed test weights, poor tillering, and yield loss (Murray et al., 2005; Price et al., 2010b). Generally, early fall infection before tillering stages has a more severe impact on yield than infection in the later growing stages (Hunger et al., 1992; Thomas and Hein, 2003). In the field, the symptoms appear first on the edge of the plot as the vector (wheat curl mite, WCM) moves with the virus from adjacent infected alternative hosts or wheat fields passively with wind (Slykhuis, 1955).

The WCM is an eriophyid spindle-shaped mite with two pairs of legs (Keifer, 1969). Feeding of WCM causes curling of the leaf edges and creates a protective microclimate for its growth and reproduction (Slykhuis, 1955; Thomas et al., 2004). The significant yield loss associated with WCM is the virus infection, as the feeding of aviruliferous mites has a minor impact on yield (Harvey et al., 1999). WCM transmits WSMV in a semi-persistent manner (Seifers et al., 2006; Siriwetwiwat, 2006). WSMV transmission efficiency varies by the stage of the mite; the nymphal stages are more efficient at transmitting than adults (Siriwetwiwat, 2006).

Common wheat viruses

Wheat streak mosaic virus (WSMV)

Wheat streak mosaic virus was first discovered in 1922 (McKinney, 1937) and became one of the most common viral pathogens of wheat. The virus is widely distributed throughout the globe (Ellis et al., 2003; Kapooria and Ndunguru, 2004; Navia et al., 2013). While wheat is the primary host of WSMV, it can infect a variety of crops including corn (*Zea mays* L.), barley

(*Hordeum vulgare* L.), rye (*Secale ceareal* L.), oat (*Avena sativa* L.), pearl millet (*Pennisetum glucum* (L.) R. Br.), and foxtail millet (*Setaria italic* (L.) P. Beauvois). Additionally, WSMV infects several grass species including downy brome (*Bromus tectorum* L.), rye brome (*B. secalinus* L.) green foxtail (*Setaria viridis* (L). Beauv), jointed goat grass (*Aegilops cylindrica* Host), and barnyard grass (*Echinochloa crus-gali* (L.) Beauv.) (Brey et al., 1998; Ito et al., 2012; Somsen et al., 1970). These grass species can serve as alternative hosts and act as a "green bridge" in the absence of wheat in the fields and as a source of infection for the following wheat growing season (Brey et al., 1998; Jiang et al., 2005; Somsen et al., 1970).

WSMV (genomic structure)

WSMV is the type species of the genus *Tritimovirus* within the family *Potyviridae* (Stenger et al., 1998). WSMV is a single-stranded, monopartite, flexuous rod-shaped, positive-sense RNA virus with a genome consisting of 9384 nucleotides and a single large open reading frame (ORF) (Stenger et al., 1998). The WSMV genome has 130 and 149 nucleotide (nt) untranslated regions (UTR) on 5' and 3' respectively, and end of the RNA the 3' end comprises of a poly-A tail whereas VPg (virus protein linked to the genome) is covalently attached to the 5' end (Stenger et al., 1998). The RNA genome is transcribed into a large polyprotein (~350kDa) of 3,035 amino acids and the protein undergoes self-cleavage by three encoded proteinases into 11 mature proteins (Tatineni and Hein, 2018). They are protein P1, helper component protease (HC-Pro), protein P3, 6K1, 6K2, cytoplasmic inclusion protein (CI), nuclear inclusion putative protease (NIa-pro), viral protein genome-linked proteinase (VPg), nuclear inclusion putative polymerase (NIb), and coat protein (CP) (Figure 1.2) (Choi et al., 2001a; Chung et al., 2008a; Stenger et al., 1998). Similar to other potyviruses, WSMV also has a conserved feature in the genome, a short ORF, pretty interesting potyvirus ORF (PIPO) embedded within the N-terminal

half of P3 encoding region in a reading frame different from the polyprotein which expresses P3N-PIPO protein using +2 frameshift (Chung et al., 2008a).

The P1 protein (40 kDa) is a serine protease that cleaves itself at its carboxy terminal and is an RNA silencing suppressor and enhances disease symptoms (Young et al. 2012). Gupta and Tatineni 2019 found that WSMV P1 binds and protects dsRNAs from the hydrolytic activity of host Dicer family proteins and plays a significant role in suppressing RNA silencing. In detail, they reported that the C-terminal region of P1 is responsible for the RNA silencing activity as the deletion of single amino acid from this region completely abolished the silencing function. In particular, the disruption of the P1 GW motif AGO-binding linear peptide resulted in the loss of silencing suppression function and revoked WSMV viability (Gupta and Tatineni 2019).

The HC-Pro in potyviruses is multifunctional. It is responsible for vector transmission, long-distance movement, polyprotein maturation, enhancement of virus particles yield and suppression of post-transcriptional gene silencing (PTGS) (Anandalakshmi et al., 1998; Blanc et al., 1998; Carrington et al., 1996; Revers and García, 2015). HC-Pro has three structural domains: the N-terminal, central portion, and C-terminal. Most of the HC-Pro functions are based on the central region. The C-terminal region is responsible for the proteolytic activity and the N-terminal region is responsible for facilitating vector transmission. HC-Pro provides a link between virus particles and vector stylets through interaction with the CP DAG motif and HC-Pro PTK motif (Blanc et al., 1998). This mechanism of HC-Pro acts as a reversible link between the viral particle (CP) and vector mouthparts called the “bridge hypothesis” (Valli et al., 2007).

WSMV HC-Pro (44 kDa) is required for WCM vector transmission and has cysteine proteinase activity (Stenger et al., 2005; Young et al., 2007). Deletion and mutation analysis of WSMV HC-Pro revealed no effect on WSMV virulence in wheat (Stenger et al., 2005); it is

dispensable for systemic movement but plays a role in replication (Stenger, et al. 2006a). According to Stenger, et al. 2006b, a series of 5' proximal nested deletions in the HC-Pro coding regions (1152 nt), a deletion of as few as 24 nt (codons 3-10) at the 5' end completely abolished transmission by WCM. Additionally, alanine substitution of cysteine residues at amino acid positions 16, 46, or 49 also eliminated vector transmission (Young et al., 2007) and emphasized the essential role of HC-Pro in vector transmission.

In potyviruses, P3 interacts with CI, NIb, and NIa (Revers and García, 2015), but the essential role of P3 is still obscure. It is found that P3 is required for viral replication and pathogenicity (Klein et al., 1994). P3N-PIPO in WSMV and in other potyviruses is required for virus cell-to-cell movement, amplification of the viral genome, virion assembly, and suppression of RNA silencing (Chung et al. 2008; Wei et al. 2010a). Among WSMV isolates, P3N-PIPO is highly conserved, and any mutation severely reduced WSMV cell to cell movement (Choi et al., 2005; Chung et al., 2008b).

Out of the 11 proteins, 6K1 and 6K (6 kDa) are the smallest. The function of these proteins is characterized in some potyviruses. Cui and Wang 2016 found that 6K1 is responsible for viral replication and co-involved with 6K2 to form viral replication vesicles. 6K2 is an integral membrane protein and induces the endoplasmic reticulum originated replication vesicles, therefore playing an essential role in forming the potyviral replication complex (Wei et al., 2010).

CI (73 kDa) is the largest protein among potyvirus-encoded proteins. It forms the characteristic pinwheel cylindrical inclusions in the cytoplasm of infected cells (Edwardson et al., 1984). CI interacts with other proteins and has ATPase and RNA helicase activities, playing a significant role in viral replication (Fernández et al., 1997). Along with P3N-PIPO, CI assists

in viral cell to cell movement (Wei et al., 2010). CI also acts as a virulence factor and interacts with different host factors (Revers and García, 2015).

NIa (49 kDa) is the major potyvirus protease of the potyviruses responsible for the proteolytic cleavage of the polyprotein (Carrington and Dougherty, 1988). NIa has two domains, the N-terminal VPg domain (NIa-VPg) and the C-terminal proteinase domain (NIa-pro). The free VPg protein is covalently attached to the 5' end of the viral genome, impersonating the a methyl-7-G cap of the mRNA, and interacts with several host factors including cap-binding proteins leading to the initiation of translation of the viral genome (Khan et al., 2008). NIa-pro is responsible for proteolytic processing and cleavage efficiency that plays an important role in controlling viral infection (Revers and García, 2015) and also form nuclear inclusions.

NIb (57 kDa) is the RNA-dependent RNA polymerase responsible for genome replication in potyviruses. It is reported that NIb interacts with the 6K-NIa-VPg or NIa-pro protein complex and is targeted to the replication vesicles where the viral RNA replication takes place (Dufresne et al., 2008; Fellers et al., 1998; Leonard et al., 2004; Li et al., 1997). By containing two independent nuclear localization signals, NIb also has nuclear translocation activities. The nucleocytoplasmic transport of protein complex (NIb/SCE1 or SUMOylated form of NIb) helps regulate NIb activity and generates a favorable environment for virus multiplication (Li et al., 1997; Xiong and Wang, 2013).

Protection of the viral genome through encapsidation is the primary function of the CP. In addition to virion assembly and disassembly, CP in potyviruses is multifunctional and performs several nonstructural functions including replication, vector transmission, virus translocation (Bol, 2005; Callaway et al., 2001; Dreher and Miller, 2006), and suppression of host RNA silencing (Qu et al., 2003; Wang and Metzlaff, 2005). CP consists of three regions, N-

terminal, central, and C-terminal domains. The central core domain is highly conserved and responsible for cell to cell movement and virion assembly (Dolja et al., 1994, 1995). The N and C termini are variable and exposed on the virion surface.

In WSMV, CP (37 kDa) is 349 amino acids long. The CP of WSMV is responsible for the virus transmission with its vector WCM, cell to cell movement, and pathogenicity (Tatineni and French, 2014). WSMV CP shows variation and tolerates deletions at the N terminal region. Tatineni et al. 2014a found that deleting of amino acids 6 to 27, 36 to 100 at N-terminus, and 65 at the C-terminal ends did not affect systemic infection. However, N-terminal amino acids 6 to 27 and 85 to 100 are still required for efficient virion assembly and cell-to-cell movement. The C-terminal 65 amino acids are required for cell-to-cell movement, but not virion assembly.

Interestingly, a WSMV mutant lacking part of the N-terminal amino acid 58 to 84, but not 36 to 57 induced more severe symptoms than wild type (Tatineni et al., 2017). In that study, authors reported that the deletion of amino acids from 58 to 84 accelerates cell-to-cell movement, enhances the accumulation of CP and genomic RNA, and alters CP-specific protein profiles in multiple hosts, including wheat, barley, maize, and rye. CP amino acids 58-100 are required for WCM transmission of WSMV and the aspartic acid residue at the C-terminal region of CP determines WCM transmission, as the mutation of aspartic acid residues at 289 or 326 (D289A or D326A) significantly reduces the mite transmission (Tatineni et al., 2018)

WSMV phylogeny

WSMV has been collected from around the world, sequenced, and the phylogenetic relationship was determined. WSMV isolates were grouped into four distinct clades based on the sequence of the coat protein genes. El Batán isolates from Mexico represent clade A (Sánchez-Sánchez et al., 2001), isolates from Europe and Russia comprise clade B, isolates from Iran

represent clade C, and several isolates from North America and Turkey constitute clade D (Rabenstein et al., 2002; Stenger et al., 2002). Clade B is characterized by the deletion of 3 nucleotides at position 8412 to 8414 resulting in the deletion of glycine residue Gly_{2761}, Clade B is referred to WSMV-ΔE (Gadiou et al., 2009). Clade B share 97.5% to 100% nucleotide sequence identity. The pair-wise nucleotide divergence between Clade A (El Batán) and Clade D representative Sidney 81 is 20%. Clade D undergoes substantial and recent divergence with most consensus sequence substitutions (Stenger et al., 2002).

Choi et al. 2001b compared the genome sequences of three WSMV isolates, two from Clade D and one from Clade A. The WSMV type strain and Sidney 81 strain (Clade D) has 97.6% (nucleotide) and 98.7% (amino acid) sequence identities; however, El Batán (Clade A) shared only ~79 (nucleotide) and ~90% (amino acid) sequence identity with type strain and Sidney 81. In that study, they mentioned that El Batán had 15 fewer amino acid residues in the coat protein and constituted the second WSMV population of North America.

Clade D isolates are a large polytomy with many branches at the basal node and further divided into sub-clades. WSMV isolates from Kansas are highly diverse falling into many sub-clades of the Clade D. Sub-clade D1 contains isolates from American Pacific Northwest (APNW), Kansas isolate, and Colorado isolate constitute D2. Isolates from Kentucky, Ohio, Missouri, and also Kansas constitute D3, and D4 contains isolates from Nebraska and Kansas (Stenger et al., 2002). The WSMV isolates reported from Australia and Argentina are similar to sub-clade D1, APNW (Dwyer et al., 2007).

WSMV field populations are complex and diverse, facilitated by distinct genetic isolation mechanisms (McNeil et al., 1996). Hall et al. 2001 described three distinct genetic isolation mechanisms; cross-protection among closely related strains, the spatial distribution of WSMV

strains in the co-infected plant, and vector transmission bottlenecks. Additionally, diversity in the WSMV field population is also driven by interactions with host and vector (acquisition and transmission), by selection for replication fitness of the virus, adaptation to various alternative host species (McNeil et al., 1996), and systemic movement bottlenecks.

Triticum mosaic virus (TriMV)

TriMV was first identified first time in Kansas infecting the WSMV-resistant wheat cultivar 'RonL' in 2006 (Seifers et al., 2008). TriMV is a member of the *Potyviridae* family and the type member of the genus *Poacevirus* (Seifers et al., 2008; Tatineni et al., 2009). This virus is transmitted mechanically and vectored by WCM individually or together with WSMV and HPWMoV (Seifers et al., 2009). TriMV possesses a single-stranded, positive-sense RNA genome with a polyprotein consisting of 3112 amino acids, a similar genomic organization to WSMV (Fellers et al., 2009; Tatineni et al., 2009). TriMV showed 46% CP amino acid sequence identity in coat protein to a closely related member of *Poacevirus,* the Sugarcane streak mosaic virus, and 23.2% identity with WSMV (Fellers et al., 2009; Tatineni et al., 2009). The distinct genomic organization of the TriMV genome from other potyviruses is a long 5' untranslated region of 738 nt with a total genome length of 10,266 nucleotides (Fellers et al., 2009). Similar to WSMV, polyprotein of TriMV also consists of proteins P1, HC-Pro, P3 with PIPO, 6K1, CI, 6K2, NIa, NIb, and CP (Fellers et al., 2009). TriMV is most commonly co-infected with WSMV in the Great Plains wheat fields (Byamukama et al., 2014) which induces synergism that exacerbates the symptoms, viral titer, and yield loss (Byamukama et al., 2014; Tatineni et al., 2014a).

High Plains wheat mosaic emaravirus (HPWMoV)

HPWMoV was first identified in 1993 as the causative agent of wheat and maize High Plains disease with severe mosaic and necrosis symptoms (Jensen et al., 1996). Later, researchers assigned different names such as maize red stripe virus, wheat mosaic virus, and High Plains wheat mosaic virus (Skare et al., 2006; Tatineni et al., 2014b). The International Committee on Taxonomy of Viruses (ICTV) renamed this virus as High Plains wheat mosaic emaravirus under the family *Fimoviridae* and genus *Emaravirus* in 2016 (https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode_id=202000012). As HPWMoV is a member of the WSM complex, it is also vectored by WCM and often found co-infected with WSMV and TriMV (Burrows et al., 2009; Byamukama et al., 2013).

HPWMoV genome is composed of eight negative-sense single-stranded RNA segments encoding a single ORF in each RNA segment (Tatineni et al., 2014b; Tatineni and Hein, 2021). RNA1 encodes a 2272 amino acid RNA-dependent RNA polymerase (RdRp), RNA2 encodes a 667 amino acid glycoprotein precursor, RNA3 encodes the nucleoprotein (NC) (Tatineni et al., 2014b). RNA3 contains two distinct variants, 3A (286 amino acids) and 3B (289 amino acids), with 12.5% sequence divergence (Stewart, 2016; Tatineni et al., 2014b). These two variants of RNA3 shared 95 - 99% within-group and 88 - 89% between-group protein identity (Stewart, 2016). RNA4 encodes a polyprotein with 364 amino acids predicted to be involved in virus movement. RNA5 and RNA6 each encodes a protein of unknown function of size 478 and 492 amino acids, respectively (Stewart, 2016; Tatineni et al., 2014b; Tatineni and Hein, 2021). The proteins encoded by RNA7 (305 aa) and RNA8 (176 aa) has been reported as RNA silencing suppressors (Gupta et al. 2018, 2019).

Management of WSM complex

The management of WSM is based mainly on the integration of cultural practices and genetic resistance. Cultural practices are general preventive measures and cannot provide sufficient protection once the virus infects the wheat field. These practices includes late planting of wheat, modification of nitrogen application in the field, and management of alternative hosts including volunteer wheat of virus and its WCM vector (Hadi et al., 2011; Miller et al., 2015). There is no effective chemical option available for the control of the WCM (Murphy, 2016; Velandia et al., 2010). Therefore, limiting the vector spread and/or reducing the source of inoculum may be done by controlling alternative hosts that act as reservoirs of WSM viruses and WCM.

Alternative hosts including volunteer wheat and grassy weeds, act as a "green bridge" to spread WCM and viruses from one growing season to the next (Jiang et al., 2005). Viruliferous mites move from infected wheat fields to alternative hosts before and during senescence. After the emergence of fall planted winter wheat, WCM infected with WSM viruses can move from alternative hosts to the newly emerged wheat. Warm fall, cool and wet summer, and the presence of alternative hosts including volunteer wheat and downy brome increase survival and reproduction of WCM and increase the risk of WCM infestation and virus infection (Jiang et al., 2005; Ranabhat et al., 2018; Singh et al., 2018).

Host resistance is an effective, efficient, and key component for the sustainable management of WSMV. Three resistance genes, *Wsm1*, *Wsm2*, and *Wsm3* have been identified (Fahim et al., 2012; Friebe et al., 2009; Haley et al., 2002; Lu et al., 2011a). *Wsm1* was transferred to wheat from intermediate wheatgrass *Thinopyrum intermedium* (Host) Barkworth & D.R. Dewey (Friebe et al., 1991) and introduced into cultivar 'Mace' (Graybosch et al., 2009).

The source of *Wsm2* is unknown and present in the germplasm line CO960293-2 (Haley et al., 2002). It has been incorporated into several cultivars, including 'RonL' (Seifers et al., 2007a), 'Snowmass' (Haley et al., 2011), 'Clara CL' (Martin et al., 2014), and 'Joe' (Zhang et al., 2016). However, these resistance genes are temperature sensitive, where resistance is less effective at higher temperatures above 20°C (Seifers et al., 2006, 2007a). *Wsm3* was derived by a Robertsonian translocation from *T. intermedium* chromosome arms 7S#3L to the short arm of chromosome 7B resulting in the T7BS.7S#3L translocation chromosome. *Wsm3* confers resistance to WSMV and TriMV at temperature as high as 24º C (Danilova et al., 2017; Liu et al., 2011).

Other major wheat viruses include Cereal yellow dwarf virus (CYDV), Wheat spindle streak mosaic virus (WSSMV), Wheat yellow mosaic virus, and Soilborne wheat mosaic virus (SBWMV) (Brakke, 1987; Hodge et al., 2020; Rotenberg et al., 2016). These viral pathogens reduce wheat production qualitatively, as well as quantitatively (Byamukama et al., 2014; Campbell et al., 1975; Choudhury et al., 2019; Cunfer et al., 1988). Hodge et al. 2020 also reported brome mosaic virus and Cockfoot mottle virus from Ohio wheat samples.

Brome mosaic virus (BMV)

Brome mosaic virus (Genus: *Bromovirus*, family: *Bromoviridae*) is the type member of a group of icosahedral, positive-strand ssRNA viruses with a tripartite linear genome. The genome consists of RNA1, RNA2, and RNA3 (Ahlquist et al., 1984; Kao and Sivakumaran, 2000). RNA1 and RNA2 are encapsidated separately in the icosahedral virion, while RNA3 is encapsidated separately with an additional sub-genomic RNA (Rao, 2006). RNA1 encodes protein 1a, which has capping and RNA helicase activities, RNA2 encodes protein 2a, a putative RNA-dependent RNA polymerase, RNA3 encodes two proteins: 3a movement protein, MP and

coat protein, CP. CP is translated or derived from RNA3 as sub-genomic RNA and referred as RNA4 (Kao and Sivakumaran 2000; Rao 2006). The capsid of all three particles contains 180 subunits of the CP arranged in icosahedral symmetry (Lucas et al., 2002).

BMV was studied extensively as a model for RNA virus biology and in recombinant DNA technology (He et al., 2021; Kao and Sivakumaran, 2000). Only a few studies were conducted to evaluate the BMV incidence on economically important crops. A 13% incidence of BMV was reported from wheat fields in Escambia county of Alabama in 2004 (Srivatsavai, 2005). BMV was detected with a high prevalence in soft red winter wheat and showed a potentially high risk to wheat production with up to 61% yield loss on soft red winter wheat when inoculated at early growth stages (Hodge et al., 2019(Hodge et al., 2020).

Nanopore sequencing

The idea of sequencing a single strand of DNA by altering the ionic current was first described by a professor from University of California Davis, David Deamer in 1989 in his notebook and later in collaboration with George Church and Daniel Branton (both at Harvard University). In 1993 Deamer, Daniel Branton, and Kasiannowicz employed α-hemolysin (α-HL), a pore-forming protein secreted by a Staphylococcus bacteria and in 1996 their results of DNA translocation through α-HL nanopore was published (Deamer et al., 2016; Kasianowicz et al., 1996). In 2012, Oxford Nanopore presented the first-ever Nanopore sequencing data. In 2014 the MinION as MinION Access Program (MAP) became available to the public with the invention of new robust membrane and data analysis methods to enable basecalling (Heger, 2014; Wang et al., 2015). Currently, Oxford Nanopore sequencing Technologies (ONT) has been an excellent method for generating high throughput long reads of both DNA and RNA in real-time with no prior amplification (Ayub et al., 2013; Kasianowicz et al., 1996).

The portability of MinION and improvements in read accuracy make ONT a cost-effective option for both high and low-scale diagnostics tools. The application of ONT for plant virus detection is rapidly progressing (Liefting et al., 2021; Mehetre et al., 2021). Recently, ONT has been used for the detection and genomic analysis of causative agents of plant viral diseases of various crops. For examples, Cassava mosaic virus (Boykin et al. 2018), Wheat streak mosaic virus (J. Fellers et al., 2019), Potato virus Y (Della Bartola et al., 2020), Sowthistle yellow vein virus (Stenger et al., 2020), Cowpea bright yellow mosaic virus (Naito et al., 2019), Plum pox virus from prunus plant (Bronzato Badial et al., 2018), Tomato yellow leaf curl virus, Watermelon chlorotic stunt virus and Tomato brown rugose fruit virus (Chalupowicz et al., 2019); Cucumber green mottle mosaic virus in tomato, Zucchini yellow mosaic virus in Butter squash, and three viruses including Dioscorea bacilliform virus, Yam mild mosaic virus and Yam chlorotic necrosis virus from Water yam plant (Filloux et al., 2018). Additionally, first report of Arabis mosaic virus in potato (Monger et al., 2020) and Tomato severe rugose virus in Tomato weed (*Physalis angulata*) as well as complete genome sequencing of Sri Lankan cassava mosaic virus (Leiva et al., 2020) were performed using ONT.

A significant issue with ONT is that it is more error-prone in single-base accuracy than the current high throughput sequencing methods including Illumina (Chen et al., 2020; Lebrigand et al., 2020). It is improving through the use of downstream high accuracy basecalling and polishing software (Cao et al., 2019). The single-read accuracy is up to 98% for the current MinION device44 (Jain et al., 2017) which is often sufficient for the identification of pathogens (Jain et al., 2016). ONT continues to improve with high accuracy base-calling software, for example, for *E. coli* where the single read consensus accuracy improved to 99.5% at 30X coverage (Loman et al., 2015). Currently, MinION achieves read length of 50kb with about 92 -

95% single read accuracy, and basecalling and polishing software are likely to continue to improve and consensus accuracy recorded up to 99.9% (Chang et al., 2020) and (https://nanoporetech.com/about-us/news/r103-newest-nanopore-high-accuracy-nanopore-sequencing-now-available-store).

Mechanical inoculation and phenotypic rating

Mechanical inoculation is a common practice for plants screened for breeding programs to test virus resistance. Biological vector inoculation is rarely applied because of the difficulty of maintaining vector colonies, uniform inoculation to all plants/field plots, as well as low vector survival after inoculation (Lu et al., 2011a; Wosula et al., 2018). Successful mechanical inoculation is essential to evaluate the level of resistance in breeding lines. Success largely depends on the infectivity of the virus particle, and the effectiveness of the inoculum depends on both concentration and viability

Historically, the resistance level of cultivars inoculated with the viral pathogens is determined based on phenotypic symptom assessment (DeWolf et al. 2019; Johnson et al. 2019; Marburger et al. 2018; Rupp 2015). Additionally, the phenotypic rating scale is often inconsistent, which hinders the scientific comparison on a larger scale. The commonly used scale is 1 to 9, resistance to susceptible (DeWolf et al. 2019; Johnson et al. 2019; Bockus et al. 2011), but other categorical scales, such as susceptible, moderately susceptible, moderately resistant, resistant have also been used (Guillen-Portal et al. 2021; Kleinjan et al. 2021). Reproducible inoculation of the virus by mechanical methods and measurement of viral load in initial inoculation and *in planta* will provide an accurate assessment as the level of symptoms does not always correspond to the level of virus titer (Ranieri et al., 1993; Roossinck, 2012).

Significance of the study and objectives

Plant viruses cause severe economic losses by reducing yield, negatively affecting quality, and continuously threatening sustainable agriculture. Due to rapid symptom development and a short life cycle, plant viruses cause a highly negative impact on crop production (Mehetre et al., 2021). Multiple virus infections in a single plant are common. Frequent monitoring of viral pathogen and accurate diagnosis of field virus diversity is essential to design management strategies. There is a great demand for accurate new techniques to identify multiple virus infection in a single plant. Currently, identification and detection rely on target-specific tests, based on either serological assays such as ELISA, or molecular approaches such as PCR, nucleic acid spot hybridization methods, or microarray techniques (Boonham et al., 2007; Gibbs and Mackenzie, 1997; James et al., 2006). Sometimes sophisticated techniques, including electron microscopy and a combination of ELISA and electron microscopy, can also be used to confirm or as investigational methods (Mumford et al., 2006).

These contemporary methods have some shortcomings. First, most of these methods require prior knowledge of the genome to use specific primers, probes, or antibodies to identify the pathogen. These primers or probes are specific to a particular strain, individual species, or small group (Chalupowicz et al., 2019). These methods follow high workload assays to predict potential pathogens infecting a new host or previously uncharacterized agents. Second, nonspecific methods such as electron microscopy are costly, time-consuming, and require a high degree of expertise, and this method only provides information on the pathogen presence. Further analysis is then required for the characterization of the pathogen. Identifying and characterizing novel viral pathogens through conventional molecular and diagnostic methods

takes several months (Susi, 2004). Therefore, a promising approach to address these shortcomings is the use of high throughput sequencing including third-generation sequencing.

ONT is a diagnostic method that does not require direct prior knowledge of the pathogen, discovers novel viruses, and sequences the whole genome for the in-depth study of the pathogens. ONT is also used as an on-site diagnostic tool. The cassava virus action project used a pocket-sized portable field diagnostic system with MinION and MinIT mobile sequencing devices and successfully diagnosed of the virus on site (Boykin et al., 2019).

For wheat virus management, cultural preventive measures, the key management practice, along with the good cultural practice, is the application of genetic resistance. However, there are a few challenges for managing wheat viruses through genetic resistance. This dissertation has the following objectives to address the challenges of wheat virus management through genetic resistance.

First, a threat to the durability of resistant cultivars is due to the presence of potential new variants or new combinations of the viruses in the field population. Mixed infections of multiple viruses are common in the natural and agricultural production systems, and disease severity in a wheat plant is generally increased (Syller 2012; Sanfaçon 2017; Tatineni et al. 2021). Unfortunately, it is hard to distinguish between single and mixed infection phenotypically. The synergistic interactions of multiple virus infections are common in wheat, leading to significant crop production losses even in moderately resistant cultivars. Therefore, the wheat virome study by using ONT for the identification of diversity of viruses (viral population) and characterization of new isolates or new virus combinations was one of the primary objectives of this study.

Second, genetic, and evolutionary characterization of the viral field population provides in-depth information about pathogens. Pathogens having high evolutionary potential are most

likely capable of overcoming genetic resistance (McDonald and Linde 2002). Using resistant wheat cultivars is the most promising strategy for wheat virus management. The knowledge of molecular characterization, the phylogenetic relationship among virus isolates, and the study of the population genetic parameters of naturally occurring virus isolates would help to design an effective breeding program for durable resistance. Therefore, establishing phylogenetic relationships and population genetic structures of known and unknown wheat viruses was the second objective of this study.

Third, exploring the viability of virus inoculum in mechanical inoculation methods to evaluate resistance in breeding lines and establishing the relationship between virus load and phenotypic expression is essential for the successful and effective screening of wheat cultivar resistance against viruses at the field level. Accurate quantification of virus load will provide an unbiased plant-virus interaction evaluation for breeding lines with no or mild symptoms but with higher virus titer as these lines can spread the virus in the fields. Therefore, the third objective of this study was to determine the viability and stability of the Wheat streak mosaic virus in inoculum with time and to establish the relation of viral load to phenotypic symptom expression. The specific objectives were:

- Implement high throughput Oxford Nanopore sequencing technique for wheat virome detection and diagnosis
- Establish phylogenetic relations of common wheat virus isolates based on whole genome sequencing
- Characterize whole genome sequences of previously uncharacterized Kansas wheat viruses

- Characterize brome mosaic virus isolate and its association with other wheat viruses in Kansas
- Establish evolutionary relationships and determine population characteristics of brome mosaic virus
- Determine stability and viability of wheat streak mosaic virus in inoculum over time
- Quantify accurate wheat streak mosaic virus titer in the inoculum and virus-infected leaves
- Establish the relationship between virus titer and phenotypic ranting scale

References

Ahlquist, P., French, R., Janda, M., and Loesch-Fries, L. S. 1984. Multicomponent RNA plant virus infection derived from cloned viral cDNA. Proceedings of the national academy of sciences. **81**:7066–7070.

Alexandratos, N., and Bruinsma, J. 2012. World agriculture towards 2030/2050: the 2012 revision.

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., and Vence, V. B. 1998. A viral suppressor of gene silencing in plants. Proceedings of the National Academy of Sciences. **95**:13079–13084.

Appel, J., Dewolf, E., Todd, T., and Bockus, W. 2015. Preliminary 2015 Kansas Wheat Disease Loss Estimates; Kansas Cooperative Plant Disease Survey Report; Kansas Department of Agriculture: Manhattan, KS, USA, 2015.

Ayub, M., Hardwick, S. W., Luisi, B. F., and Bayley, H. 2013. Nanopore-based identification of individual nucleotides for direct RNA sequencing. Nano letters. **13**:6144–6150.

Blanc, S., Ammar, E., Garcia-Lampasona, S., Dolja, V., Llave, C., Baker, J., and Pirone, T. P. 1998. Mutations in the potyvirus helper component protein: effects on interactions with virions and aphid stylets. Journal of General Virology. **79**:3119–3122.

Bockus, W. W., Wolf, E. D. D., Gill, B. S., Jardine, D. J., Stack, J. P., Bowden, R. L., Fritz, A. K., and Martin, T. J. 2011. Historical durability of resistance to wheat diseases in Kansas. Plant Health Progress. **12**:25.

Bol, J. F. 2005. Replication of Alfamo-and Ilarviruses: role of the coat protein. Annual Review of Phytopathology. **43**:39–62.

Boonham, N., Tomlinson, J., and Mumford, R. 2007. Microarrays for rapid identification of plant viruses. Annual Review of Phytopathology. **45**:307–328.

Boykin, L. M., Sseruwagi, P., Alicai, T., Ateka, E., Mohammed, I. U., Stanton, J.A. L., Kayuki, C., Mark, D., Fute, T., Erasto, J. and Bachwenkizi, H. 2019. Tree lab: Portable genomics for early detection of plant viruses and pests in sub-Saharan Africa. Genes. **10**:632.

Brakke, M. 1987. Virus diseases of wheat. Wheat and wheat improvement. **13**:585–624.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L., D'Amore, R., Allen, A. M., McKenzie, N., Kramer, M., Kerhornou, A., Bolser, D., and Kay, S. 2012. Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature. **491**:705–710.

Brey, C. W., Johnson, G. D., and Blodgett, S. L. 1998. Survey of Montana Grasses for Wheat Curl Mite (Acari: Eriophyidae), the Vector of Wheat Streak Mosaic Virus. Journal of Agricultural Entomology. **15**:173–181.

Bronzato Badial, A., Sherman, D., Stone, A., Gopakumar, A., Wilson, V., Schneider, W., and King, J. 2018. Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. Plant Disease. **102**:1648–1652.

Burrows, M., Franc, G., Rush, C., Blunt, T., Ito, D., Kinzer, K., Olson, J., O'Mara, J., Price, J., Tande, C., Ziems, A, and Stack, J. 2009. Occurrence of viruses in wheat in the Great Plains region, 2008. Plant Health Progress. **10**:14.

Byamukama, E., Seifers, D., Hein, G., De Wolf, E., Tisserat, N., Langham, M. A. C., Osborne, L. E., Timmerman, A., and Wegulo, S. N. 2013. Occurrence and distribution of Triticum mosaic virus in the central Great Plains. Plant Disease. **97**:21–29.

Byamukama, E., Wegulo, S., Tatineni, S., Hein, G., Graybosch, R., Baenziger, P. S., and French, R. 2014. Quantification of yield loss caused by Triticum mosaic virus and Wheat streak mosaic virus in winter wheat under field conditions. Plant Disease. **98**:127–133

Callaway, A., Giesman-Cookmeyer, D., Gillock, E., Sit, T., and Lommel, S. 2001. The multifunctional capsid proteins of plant RNA viruses. Annual review of phytopathology. **39**:419–460.

Campbell, L., Heyne, E., Gronau, D., and Niblett, C. 1975. Effect of soilborne wheat mosaic virus on wheat yield. Plant Disease Reporter. **59**:472–476.

Cao, Y., Li, J., Chu, X., Liu, H., Liu, W., and Liu, D. 2019. Nanopore sequencing: A rapid solution for infectious disease epidemics. Science China Life Sciences. **62**:1101–1103.

Carrington, J. C., and Dougherty, W. G. 1988. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proceedings of the National Academy of Sciences. **85**:3391–3395.

Carrington, J. C., Kasschau, K. D., Mahajan, S. K., and Schaad, M. C. 1996. Cell-to-cell and long-distance transport of viruses in plants. The Plant Cell. **8**:1669.

Chalupowicz, L., Dombrovsky, A., Gaba, V., Luria, N., Reuven, M., Beerman, A., Lachman, O., Dror, O., Nissan, G. 2019. Diagnosis of plant diseases using the Nanopore sequencing platform. Plant Pathology. **68**:229–238.

Chang, J. J. M., Ip, Y. C. A., Bauman, A. G., and Huang, D. 2020. MinION-in-ARMS: nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from autonomous reef monitoring structures. Frontiers in Marine Science. **7**:448.

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., and Gautier, M. F. 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (Triticum and Aegilops). The Plant Cell. **17**:1033–1045.

Chen, Z., Erickson, D. L., and Meng, J. 2020. Benchmarking hybrid assembly approaches for genomic analyses of bacterial pathogens using Illumina and Oxford Nanopore sequencing. BMC genomics. **21**:1–21

Choi, I. R., Hall, J., Henry, M., Zhang, L., Hein, G., French, R., and Stenger, D. C. 2001a. Contributions of genetic drift and negative selection on the evolution of three strains of wheat streak mosaic tritimovirus. Archives of virology. **146**:619–628.

Choi, I. R., Hall, J., Henry, M., Zhang, L., Hein, G., French, R., and Stenger, D. C. 2001b. Contributions of genetic drift and negative selection on the evolution of three strains of wheat streak mosaic tritimovirus. Archives of virology. **146**:619–628.

Choi, I. R., Horken, K. M., Stenger, D. C., and French, R. 2005. An internal RNA element in the P3 cistron of Wheat streak mosaic virus revealed by synonymous mutations that affect both movement and replication. Journal of general virology. **86**:2605–2614.

Choudhury, S., Larkin, P., Meinke, H., Hasanuzzaman, M., Johnson, P., and Zhou, M. 2019. Barley yellow dwarf virus infection affects physiology, morphology, grain yield and flour pasting properties of wheat. Crop and Pasture Science. **70**:16–25.

Chung, B. Y.-W., Miller, W. A., Atkins, J. F., and Firth, A. E. 2008a. An overlapping essential gene in the Potyviridae. Proceedings of the National Academy of Sciences. **105**:5897–5902.

Chung, B. Y.-W., Miller, W. A., Atkins, J. F., and Firth, A. E. 2008b. An overlapping essential gene in the Potyviridae. Proceedings of the National Academy of Sciences. **105:**5897–5902.

Cui, H., and Wang, A. 2016. Plum pox virus 6K1 protein is required for viral replication and targets the viral replication complex at the early stage of infection. Journal of virology. **90**:5119–5131.

Cunfer, B. M., Demski, J. W., and Bays, D. C. 1988. Reduction in plant development, yield, and grain quality associated with wheat spindle streak mosaic virus. Phytopathology. **78**:198–204.

Danilova, T. V., Zhang, G., Liu, W., Friebe, B., and Gill, B. S. 2017. Homoeologous recombination-based transfer and molecular cytogenetic mapping of a wheat streak mosaic virus and Triticum mosaic virus resistance gene Wsm3 from *Thinopyrum intermedium* to wheat. Theoretical and applied genetics. 130:549–556.

Deamer, D., Akeson, M., and Branton, D. 2016. Three decades of nanopore sequencing. Nature biotechnology. **34**:518–524.

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.

DeWolf, E.D., Lollato, R. and Whitworth, R.J. 2019. Wheat Variety Disease and Insect Ratings 2019. Kansas State University, K-State Research and Extension, MF991. Wheat rating

Dolja, V., Haldeman, R., Robertson, N., Dougherty, W., and Carrington, J. 1994. Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. The EMBO Journal. **13**:1482–1491.

Dolja, V. V., Haldeman-Cahill, R., Montgomery, A. E., Vandenbosch, K. A., and Carrington, J. C. 1995. Capsid protein determinants involved in cell-to-cell and long-distance movement of tobacco etch potyvirus. Virology. **206**:1007–1016.

Dreher, T. W., and Miller, W. A. 2006. Translational control in positive strand RNA plant viruses. Virology. **344**:185–197.

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

Dwyer, G. I., Gibbs, M. J., Gibbs, A. J., and Jones, R. A. 2007. Wheat streak mosaic virus in Australia: relationship to isolates from the Pacific Northwest of the USA and its dispersion via seed transmission. Plant Disease. **91**:164–170.

Edwardson, J., Christie, R., and Ko, N. 1984. Potyvirus cylindrical inclusions– Subdivision-IV. Phytopathology. **74**:1111–1114.

Ellis, M., Rebetzke, G., Mago, R., and Chu, P. 2003. First report of Wheat streak mosaic virus in Australia. Australasian Plant Pathology. **32**:551–553.

Fahim, M., Mechanicos, A., Ayala-Navarrete, L., Haber, S., and Larkin, P. 2012. Resistance to Wheat streak mosaic virus–a survey of resources and development of molecular markers. Plant pathology. **61**:425–440.

FAO, F. 2017. The future of food and agriculture–Trends and challenges. Food and Agriculture Organization Rome.

Feldman, M., Bonjean, A., and Angus, W. 2001. The world wheat book: a history of wheat breeding. Edited by: Bonjean AP, Angus WJ. :3–53.

Fellers, J. P., Seifers, D., Ryba-White, M., and Joe Martin, T. 2009. The complete genome sequence of Triticum mosaic virus, a new wheat-infecting virus of the High Plains. Archives of virology. **154**:1511–1515.

Fellers, J., Wan, J., Hong, Y., Collins, G. B., and Hunt, A. G. 1998. In vitro interactions between a potyvirus-encoded, genome-linked protein and RNA-dependent RNA polymerase. Journal of General Virology. **79**:2043–2049.

Fellers, J., Webb, C., Fellers, M., Shoup Rupp, J., and De Wolf, E. 2019. Wheat virus identification within infected tissue using nanopore sequencing technology. Plant Disease. **103**: 2199-2203.

Fernández, A., Guo, H. S., Sáenz, P., Simón-Buela, L., de Cedrón, M. G., and García, J. A. 1997. The motif V of plum pox potyvirus CI RNA helicase is involved in NTP hydrolysis and is essential for virus RNA replication. Nucleic acids research. **25**:4474–4480.

Filloux, D., Fernandez, E., Loire, E., Claude, L., Galzi, S., Candresse, T., Winter, S., Jeeva, M. L., Makeshkumar, T., Martin, D. P., and Roumagnac, P. 2018. Nanopore-based detection and characterization of yam viruses. Scientific reports. **8**:1–11.

Friebe, B., Mukai, Y., Dhaliwal, H., Martin, T., and Gill, B. 1991. Identification of alien chromatin specifying resistance to wheat streak mosaic and greenbug in wheat germ plasm by C-banding and in situ hybridization. Theoretical and applied genetics. **81**:381–389.

Friebe, B., Qi, L., Wilson, D., Chang, Z., Seifers, D., Martin, T., Fritz, A., and Gill, B.S. 2009. Wheat–*Thinopyrum intermedium* recombinants resistant to Wheat streak mosaic virus and Triticum mosaic virus. Crop science. **49**:1221–1226.

Gadiou, S., Kudela, O., Ripl, J., Rabenstein, F., Kundu, J. K., and Glasa, M. 2009. An amino acid deletion in wheat streak mosaic virus capsid protein distinguishes a homogeneous group of European isolates and facilitates their specific detection. Plant Disease. **93**:1209–1213.

Gibbs, A., and Mackenzie, A. 1997. A primer pair for amplifying part of the genome of all potyvirids by RT-PCR. Journal of Virological methods. **63**:9–16.

Graybosch, R. A., Peterson, C., Baenziger, P. S., Baltensperger, D. D., Nelson, L. A., Jin, Y., Kolmer, J., Seabourn, B., French, R., Hein, G. and Martin, T. J. 2009. Registration of 'Mace'hard red winter wheat. Journal of Plant Registrations. **3**:51–56.

Guillen-Portal, F., Garetson, R., Bell, J., Trostle, C., Kimura, E., McGinty, J., Noland, R., Drake, D., Ramirez, J., Ibrahim, A., Rudd, J., Sutton, R., Devkota, R., Baker, J., Baker, S., Opena, G., Simoneaux, B., Naylor C., and Braley, A. 2021. 2021 Texas wheat uniform variety trial results. varietyesting.amu.eud/wheat. Texas A&M AgriLife research and extention, soil and crop sciences. SCSC-2021-11.

Gupta, A. K., Hein, G. L., Graybosch, R. A., and Tatineni, S. 2018. Octapartite negative-sense RNA genome of High Plains wheat mosaic virus encodes two suppressors of RNA silencing. Virology. **518**:152–162.

Gupta, A. K., Hein, G. L., and Tatineni, S. 2019. P7 and P8 proteins of High Plains wheat mosaic virus, a negative-strand RNA virus, employ distinct mechanisms of RNA silencing suppression. Virology. **535**: 20-31

Gupta, A. K., and Tatineni, S. 2019. RNA silencing suppression mechanisms of Triticum mosaic virus P1: dsRNA binding property and mapping functional motifs. Virus Research. :197640.

Hadi, B., Langham, M., Osborne, L., and Tilmon, K. 2011. Wheat streak mosaic virus on wheat: biology and management. Journal of Integrated Pest Management. **2**: J1–J5.

Haley, S. D., Johnson, J. J., Peairs, F. B., Stromberger, J. A., Heaton, E. E., Seifert, S. A., et al. 2011. Registration of Snowmass wheat. Journal of Plant Registration. **5**: 87-90.

Haley, S. D., Martin, T. J., Quick, J. S., Seifers, D. L., Stromberger, J. A., Clayshulte, S.R., Clifford, B. L., Peairs, F. B., Rudolph, J. B., Johnson, J. J., and Gill, B. S. 2002. Registration of CO960293-2 wheat germplasm resistant to Wheat streak mosaic virus and Russian wheat aphid. (Registrations of Germplasm). Crop science. **42**:1381–1383.

Hall, J. S., French, R., Hein, G. L., Morris, T. J., and Stenger, D. C. 2001. Three distinct mechanisms facilitate genetic isolation of sympatric wheat streak mosaic virus lineages. Virology. **282**:230–236.

Halvin, J. L., Beaton, J., Tisdale, S., and Nelson, W. 2005. Soil fertility and fertilizers: an introduction to nutrient management. Prentice Hall, New Jersey.

Harvey, T., Seifers, D., Martin, T., Brown-Guedira, G., and Gill, B. 1999. Survival of wheat curl mites on different sources of resistance in wheat. Crop science. **39**:1887–1889.

He, G., Zhang, Z., Sathanantham, P., Diaz, A., and Wang, X. 2021. Brome Mosaic Virus (Bromoviridae). Encyclopedia of Virology. :252.

Heger, M. 2014. At AGBT, first data from Oxford Nanopore presented as company issues early access invites. In Sequence/GenomeWeb.

Hodge, B., Paul, P., and Stewart, L. R. 2020. Occurrence and High-Throughput Sequencing of Viruses in Ohio Wheat. Plant Disease. **104**:1789–1800.

Hodge, B., Salgado, J., Paul, P., and Stewart, L. 2019. Characterization of an Ohio isolate of Brome mosaic virus and its impact on the development and yield of soft red winter wheat. Plant Disease. **103**:1101–1111.

Hollandbeck, F. G., Dewolf, E., and Todd, T. 2017. Kansas Cooperative Plant Disease Survey Report, Preliminary 2017 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Hollandbeck, F. G., Dewolf, E., and Todd, T. 2020. Kansas Cooperative Plant Disease Survey Report, Preliminary 2020 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Hollandbeck, F. G., Onofre, K. A., Dewolf, E., and Todd, T. 2021. Kansas Cooperative Plant Disease Survey Report, Preliminary 2021 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Hunger, R., Sherwood, J., Evans, C., and Montana, J. 1992. Effects of planting date and inoculation date on severity of wheat streak mosaic in hard red winter wheat cultivars. Plant Disease. **76**:1056–1060.

Ito, D., Miller, Z., Menalled, F., Moffet, M., and Burrows, M. 2012. Relative susceptibility among alternative host species prevalent in the Great Plains to Wheat streak mosaic virus. Plant Disease. **96**:1185–1192.

Jain, M., Olsen, H. E., Paten, B., and Akeson, M. 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome biology. **17**:1–11.

Jain, M., Tyson, J. R., Loose, M., Ip, C. L., Eccles, D. A., O'Grady, J., Mall, S., Leggett, R. M., Wallerman, O., Jansen, H. J., and Zalunin, V. 2017. MinION Analysis and Reference Consortium: Phase 2 data release and analysis of R9. 0 chemistry. F1000Research. 6.

James, D., Varga, A., Pallas, V., and Candresse, T. 2006.Strategies for simultaneous detection of multiple plant viruses. Canadian Journal of Plant Pathology. **28**:16–29.

Jensen, S., Lane, L., and Seifers, D. 1996. A new disease of maize and wheat in the High Plains. Plant Disease. **80**:1387–1390.

Jiang, W., Garrett, K., Peterson, D., Harvey, T., Bowden, R., and Fang, L. 2005. The window of risk for emigration of Wheat streak mosaic virus varies with host eradication method. Plant disease. **89**:853–858.

Johnson, J., Haley, S., Jones, J., Asfeld, E., Meyer, R., Trujillo, W., Kann, D., Roesch, K., Spring, J., Larson, K., Vigil, M., Pettinger, B. 2019. Colorado Winter Wheat Variety Performance Trials. Tech. Rep. 19-2. Agricultural Experiment Station, Colorado State University, Ft. Collins, Colorado.

Kajla, M., Yadav, V. K., Khokhar, J., Singh, S., Chhokar, R., Meena, R. P., and Sharma, R. K. 2015. Increase in wheat production through management of abiotic stresses: a review. Journal of Applied and Natural Science. **7**:1070–1080.

Kao, C. C., and Sivakumaran, K. 2000. Brome mosaic virus, good for an RNA virologist's basic needs. Molecular plant pathology. **1**:91–97.

Kapooria, R., and Ndunguru, J. 2004. Occurrence of viruses in irrigated wheat in Zambia. EPPO Bulletin. **34**:413–419.

Kashyap, P. L., Kumar, S., Jasrotia, P., Singh, D. P., and Singh, G. P. 2020. Nanotechnology in wheat production and protection. In Environmental Nanotechnology Volume *4*, Springer, p. 165–194

Kasianowicz, J. J., Brandin, E., Branton, D., and Deamer, D. W. 1996. Characterization of individual polynucleotide molecules using a membrane channel. Proceedings of the National Academy of Sciences. **93**:13770–13773.

Keifer, H. 1969. Eriophyid Studies C-3. Department of Agriculture. Agricultural Research Service.

Khan, M. A., Miyoshi, H., Gallie, D. R., and Goss, D. J. 2008. Potyvirus genome-linked protein, VPg, directly affects wheat germ in vitro translation interactions with translation initiation factors eIF4F and eIFiso4F. Journal of Biological Chemistry. **283**:1340–1349.

Klein, P. G., Klein, R. R., Rodriguez-Cerezo, E., Hunt, A. G., and Shaw, J. G. 1994. Mutational analysis of the tobacco vein mottling virus genome. Virology. **204**:759–769.

Kleinjan, J., Graham, C., Sehgal, S., Ali, S., Kirby, K., Hawks, S., and Swan, B. 2021. 2021 South Dakota winter wheat variety trial results regional summaries. South Dakota State University Extension, South Dakota Agricultural Experiment Station at SDSU, 2021 South Dakota Board of Regents. https://extension.sdstate.edu/sites/default/files/2021-08/S-0002-2021-01-WW-Regional_Summary-v2.pdf

KS wheat. 2017. Kansas wheat. Wheat streak mosaic is timely issue at wheat. http://kswheat.com/news/2017/08/09/wheat-streak-mosaic-is-timely-issue-at-wheatu-event (accessed on 2/12/2022)

Lebrigand, K., Magnone, V., Barbry, P., and Waldmann, R. 2020. High throughput error corrected Nanopore single cell transcriptome sequencing. Nature communications. **11**:1–8.

Leiva, A. M., Siriwan, W., Lopez-Alvarez, D., Barrantes, I., Hemniam, N., Saokham, K.,and Cuellar, W. J. 2020. Nanopore-based complete genome sequence of a Sri Lankan cassava mosaic virus (*Geminivirus*) strain from Thailand. Microbiology Resource Announcements. **9**: e01274-19.

Leonard, S., Viel, C., Beauchemin, C., Daigneault, N., Fortin, M. G., and Laliberte, J. F. 2004. Interaction of VPg-Pro of Turnip mosaic virus with the translation initiation factor 4E and the poly (A)-binding protein in planta. Journal of General Virology. **85**:1055–1063.

Li, X. H., Valdez, P., Olvera, R. E., and Carrington, J. C. 1997. Functions of the tobacco etch virus RNA polymerase (NIb): subcellular transport and protein-protein interaction with VPg/proteinase (NIa). Journal of virology. **71**:1598–1607.

Liefting, L. W., Waite, D. W., and Thompson, J. R. 2021. Application of Oxford Nanopore Technology to Plant Virus Detection. Viruses. **13**:1424.

Liu, W., Seifers, D., Qi, L., Friebe, B., and Gill, B. 2011. A compensating wheat–*Thinopyrum intermedium* Robertsonian translocation conferring resistance to wheat streak mosaic virus and Triticum mosaic virus. Crop science. **51**:2382–2390.

Loman, N. J., Quick, J., and Simpson, J. T. 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. Nature methods. **12**:733–735.

Lu, H., Price, J., Devkota, R., Rush, C., and Rudd, J. 2011. A dominant gene for resistance to Wheat streak mosaic virus in winter wheat line CO960293-2. Crop science. **51**:5–12.

Lucas, R. W., Larson, S. B., and McPherson, A. 2002. The crystallographic structure of brome mosaic virus. Journal of molecular biology. **317**:95–108.

Mahmood, T., Hein, G. L., and Jensen, S. 1998. Mixed infection of hard red winter wheat with high plains virus and wheat streak mosaic virus from wheat curl mites in Nebraska. Plant Disease. **82**:311–315.

Marburger, D., Calhoun, R., Carver, B., Hunger, B., Watson, B., and Gillespie, C. 2018. Oklahoma small Grains variety performance tests 2017-2018. OSU cooperative extension service, CR-2141 and CR-2143. Oklahoma State University, Stillwater, Oklahoma

Martin, T. J., Zhang, G., Fritz, A. K., Miller, R., and Chen, M. S. 2014. Registration of 'Clara CL' wheat. Journal of Plant Registrations. **8**:38–42.

McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annual review of phytopathology. **40**:349.

McKinney, H. H. 1937. Mosaic diseases of wheat and related cereals. US Department of Agriculture.

McNeil, J. E., French, R., Hein, G. L., Baenziger, P. S., and Eskridge, K. M. 1996. Characterization of genetic variability among natural populations of wheat streak mosaic virus. Phytopathology. **86**:1222–1227.

Mehetre, G. T., Leo, V. V., Singh, G., Sorokan, A., Maksimov, I., Yadav, M. K., Upadhyaya, K., Hashem, A., Alsaleh, A. N., Dawoud, T. M., and Almaary, K. S. 2021. Current developments and challenges in plant viral diagnostics: A systematic review. Viruses. 13:412.

Middleton, C. P., Senerchia, N., Stein, N., Akhunov, E. D., Keller, B., Wicker, T., and Kilian, B. 2014. Sequencing of chloroplast genomes from wheat, barley, rye, and their relatives provides a detailed insight into the evolution of the Triticeae tribe. PLoS One. **9**: e85761

Miller, Z. J., Lehnhoff, E. A., Menalled, F. D., and Burrows, M. 2015. Effects of soil nitrogen and atmospheric carbon dioxide on Wheat streak mosaic virus and its vector (*Aceria tosichella* Kiefer). Plant Disease. **99**:1803–1807.

Monger, W., Goodfellow, H., and Back, E. 2020. First report of Arabis mosaic virus in potato (*Solanum tuberosum*), identified by nanopore sequencing. New Disease Report. **41**:2044–0588.

Mumford, R., Jarvis, B., Harju, V., Boonham, N., and Skelton, A. 2006. First report of Broad bean wilt virus 2 in the UK: findings in foxglove and salvia. Plant Pathology. **55**:819–819.

Murphy, C. Y. 2016. Chemical control and disease reservoir studies of the wheat curl mite (*Aceria tosichella* Keifer), vector to wheat streak mosaic virus. MS thesis, Montana State University, Bozeman, MT

Murray, G. M., Knihinicki, D., Wratten, K., and Edwards, J. 2005. Wheat streak mosaic and the wheat curl mite. NSW Department of Primary Industries, Orange NSW Australia. Primefact. 99.

Mylonas, I., Stavrakoudis, D., Katsantonis, D., and Korpetis, E. 2020. Better farming practices to combat climate change. In Climate Change and Food Security with Emphasis on Wheat, Elsevier, p. 1–29.

Naito, F. Y., Melo, F. L., Fonseca, M. E. N., Santos, C. A., Chanes, C. R., Ribeiro, B. M., Gilbertson, R. L., Boiteux, L. S., and de Cassaia Pereira-Carvalho, R. 2019. Nanopore sequencing of a novel bipartite New World begomovirus infecting cowpea. Archives of virology. **164**:1907–1910.

Navia, D., de Mendonça, R. S., Skoracka, A., Szyd\lo, W., Knihinicki, D., Hein, G. L., da Silva Pereira, P. R. V., Truol, G., and Lau, D. 2013. Wheat curl mite, *Aceria tosichella*, and transmitted viruses: an expanding pest complex affecting cereal crops. Experimental and Applied Acarology. **59**:95–143.

Ordon, F., Habekuss, A., Kastirr, U., Rabenstein, F., and Kühne, T. 2009. Virus resistance in cereals: sources of resistance, genetics, and breeding. Journal of phytopathology. **157**:535–545.

Price, J., Workneh, F., Evett, S., Jones, D., Arthur, J., and Rush, C. 2010. Effects of Wheat streak mosaic virus on root development and water-use efficiency of hard red winter wheat. Plant Disease. **94**:766–770.

Qu, F., Ren, T., and Morris, T. J. 2003. The coat protein of turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. Journal of virology. **77**:511–522.

Rabenstein, F., Seifers, D. L., Schubert, J., French, R., and Stenger, D. C. 2002. Phylogenetic relationships, strain diversity and biogeography of Tritimoviruses. Journal of General Virology. **83**:895–906.

Ranabhat, N. B., Seipel, T., Lehnhoff, E. A., Miller, Z. J., Owen, K. E., Menalled, F. D., Burrows, M. E. 2018. Temperature and alternative hosts influence *Aceria tosichella* infestation and wheat streak mosaic virus infection. Plant Disease. **102**:546–551.

Ranieri, R., Lister, R. M., and Burnett, P. A. 1993. Relationships between Barley Yellow Dwarf Virus Titer and Symptom Expression in Barley. Crop Science. **33**:968–973.

Rao, A. 2006. Genome packaging by spherical plant RNA viruses. Annual Review of Phytopathology. **44**:61–87.

Revers, F., and García, J. A. 2015. Molecular biology of potyviruses. In Advances in virus research, Elsevier, p. 101–199.

Roossinck, M. J. 2012. Plant virus metagenomics: biodiversity and ecology. Annual review of genetics. **46**:359–369.

Rotenberg, D., Bockus, W. W., Whitfield, A. E., Hervey, K., Baker, K. D., Ou, Z., Laney, A. G., DeWolf, E. D., and Appel, J. A. 2016. Occurrence of viruses and associated grain yields of paired symptomatic and non-symptomatic tillers in Kansas winter wheat fields. Phytopathology. **106**:202–210.

Rupp, J. L. S. 2015. RNA interference mediated virus resistance in transgenic wheat. PhD Dissertation. Kansas State University, Manhattan, KS

Sánchez-Sánchez, H., Henry, M., Cárdenas-Soriano, E., and Alvizo-Villasana, H. 2001. Identification of Wheat streak mosaic virus and its vector *Aceria tosichella* in Mexico. Plant Disease. **85**:13–17.

Sanfaçon, H. 2017. Grand challenge in plant virology: understanding the impact of plant viruses in model plants, in agricultural crops, and in complex ecosystems. Frontiers in microbiology **8**: 860.

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. 2019. The global burden of pathogens and pests on major food crops. Nature ecology & evolution. **3**:430–439.

Seifers, D. L., Harvey, T. L., Martin, T., and Jensen, S. G. 1997. Identification of the wheat curl mite as the vector of the High Plains virus of corn and wheat. Plant Disease. **81**:1161–1166

Seifers, D. L., Martin, T., Harvey, T. L., Fellers, J. P., and Michaud, J. 2009. Identification of the wheat curl mite as the vector of Triticum mosaic virus. Plant Disease. **93**:25–29.

Seifers, D. L., Martin, T., Harvey, T. L., Fellers, J. P., Stack, J. P., Ryba-White, M., Haber, S. Krokhin, O., Spicer, V., Lovat, N., and Yamchuk, A. 2008. Triticum mosaic virus: A new virus isolated from wheat in Kansas. Plant Disease. **92**:808–817.

Seifers, D., Martin, T., Harvey, T., and Haber, S. 2007. Temperature-sensitive Wheat streak mosaic virus resistance identified in KS03HW12 wheat. Plant Disease. **91**:1029–1033.

Seifers, D., Martin, T., Harvey, T., Haber, S., and Haley, S. 2006. Temperature sensitivity and efficacy of Wheat streak mosaic virus resistance derived from CO960293 wheat. Plant Disease. **90**:623–628.

Shewry, P. R. 2009. Wheat. Journal of experimental botany. **60**:1537–1553.

Singh, K., Wegulo, S. N., Skoracka, A., and Kundu, J. K. 2018. Wheat streak mosaic virus: a century old virus with rising importance worldwide. Molecular plant pathology. **19**:2193–2206.

Siriwetwiwat, B. 2006. Interactions between the wheat curl mite, *Aceria tosichella* Keifer (Eriophyidae), and wheat streak mosaic virus and distribution of wheat curl mite biotypes in the field. The University of Nebraska-Lincoln.

Skare, J. M., Wijkamp, I., Denham, I., Rezende, J. A., Kitajima, E. W., Park, J. W., Desvoyes, B., Rush, C. M., Michels, G., Scholthof, K. B. G., and Scholthof, H. B. 2006. A new eriophyid mite-borne membrane-enveloped virus-like complex isolated from plants. Virology. **347**:343–353.

Slykhuis. 1955. *Aceria tulipae* Keifer (Acarina: Eriophyidae) in relation to spread of wheat streak mosaic virus. Phytopathology. **45**:116–128.

Somsen, H. W., and Sill, W. H. 1970. The wheat curl mite, *Aceria tulipae* Keifer, in relation to epidemiology and control of wheat streak mosaic. Agricultural Experiment Station.

Srivatsavai, V. 2005. Identification, distribution, and vector biology of brome mosaic virus of wheat in Alabama MS Thesis. University of Alabama.

Stenger, D. C., Burbank, L. P., Wang, R., Stewart, A. A., Mathias, C., and Goodin, M. M. 2020. Lost and found: Rediscovery and genomic characterization of sowthistle yellow vein virus after a 30+ year hiatus. Virus Research. **284**:197987.

Stenger, D. C., French, R., and Gildow, F. E. 2005. Complete deletion of Wheat streak mosaic virus HC-Pro: a null mutant is viable for systemic infection. Journal of virology. **79**:12077–12080.

Stenger, D. C., Hall, J. S., Choi, I. R., and French, R. 1998. Phylogenetic relationships within the family Potyviridae: wheat streak mosaic virus and brome streak mosaic virus are not members of the genus *Rymovirus*. Phytopathology. **88**:782–787.

Stenger, D. C., Hein, G. L., and French, R. 2006a. Nested deletion analysis of Wheat streak mosaic virus HC-Pro: mapping of domains affecting polyprotein processing and eriophyid mite transmission. Virology. **350**:465–474.

Stenger, D. C., Seifers, D. L., and French, R. 2002. Patterns of polymorphism in Wheat streak mosaic virus: sequence space explored by a clade of closely related viral genotypes rivals that between the most divergent strains. Virology. **302**:58–70.

Stenger, D. C., Young, B. A., and French, R. 2006b. Random mutagenesis of wheat streak mosaic virus HC-Pro: non-infectious interfering mutations in a gene dispensable for systemic infection of plants. Journal of general virology. **87**:2741–2747.

Stewart, L. R. 2016. Sequence diversity of wheat mosaic virus isolates. Virus research. **213**:299–303.

Susi, P. 2004. Black currant reversion virus, a mite-transmitted nepovirus. Molecular Plant Pathology. **5**:167–173.

Syller, J. 2012. Facilitative and antagonistic interactions between plant viruses in mixed infections. Molecular plant pathology **13**: 204–216.

Tatineni, S., Alexander Jeff, and Qu Feng 2021. Differential Synergistic Interactions among Four Different Wheat-infecting Viruses. Frontiers in microbiology.**12**: 800318.

Tatineni, S., Elowsky, C., and Graybosch, R. A. 2017. Wheat streak mosaic virus Coat Protein Deletion Mutants Elicit More Severe Symptoms Than Wild-Type Virus in Multiple Cereal Hosts. Molecular plant-microbe interactions: MPMI. **30**:974–983.

Tatineni, S., and French, R. 2014. The C-terminus of Wheat streak mosaic virus coat protein is involved in differential infection of wheat and maize through host-specific long-distance transport. Molecular plant-microbe interactions. **27**:150–162.

Tatineni, S., and Hein, G. L. 2018. Genetics and mechanisms underlying transmission of Wheat streak mosaic virus by the wheat curl mite. Current opinion in virology. **33**:47–54.

Tatineni, S., and Hein, G. L. 2021. High Plains wheat mosaic virus: An enigmatic disease of wheat and corn causing the High Plains disease. Molecular Plant Pathology. **22**:1167–1179.

Tatineni, S., Kovacs, F., and French, R. 2014a. Wheat streak mosaic virus infects systemically despite extensive coat protein deletions: Identification of virion assembly and cell-to-cell movement determinants. Journal of virology. **88**:1366–1380.

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

Tatineni, S., McMechan, A. J., Wosula, E. N., Wegulo, S. N., Graybosch, R. A., French, R., and Hein, G. L. 2014b. An eriophyid mite-transmitted plant virus contains eight genomic RNA segments with unusual heterogeneity in the nucleocapsid protein. Journal of virology. **88**:11834–11845.

Tatineni, S., Ziems, A. D., Wegulo, S. N., and French, R. 2009. Triticum mosaic virus: a distinct member of the family Potyviridae with an unusually long leader sequence. Phytopathology. **99**:943–950.

Thomas, J. A., and Hein, G. L. 2003. Influence of volunteer wheat plant condition on movement of the wheat curl mite, *Aceria tosichella*, in winter wheat. Experimental & applied acarology. **31**:253–268.

Thomas, J., Hein, G., and Lyon, D. 2004. Spread of wheat curl mite and Wheat streak mosaic virus is influenced by volunteer wheat control methods. Plant health progress. **5**:2.

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

Valli, A., Lopez-Moya, J. J., and Garcia, J. A. 2007. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. Journal of General Virology. **88**:1016–1028.

Velandia, M., Rejesus, R. M., Jones, D. C., Price, J. A., Workneh, F., and Rush, C. M. 2010. Economic impact of Wheat streak mosaic virus in the Texas High Plains. Crop Protection. **29**:699–703.

Vogel, J., Garvin, D., Mockler, T. C., Schmutz, J., Rokhsar, D., and Bevan, M. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. **463**:7282

Wang, M. B., and Metzlaff, M. 2005. RNA silencing and antiviral defense in plants. Current opinion in plant biology. **8**:216–222.

Wang, Y., Yang, Q., and Wang, Z. 2015. The evolution of nanopore sequencing. Frontiers in genetics. **5**:449.

Wei, T., Zhang, C., Hong, J., Xiong, R., Kasschau, K. D., Zhou, X., Carrington, J. C., and Wang, A. 2010. Formation of complexes at plasmodesmata for potyvirus intercellular movement is mediated by the viral protein P3N-PIPO. PLoS pathogens. **6**. e1000962.

Wosula, E., McMechan, A. J., Knoell, E., Tatineni, S., Wegulo, S. N., and Hein, G. L. 2018. Impact of timing and method of virus inoculation on the severity of wheat streak mosaic disease. Plant Disease. **102**:645–650.

Xiong, R., and Wang, A. 2013. SCE1, the SUMO-conjugating enzyme in plants that interacts with NIb, the RNA-dependent RNA polymerase of Turnip mosaic virus, is required for viral infection. Journal of virology. **87**:4704–4715.

Young, B. A., Stenger, D. C., Qu, F., Morris, T. J., Tatineni, S., and French, R. 2012. *Tritimovirus* P1 functions as a suppressor of RNA silencing and an enhancer of disease symptoms. Virus research. **163**:672–677.

Young, B., Hein, G. L., French, R., and Stenger, D. C. 2007. Substitution of conserved cysteine residues in wheat streak mosaic virus HC-Pro abolishes virus transmission by the wheat curl mite. Archives of virology. **152**:2107–2111.

Zhang, G., Martin, T. J., Fritz, A. K., Miller, R., Chen, M.-S., Bowden, R. L., and Bai, G. 2016. Registration of ‘Joe’ hard white winter wheat. Journal of Plant Registrations. **10**:283–28.

Figure 1.1. A wheat field in Ness County Kansas infected with wheat streak mosaic complex in 2019. Symptoms of yellow mosaic can be seen in the field and highlighted part of the picture; total loss of yield was observed in the individual field (Photo credit: Ranabhat 2019).

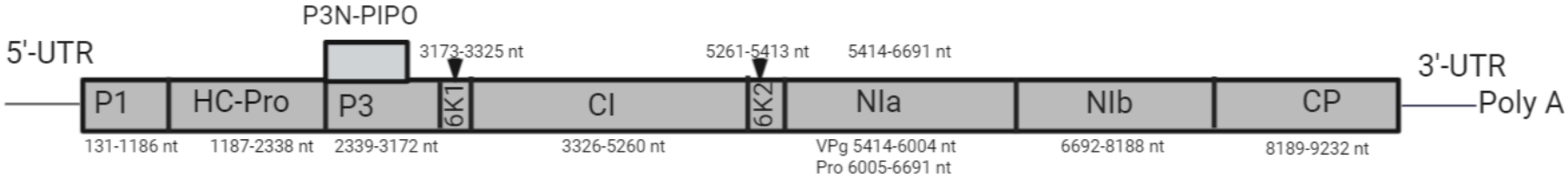


Figure 1.2. Genomic organization of Wheat streak mosaic virus (WSMV) representing a typical genomic organization of potyviruses The poly protein is composed of 10 proteins: P1 (P1 protein: 40 kDa); HC-pro (Helper component protease: 44 kDa); P3 (P3 protein: 32 kDA); 6K1 and 6K2 (6 kDA protein), CI (Cytoplasmic inclusion protein: 73 kDa); VPg (viral protein genome-linked proteinase: 23 kDa); NIa (Nuclear inclusion putative protease: 26 kDa); NIb (Nuclear inclusion putative polymerase: 57 kDa); CP (Coat protein: 37 kDa). nt, nucleotides; UTR, untranslated region (Graphic credit: Ranabhat 2020)

Chapter 2 - Wheat virome metagenomic and phylogenetic analysis using high-throughput Nanopore sequencing

Abstract

Wheat viruses have been recognized as common yield-reducing factors of wheat. The synergistic impact of several concurrently infecting wheat viruses is responsible for millions of dollars lost in wheat production. Genetic resistance along with good cultural practices is one of the key management practices of wheat viruses. The primary threat to durable resistance is the presence of potential new virus variants. Contemporary target-specific diagnosis methods use specific antibodies, primers, and probes to detect viral pathogens. In this work, high-throughput Oxford Nanopore sequencing technique (ONT) was used to detect and determine the viral population structure in wheat samples. A survey was conducted in 2019, 2020, and 2021 in major wheat-growing counties of Kansas, and wheat leaves showing virus-like symptoms were collected. Total RNA was extracted, and cDNA sequencing libraries were made using a PCR-cDNA barcoding kit and loaded onto ONT MinION flow cells. Sequencing reads were aligned to cereal virus references. We identified eight wheat viruses belonging to the genera: *Tritimovirus*, *Poacevirus*, *Emaraviurs*, *Bromovirus*, *Luteovirus*, *Polerovirus*, *Bymovirus*, and *Furovirus.* We recorded mixed infections of two to five viruses in a single sample. Wheat streak mosaic virus (WSMV) + triticum mosaic virus (TriMV) mixed infection was most predominant (16.7%), followed by WSMV + TriMV + BMV (brome mosaic virus) (11.9%) and WSMV single infection (11.9%). BMV was identified for the first time in a Kansas wheat field infected with other wheat viruses. Phylogenetic analysis of the whole genomes of WSMV revealed the wide distribution of isolates with European and recombinant isolates. A phylogenetic analysis of HPWMoV Kansas isolates was conducted based on RNA 3A and 3B. The whole-genome

sequences of soilborne wheat mosaic virus isolates were also characterized. Overall, the information of genetic variability, the phylogenetic relationship among isolates, and reports of new isolates of viruses and their co-infections will help recommend sustainable management practices for wheat viruses through host genetic resistance.

Introduction

Kansas is one of the top wheat-producing states in the United States. It produced 7.654 million tons in 2020, and 9.906 million tons in 2021, ranked second and first leading producer (USDA, National Agricultural Statistics Service, 2019 and 2020). Many biotic agents, such as fungal, bacterial, and viral pathogens, insects, and nematodes, cause significant loss in wheat production. It is hard to estimate crop loss accurately due to biotic factors only as abiotic factors are always associated with. However, researchers in Kansas estimated the history of cumulative disease loss for wheat from 1976 to 2020 that was ranged from 0.2% to 22.2% of potential yield loss; in 2020 and 2021, the cumulative loss of wheat by disease (excluding nematodes) was estimated to be 10.8% and 16.2% or about 31.8 million bushels and 70.4 million bushels respectively (Hollandbeck et al 2020 and 2021).

The frequent monitoring and characterization of viral field populations provide information about new isolates, dominant or new co-occurring virus combinations, or novel viruses. Diagnosis of virus-like symptoms in plants has been mostly dominated by specific methods with known antibodies, primers, and probes. These conventional diagnostic techniques, such as serological tests and reverse-transcription PCR, can only test for known viruses. Because of the portability, rapid results, accurate identification of multiple pathogens, short library preparation time, Oxford Nanopore sequencing technology (ONT) is a commonly used cutting-edge technology (Phannareth et al., 2020). This technology has been used as a surveillance tool

for detecting fungal, bacterial (Bronzato Badial et al., 2018), as well as plant viral pathogens (Fellers et al., 2019).

Viral pathogens are a common threat to wheat production. These viruses are vectored by aphid (e.g., Barley yellow dwarf virus, Cereal yellow dwarf virus), mite (e.g., Viruses of wheat streak mosaic complex), plasmodiophorid *Polymyxa graminis* (e.g., Soilborne wheat viruses), or leafhoppers (*Wheat dwarf virus*) (Ordon et al., 2009). In Kansas, the most common viral diseases of wheat are Barley yellow dwarf, 'Wheat steak mosaic' (WSM) complex and soilborne viruses. Barley yellow dwarf is a disease caused by the infection of group of Barley yellow dwarf viruses (BYDV) of Genus *Luteovirus* (e.g., Barley yellow dwarf virus MAV, Barley yellow dwarf virus PAS, Barley yellow dwarf virus PAV), and Genus *Sobemovirus* (e.g., Barley yellow dwarf virus GPV and Barley yellow dwarf virus SGV) (Virus Taxonomy 2020 Release) each vectored by a different aphid. WSM disease can be caused by three distinct viral pathogens including wheat streak mosaic virus (WSMV), Triticum mosaic virus (TriMV), and High plains wheat mosaic emaravirus (HPWMoV). All three viruses are transmitted by an eriophyid mite, the wheat curl mite (*Aceria tosichella* Keifer) within and between the fields. In the Great Plains, WSM causes approximately 2% estimated loss over 5 years average; however, total loss in localized fields is common (Burrows et al. 2009; Ranabhat, N. personal observation 2019). Annual yield loss caused by WSM is variable. In an epidemic year such as in 2017 Kansas wheat producers lost an estimated 5.6% of total yield (19.2 million bushels of wheat) worth $76.8 million (KSwheat 2017, Hollandbeck et al. 2017).

Other major wheat viruses include Cereal yellow dwarf virus (CYDV), Wheat spindle streak mosaic virus (WSSMV), Wheat yellow mosaic virus, and Soilborne wheat mosaic virus (SBWMV) (Brakke, 1987; Hodge et al., 2020; Rotenberg et al., 2016). These viral pathogens

reduce wheat production qualitatively and quantitatively (Byamukama et al., 2014; Campbell et al., 1975; Choudhury et al., 2019; Cunfer et al., 1988). Hodge et al. 2020 also reported brome mosaic virus and Cockfoot mottle virus from Ohio. The Oho isolate, BMV_OH, reduced grain yield by 61% in soft red winter wheat inoculated mechanically at an early growth stage compared to non-inoculated control (Hodge et al., 2019).

Surveys of wheat viruses have been conducted in many states of Great Plains (Burrows et al., 2009) and Kansas only (Rotenberg et al., 2016). These studies reported that BYDV-PAV, CYDV-RPV, WSMV, TriMV, HPWMoV, SBWMV, and WSSMV were the most common wheat viruses in Kansas. Rotenberg et al. 2016 found that BYDV-PAV and WSMV were the most prevalent co-occurring viruses across the state. Other co-occurring virus combinations in the single plant were WSMV + TriMV, WSMV + HPWMoV, WSMV + TriMV + HPWMoV, BYDV + SBWMV, BYDV + WSMV + HPWMoV + CYDV-RPV, WSMV + SBWMV, BYDV + SBWMV, BYDV + WSSM in Kansas wheat fields (Burrows et al., 2009; Rotenberg et al., 2016). However, these surveys were based only on detecting samples by ELISA of only known viruses.

High throughput next generation sequencing with barcoding can be used to characterize large numbers of field samples. This method could be cost-effective and precise in metagenomics studies and future plant disease diagnostic areas. Comparing whole-genome (complete or near-complete) sequences among isolates will increase the likelihood of describing genomic variability among them. In this study, we established the current statewide prevalence and distribution of wheat viruses in Kansas wheat fields. Using ONT, we determined that WSMV, TriMV and HPWMoV were still the predominant viruses. But we identified brome mosaic virus for the first time in Kansas wheat fields. The information on wheat virus

prevalence, characterization, statewide distribution, and status of co-infection of viruses will better manage the wheat viruses through genetic resistance.

Materials and methods

Winter wheat field survey

A survey of major wheat-growing counties of Kansas was carried out during the wheat growing season from May to July 2019, 2020, and 2021. Plants with virus-like symptoms of yellow discoloration and/or streaking or mosaic patterns on leaves were sampled during stem elongation and head development growth stage (Feekes 6 to 10). A total of 84 symptomatic wheat leaves (2019 = 46, 2020 = 25, and 2021= 13) were collected from winter wheat fields in 47 different counties of Kansas (Figure 2.1). Leaf tissue of each sample was stored at - 20°C until the tissue could be processed for RNA extraction.

RNA extraction

Total RNA from each sample (200 mg of tissue) was extracted using a *mir*Vana miRNA extraction kit (Ambion Catalog number: AM1560, Thermo Fisher Scientific, MA, USA) according to the manufacturer's instructions. RNA concentration was measured by a NanoDrop spectrophotometer (NanoDrop Technologies, Rockland, DE, USA). Extracted total RNA (7 µg) was cleaned with 1µl of DNase using Turbo DNase-Free ™ kit (AM 1907, Ambion®, Thermo Fisher Scientific, MA, USA) according to the manufacturer's instructions in a 50 µL reaction to get rid of host genomic DNA.

Reverse transcription and strand-switching

MinION library preparation for multiple samples was achieved using an Oxford PCR-cDNA Barcoding (SQK-PCB109) kit following the manufacturer's instruction (Oxford Nanopore Technologies, Oxford, U.K.) with the following details. A total of 1 µl (100 to

120ng/µl) total RNA was used with 1 µl of 2 µM VN primers (VNP, SQK-PCB109, ONT), 1 µl of 10 mM dNTPs (Invitrogen, catalog number: 1875160) and 8 µl of nuclease-free water (NFW) were incubated at 65°C for 5 minutes. Then the mixture of 4 µl 5x RT buffer (Invitrogen, catalog number 18090200), 1 µl of RNaseOUT (Invitrogen, catalog number 10777019), 2 µl of 10µM Strand-Switching primer (SSP, SQK-PCB109, ONT), and 1 µl of NFW were added to the incubated mixture and incubated at 42°C for 2 minutes. 1 µl of Maxima H minus Reserve Transcriptase (Invitrogen, catalog number 18090200), was added and that 20 µl total reaction volume was incubated for 90 minutes at 42°C and followed by heat inactivation for 5 minutes at 85°C.

PCR and barcoding

For up to 12 samples, 5 µl of reverse-transcribed cDNA sample was used with 1.5 µl Barcode Primers (BP01 to BP12, SQK-PCB109, ONT), 18.5 µl of NFW, and 25 µl of 2x LongAmp Taq Master mix (New England Biolabs, catalog number #M0287) mix with a total volume of 50 µl. The thermal cycler program was 95°C for 30secs for initial denaturation, 15 cycles of 95°C for 15secs, 62°C for 15secs, 65°C for 12 minutes, and final extension of 6 minutes at 65°C to amplify ~10kb products. After completing PCR, 1 µl of NEB exonuclease 1 (New England Biolabs, catalog number #M0293) was added to each PCR tube and incubated for 15 minutes at 37°C, followed by 80°C for 15 minutes. Then 0.8x volume of AMPure XP beads (Beckman Coulter, #A63881) was added to the reaction and incubated for 5 minutes in a rotator mixer. The beads were washed with freshly prepared 70% ethanol following the manufacturer's instruction (Oxford Nanopore Technologies, Oxford, U.K.). cDNA library was eluted in 12 µl of Elution buffer (EB, SQK-PCB109, ONT), and the concentration of the DNA was measured by nanodrop. After quantification, cDNA barcoded samples were pooled to a final volume of 11 µl

and 1 µl of Rapid Adaptor (RAP, SQK-PCB109, ONT) was added to the amplified cDNA library and incubated at 22°C for 15 minutes. The prepared library was loaded on the MinION R 9.4.1 flow cell (Oxford Nanopore) following the manufacturer's priming and loading (Oxford Nanopore Technologies, Oxford, U.K.).

Bioinformatics analysis

Raw fast5 format data from MinKNOW software (version: 21.06.13), guppy basecaller high accuracy (version: 3.1.5, dna_r8.4.1_450bps_hac.cfg, R 9.4.1, Oxford Nanopore) (Wick et al., 2019) was used to translate the raw electrical signal to the nucleotide sequence. The fastq format data obtained after base-calling was separated by barcode via guppy barcode. After separation, adapters were trimmed using porechop v0.2.3 (Wick et al., 2017). Nanofilt v2.3.0 (De Coster et al., 2018) was used to set min-length to 75 and max-length to 30000. Following porechop and nanofilt processing, reads were imported and mapped against cereal virus reference genomes. Using CLC Genomic Workbench® v21.0.4 (Qiagen, MD, United States) by aligning reads against reference genome of the most common cereal viruses (Supplementary Table A1). Reference was made by assembling the whole genomes sequences viruses. The following parameters were used in CLC workbench, reads were mapped to reference using the parameters following resequencing analysis with masking mode = no masking, match score = 1, mismatch cost = 2, cost of insertions and deletions = Linear gap cost, insertion cost = 3, deletion cost = 3, length fraction = 0.5, similarity fraction = 0.8, global alignment = no, non-specific match handling = map randomly, output mode = create stand-alone read mappings, create report = yes, collect unmapped reads = no. The viruses with incomplete sequences were blasted in the NCBI nucleotide (blastn) database to confirm the presence of wheat viruses. The minimum sequence read length was 1000 bp were considered as lower limit to count.

Recombinant analysis

Whole-genome consensus sequences of WSMV or TriMV of this study (Supplementary Table A2 and A3) and the complete reference genome sequences obtained from GenBank (Supplementary Table A4) were aligned using muscle alignment in Mega X (Kumar et al., 2018). Seven different algorithms were used in the RDP5 program (Martin et al., 2021) to examine the recombinant isolates. These algorithms were RDP (Martin and Rybicki, 2000), Bootscan (Martin et al., 2005), GENECONV (Padidam et al., 1999), MaxChi (Smith, 1992), 3SEQ (Lam et al., 2018), Chimaera (Posada and Crandall, 2001), and SiScan (Gibbs et al., 2000). Putative recombinants and potential parents were determined if at least four out of seven algorithm methods were significant ($P < 0.01$).

Phylogenetic analysis

Realignment of nucleoprotein sequences of all WSMV isolates except putative recombinants was performed with a muscle program in Mega X (Kumar et al., 2018), including Oat necrotic mottle virus (for WSMV), Sugarcane streak mosaic virus, and Calademia virus A (for TriMV) (Supplementary Table A5) and Raspberry leaf blotch virus (for HPWMoV) (Supplementary Table A7) as outgroups. The best-fitted nucleotide substitution models were determined by the maximum likelihood fits (Kumar et al., 2018; Nei and Kumar, 2000) necessary for constructing the phylogenetic tree. The best-fitted model was selected based on the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) scores (Guindon and Gascuel, 2003). These models were GTR + G + I (General Time Reversible model with Gamma distributed and Invariant sites) for WSMV, GTR + G (General Time Reversible model with Gamma distributed rate) for TriMV, and T92 + G (Tamura-3-parameter model with Gamma distributed rate) for HPWMoV (Kumar et al., 2018). Maximum likelihood phylogenetic

trees were constructed using Mega X with parameters as follows: number of Bootstrap replications of 1000, nucleotide substitution model as mentioned above for different viruses, number of threads of 4.

Results

Virome analysis of symptomatic samples using Nanopore sequencing

After mapping with cereal viral reference genomes, we identified eight different wheat viruses in samples collected from 47 counties in Kansas including WSMV, TriMV, BMV, HPWMoV, BYDV, CYDV, WSSMV, and SBWMV. Figure 2.2 illustrates the distribution of wheat viruses across the major wheat-growing counties of Kansas. The most dominant co-infection was WSMV + TriMV (16.7%) followed by WSMV + TriMV + BMV (11.9%) and single virus infection, WSMV only (11.9%) (Figure 2.3). We found co-infection of wheat viruses was more common than single virus infection. TriMV was found co-infected with all other viruses except SBWMV, and we found one sample infected with TriMV only. BMV was found co-infected with all other wheat viruses however most commonly co-infected with WSMV, TriMV, and HPWMoV. 2.4% of samples that had five viruses (WSMV + TriMV + HPWMoV + BMV + BYDV) infected in a single sample. WSSMV was found co-infected with all other eight viruses found in different samples, but SBWMV was found co-infected only with WSSMV, WSMV, and BMV in one sample (Figure 2.3). Pawnee and Riley were the only two counties where SBWMV was found. However, WSSMV was also recorded from Barton, Kingman, and Reno counties along with Pawnee and Riley counties (Figure 2.2).

The total incidence of a virus was calculated by the infection of the samples by a virus alone or co-infected with other viruses. The individual virus incidence based on 84 total samples collected from different counties, WSMV was the most dominant virus (94.43%) identified in all

47 counties sampled, followed by TriMV (58.33%) identified in 33 counties out of 47 sampled, and BMV (44.05%) detected in 30 counties out of 47 sampled (Table 2.1, Figure 2.2). BYDV (34.52%) and HPWMoV (27.38%) were identified in 22 and 21 counties, respectively, out of the sampled 47 major wheat-growing counties of Kansas sampled. WSSMV, CYDV, and SBWMV were found in only a few samples; they comprised only 8.33%, 3.57%, and 2.38% incidence.

Whole-genome sequencing of wheat viruses

An average of 4.48 x 10^5 raw reads (Supplementary Table A2 and A3) was obtained after mapping to a cereal virus reference genome database. An average of 9.78 x 10^3 reads of WSMV were obtained after mapping with a reference sequence with an average coverage of 991.82 X (Supplementary Table A4). 36 Complete sequences of nucleoprotein of WSMV were obtained (Supplementary Table A4) and deposited in GenBank. The remaining samples did not produce enough coverage and/or were missing a few nucleotides in the coding region of the genome were excluded for further study. For TriMV, an average of 6.06 x 10^3 reads with an average coverage of 301.56 X was obtained after mapping with the reference genome (Supplementary Table A5). 11 TriMV complete and near-complete (only missing a few nucleotides in the 5' untranslated region of the genome) genome sequences were deposited in GenBank (Supplementary Table A4). Complete and near-complete one RNA1 and RNA2, four RNA3A, three RNA3B, four RNA4, one RNA5, two RNA6, four RNA7, and two RNA8 genomes of HPWMoV with an average of 4.5 x 10^3 long reads with an average coverage of 1126.31 X were obtained and deposited in the GenBank (Supplementary Table A7). For BMV, complete and near-complete genome sequences of two RNA1, two RNA2, and seven RNA3 were obtained with an average of 1.43 x 10^5 long reads and an average coverage of 40717.19 X and deposited in GenBank (Supplementary Table B3). Two complete genomes of RNA1 and RNA2 with an average of 6.93

x 10^3 long reads with an average coverage of 410.71 X (Supplementary Table A9) of SWMV were obtained. Coverage of BYDV and WSSMV was insufficient to produce useful full genome sequences and were omitted from GenBank submission.

Sequence alignments

WSMV

The whole-genome sequences were aligned with sequences obtained from GenBank (Supplementary Table A6). Most of the isolates showed high nucleotide similarities (≥ 97%) with other already characterized isolates from the US except for three isolates (19RH1, 19SV, and 20GO). 19RH1 showed only 88.4%, 19SV showed 90.4%, but 20GO showed 94.5% nucleotide similarities with WSMV type strain and other US isolates (Supplementary Table A4). Most isolates showed lower nucleotide similarities (~ 88%) with Central European isolates. However, 19RH1 showed > 97%, 19SV showed > 95% and 20GO showed ≥ 92% similarities with central European isolates. 19RH1 had an in-frame three-nucleotide GCA deletion at nucleotide position 8412 to 8414 of coat protein, leading to the loss of a glycine residue at position 2761 in the polyprotein.

TriMV

The whole-genome sequences of 11 isolates TriMV were aligned with the complete genome sequence of the TriMV retrieved from the GenBank (Supplementary Table A5). TriMV isolates from this study showed high similarities (> 99%) with isolates from Colorado, Nebraska, and Kansas (Supplementary Table A5).

HPWMoV

The complete and near-complete genome sequences of eight RNA segments of HPWMoV isolates obtained in this study were aligned with reference genomes and other

published isolates retrieved from GenBank (Supplementary Table A8). All RNA segments had high nucleotide similarities (> 99%) with reference genomes, but 20MC2 RNA3A and RNA7 sequences showed 96.3% and 96.6% similarities, respectively, with the reference genome (Supplementary Table A7). Isolate 20RH2 RNA7 showed only 85.5% similarity with the RNA7 reference genome (Supplementary Table A7). Two variant sequences of RNA 3A and RNA 3B were found. RNA 3A encodes 286 amino acids of 33.2 kDa, and RNA 3B encodes for the 289 amino acids of 33.4 kDa nucleocapsid proteins. RNA 3A and RNA 3B variants found in this study had an average of within group and between group percent identity 12.43% sequence divergence between these variants. The alignment of protein sequences of 3A and 3B obtained from this study and with sequence retrieved from GenBank showed a 3-amino-acid insertion in RNA 3B at the positions of 23, 24, and 287 C-terminus of the protein in all isolates used (Supplementary Figure A1). These insertions differentiate RNA 3B from 3A.

SBWMV

The complete genome sequence of both RNA1 and RNA2 of SBWMV isolates obtained from Pawnee and Riley counties were aligned with RNA1 and RNA2 reference genomes separately (Supplementary Table A9). RNA1 genomes of SBWMV isolates from Pawnee and Riley County were more than 98% and 96% nucleotide similarity with the reference genome, respectively. RNA2 genomes of both isolates from Pawnee and Riley counties were more than 98% similar to the reference genome (Supplementary Table A9). RNA1 of Riley and Pawnee County isolates encodes three proteins: measuring 149.9/150 kDa (from 102 to 4064 nt, 1320 amino acids), 54.7 /54.6 kDa (from 4185 to 5588, 467 amino acids), and 37.2 /37.3 kDa (5653 – 6636, 327aa) respectively. RNA1 of Riley and Pawnee County isolates consists of 7096 and 6995 nucleotides respectively. The reference genome was 7099 nucleotides long, around 100 bps

were missing from the 3' untranslated region of Pawnee County isolates. The RNA2 is the shorter particle. Both counties' isolates contain 3590/3591 nucleotides and encode for three proteins: measuring 19.3 kDa (from 334 to 864 nt, 176 amino acids), 54.0 kDa (from 1141 to 2598, 458 amino acids), and 18.8 kDa (from 2665 to 3189, 174 amino acids).

Recombinant analysis of WSMV and TriMV

A 37.8% of total samples of WSMV (14 out of 37) were identified as potential recombinants by at least five algorithms of the RDP5 program at a significant value of $P < 0.05$ (Supplementary Table A10). These algorithms also provided the potential major and minor parents. These recombinant isolates were 19SV, 19ST, 19RA3, 19TR1, 19FI, 19GH1, 10SW, 20GO, 20WA, 20EW, 20TR2, 20GH2 and 20JW3. We also analyzed the recombinant analysis of TriMV isolates obtained from this study and reference genomes retrieved from the GenBank (Supplementary Table A6). No potential TriMV recombinant isolates were predicted by the RDP5 program.

Phylogenetic analysis

WSMV

The complete nucleoprotein or coding sequence of 23 WSMV isolates (9 from 2019 and 2020, 5 from 2021) from this survey and 22 reference isolates obtained from GenBank (Supplementary Table A6) were used to construct a phylogenetic tree. Constructing phylogenetic trees without recombination analysis leads to conflicting phylogenetic signals (Braidwood et al., 2019). Therefore, before phylogenetic analysis, recombinant analysis was performed, and 14 potential recombinant isolates detected by the RDP5 program were excluded from phylogenetic analysis. Oat necrotic mottle virus (ONMV) was used as an outgroup. The phylogenetic tree constructed from 45 complete nucleoprotein sequences of WSMV isolates consists of four main

clades (Figure 2.5). Clade A represents isolate ‘El- Batán’ from Mexico. Clade B includes European isolates characterized by a deletion of the glycine residue at position 2761 in the coat protein region because of the deletion of the GCA codon at nucleotide positions 8412 to 8414. The isolate (19RH1) collected from Rush County of KS was clustered with European isolates. Clade C represents an isolate from Iran. Clade D includes isolates from the United States, Argentina, and Turkey. Clade D isolates were further divided into four subclades (D1 to D4). D1 contained isolates from the American Pacific Northwest and WSMV type isolate, D2 was constituted by only isolates from this study. Isolates from Colorado, other already detected Kansas isolates and other Kansas isolates from this study comprised a small group and polytomies between sub-clades D1 and D2. D3 also contained the isolates from Kansas only with one already detected Kansas isolate. D4 included isolates from Kansas, Nebraska, Idaho, and Turkey.

TriMV

The phylogenetic tree was constructed using a complete nucleoprotein sequence with 11 TriMV Kansas isolates from this study (Supplementary Table A5) and 6 TriMV isolates from sequences retrieved from GenBank (Supplementary Table A6). Sugarcane mosaic virus (YN-YZ211) and Caladenia virus A (CalVA KP1) were used as outgroups. The topology of the tree consists of three clades: A, B, and C (Figure 2.6). Clade A consisted of single isolates from Colorado. Clade B contained two isolates, 19 MT and 20GL2, collected during this study. Clade C consisted of one isolate from Nebraska and four isolates previously collected from Kansas and 9 isolates collected during this study. Clade C comprises one subclade C1, including two isolates from Wichita and Lane County collected in 2021 and one from Seward County, Kansas collected in 2019.

HPWMoV

The coding sequence of HPWMoV nucleocapsid protein RNA3 and its two variants, RNA3A and RNA 3B were used to construct a phylogenetic tree (Figure 2.7). Five RNA3, three RNA3A, and RNA3B nucleocapsid protein sequences obtained from GenBank (Supplementary Table A8), four RNA3A, and three RNA3B sequences obtained from this study were included in the phylogenetic tree. Raspberry leaf blotch virus (RLBV) RNA3 nucleoprotein sequence was used as an outgroup. Because of the 95 - 99% within-group sequence identity and 87 - 89% between-group identity, RNA3 clustered separately in the middle of the topology between RNA3A and RNA3B. RNA3A isolates from this study and previously sequenced Nebraska and Kansas isolates were clustered together with a common node of significant bootstrap support (Figure 2.7). RNA3B GG1 Ohio isolates form a separate cluster. However, RNA3B isolates from this study (20MC2 and 20SC2) and previously sequenced Kansas isolate (KS7) clustered together. One Nebraska isolate, and 20KE2 isolate, formed a single polytomy within the RNA3B cluster.

Discussion

Oxford Nanopore sequencing (ONT) has great potential to identify viral diseases in wheat field samples and could be used as a disease diagnosis tool (Fellers et al., 2019). ONT is a promising method for generating high throughput long reads of both DNA/RNA in real-time with no prior amplification (Ayub et al., 2013; Piombo et al., 2021). This method has recently been used as a long-read sequencing technique to characterize many plant viruses including maize viruses (Adams et al., 2017), potato virus Y (Della Bartola et al., 2020), plum pox virus (Bronzato Badial et al., 2018), yam viruses (Filloux et al., 2018), and wheat viruses (Fellers et al., 2019). Cost per sample processing can be a major constraint for diagnostic labs to adopt

technology. One way of reducing the cost of ONT is to evaluate multiple samples in a single run using barcodes. In this study, we loaded up to 12 barcoded samples per MinION flow cell run and obtained a total of 5 to 16.64 million reads. Improved protocols using ONT as a diagnostic tool for plant viruses have been conducted, including use of a ribodepletion kit to remove host ribosome and increase coverage of viral pathogen (Liefting et al. 2021). ONT could be a feasible method for accurate diagnosis of plant viruses and could address the shortcoming of target-specific methods including the requirement of prior knowledge of the genome, limited ability to detect multiple pathogens simultaneously, and low sensitivity.

Aside from disease surveillance, ONT is a powerful choice for viral metagenomics or pathogenomic and transcriptomic studies. ONT generates long reads that potentially reduce the biases that arise during metagenomics studies (Wommack et al., 2008). The primary bias in using short-read sequencing platforms is the likelihood of generating chimeras from short reads from different virus variants during *de novo* assembly (Roux et al., 2017; Wommack et al., 2008). In contrast, the long-reads of ONT reduce the chances of having unassigned reads during contig assembly and increase the probability of getting true (unbiased) genetic variability by adjusting requirements during sequencing and post-sequencing bioinformatics pipeline (Filloux et al., 2018). ONT still suffers from high error rates and inconsistencies in coverage or number of reads while loading multiple samples, but high accuracy base calling packages and downstream computational methods are available to correct ONT sequencing data for deep sequence analysis and metagenomics (Sahlin and Medvedev, 2021). The ONT platform shows great promise for downstream analysis, including genetic modifications, recombination, adaptive evolution or resistant breaking mutations, and metagenomics (Brown et al., 2017).

Plant viral diagnostics is complicated by mixed infections of multiple viruses. Accurate diagnosis of plant viruses is essential to reduce disease spread and manage them effectively. This study identified positive-sense ssRNA viruses, bipartite positive-sense RNA viruses, a tripartite positive-sense RNA virus, and an octapartite negative-sense RNA virus in a single sample. This result shows ONT is an efficient means of detecting multiple viruses with different genomic structures in a mixed infection. Co-infection of plant viruses in a single plant is common, leading to a synergistic negative impact on the host (Sanfaçon, 2017; Syller, 2012; Tatineni et al., 2021). Maize lethal necrosis, maize chlorotic mottle virus and sugarcane mosaic virus cause synergistic impacts (Redinbaugh and Stewart, 2018). Synergistic interactions between isolates of casava mosaic virus have been documented (Pita et al., 2001). Past research has shown that co-infection increases vector transmission efficiency and systemic movement and gave a fitness advantage due to the synergistic effect of co-infection and high titers of WSMV and TriMV (Seifers et al., 2009; Tatineni et al., 2010). Additionally, the transmission efficiency of TriMV and HPWMoV mixed infection with WSMV was higher compared to a single infection of these viruses (Seifers et al., 2009, 2002). Mixed infection leads to mutual benefits between the co-infected viruses and causes greater impact on yield loss. Greenhouse and field studies showed significant yield loss as a result of co-infection of WSMV and TriMV (Byamukama et al., 2012; Wegulo et al., 2012). In this study, we targeted wheat RNA viruses by using Oligo(dT) primer and could also identify the viruses with “multiple adenylation” in the genome other than viral genomes with poly(A) tails. However, future studies using random primers, polyA tailing, and other ONT kits such as “What’s In My Pot” (EPI2ME WIMP workflow) could diagnose multiple RNA and DNA viruses in a single run.

Understanding viral population structure would help us develop or recommend better sources of genetic resistance. These results show that the field viral populations are very diverse. Wheat breeding programs mainly focused on screening the wheat cultivars with a single virus, especially wheat streak mosaic virus. With this knowledge of the diversity of field viral populations, breeders should consider screening their cultivars with multiple viruses. The selection of virus resistance cultivars through varietal screening programs should be dynamic based on frequent monitoring and characterization of the field viral population. Screening nurseries can mechanically inoculate with multiple viruses or set their breeding nursery in the field with a history of uniform natural infections. For best natural inoculation with field viral populations, one can establish the nursery next to the field with volunteer wheat or simulate the volunteer wheat around the varietal screening nursery.

The sampling method of this study targeted symptomatic wheat leaves. Additionally, sampling was done mostly from the edge of the fields. Therefore, this study was inadvertently biased towards multiple infection. However, future studies with a random sampling of symptomatic and neighboring non-symptomatic wheat leaves with a thorough investigation of the distribution might help fully understand the distribution of wheat viruses in the wheat growing reasons of Kansas.

Some samples of this work were collected from cultivars with known resistance genes from Lane and Wichita counties. We observed that wheat cultivars with *Wsm2*, such as ‘Joe’ and ‘Guardian’ (*Wsm2* + curl mite resistance gene) were heavily infected with high virus load. The counties where WSM complex viruses and Brome mosaic virus was identified were superimposed with the area of adaption of these varieties. Both varieties were infected with WSMV and TriMV. This result leads to further investigation of whether that is due to high

temperatures, as *Wsm2* is temperature sensitive, meaning resistance does not work effectively at higher temperatures (Seifers et al., 2007, 2006), or due to the presence of resistant breaking isolates of WSMV. One *Wsm2* resistant breaking isolate (KSH294) was already identified from foxtail in 2019 at Hays Kansas (Kumssa et al., 2019). Therefore, the presence of similar isolates in the field is possible. A growth chamber study to evaluate the infection of the isolates identified from 'Joe' and 'Guardian' in different temperature regimes will provide information on the effect of temperature or the ability of the isolates to overcome *Wsm2* resistance at temperatures below 24°C.

The deployment of wheat cultivars with a single resistance gene in a large geographic area increases the selection pressure on viruses. Joe has been cultivated in Kansas as one of the top five most planted wheat cultivars by the total area of plantation. Phylogenetic analysis of WSMV showed the diversity among Kansas isolates as they clustered into separate clades and sub-clades. This phylogenetic relationship of WSMV isolates is consistent with previous phylogenetic clustered WSMV isolates based on the CP sequence (Rabenstein et al., 2002; Stenger et al., 2002) and recent full genome sequence (Redila et al., 2021). The isolate collected from Rush County (19RH1) was clustered in Clade B with the European isolates. European isolates were reported from the Pacific Northwest region of the United States (Robinson and Murray 2013) and also from Great Plains (Redila et al., 2021). In the applied aspect, the diversity of WSMV isolates, including the European isolate and the putative recombinants in Kansas, could have significant implications for breeding programs because different WSMV isolates may interact differently with resistance genes. Regarding selection pressure, the presence of genetically variable WSMV isolates in Kansas could increase the likelihood of the evolution of resistance-breaking isolates. Breeders should consider using tolerant cultivars or cultivars with

stack of minors and major resistance genes to reduce the selection pressure in viral populations. A high virus load with no to little impact on overall yield is not much of great concern. This mechanism is less likely to break down as it causes less selection pressure. Future research targeting tolerant cultivars in breeding programs or in-depth investigation of the interaction of wheat viruses with tolerant cultivars would help make conclusive recommendations toward wheat virus management.

Recombination is a common phenomenon in plant viruses for evolutionary advantage. Evidence of multiple recombinants were reported from other potyviruses including plum pox virus (Cervera et al., 1993), isolates of potato potyvirus Y (Glais et al., 2002; Revers et al., 1996), bean common mosaic virus, zucchini yellow mosaic virus (Revers et al., 1996), and watermelon mosaic virus (Desbiez and Lecoq, 2008). In this study, recombinant analysis showed that about 38% of the total characterized isolates were recombinants. The putative recombinant isolates from this study were recombined with isolates from Europe (Clade B) and with isolates from United States (Clade D) as the major and minor parents (Supplementary Table A10). The presence of recombinant isolates in the field is concerning as the *Wsm2* resistant breaking isolate (KSH294) was reported as a potential recombinant isolate (Redila et al., 2021). This result suggests the recombination plays important role in the evolution of WSMV strains that might cause infection in cultivars of wheat with resistance genes. Additionally, the possibility of expanding the host range or evolving to be more aggressive or virulent exists. However, there was no information of host resistance genes for those recombinant isolates reported in this study. Future studies require a systemic survey of wheat cultivars with known resistance genes or defense mechanisms to understand the virus-virus-host interaction.

From this study, the whole genome sequence of all eight segments of HPWMoV was obtained from different Kansas counties. Among RNA3 variants sequenced in this study, RNA3B was more diverse than RNA3A. The sequence divergence between RNA3A and 3B and within-group and between-group identity among RNA3A, and 3B is consistent with the sequence divergence between RNA 3A and 3B reported in Ohio isolates (Stewart, 2016) and Nebraska isolates (Tatineni et al., 2014). The phylogenetic relationship showed Kansas and Nebraska isolates have more similarities than Ohio isolates. However, RNA3A isolates obtained from this study and isolates from Nebraska and Ohio clustered together, showing lower variability in RNA3A compared to RNA3B. The comparison of isolates from these three states shows the location-wise variability of HPWMoV RNA3B, but that is not true with RNA3A. In contrast, Stewart 2016 showed that neither location nor collection host was useful in predicting the nucleoprotein variability of HPWMoV. A significant number of HPWMoV isolates from different hosts, locations, and vector types would be an interesting future comparison that helps better understand the variability in HPWMoV isolates.

The whole genomes of RNA1 and RNA2 of SBWMV isolate characterized in this study are the two first isolates ever characterized from Kansas. The molecular characterization aids in the correct diagnosis and any future phylogenetic and evolutionary studies of SBWMV. Accurate diagnosis of SBWMV is crucial because the vector of this virus, *Polymyxa graminis,* produces resting spores that can remain dormant and invasive in soil for up to 30 years and then infect the host in favorable conditions (Cadle-Davidson and Gray, 2006). The vector can be easily distributed through water and farm machinery (Niblett et al. 1976). It is difficult to control this disease once it infects a wheat field. Therefore, molecular characterization of SBWMV isolates helps accurate diagnosis and helps to develop resistant cultivars through a breeding program.

Overall, the results obtained from this study demonstrate the potential cost-effective use of third generation long reads ONT to analyze multiple samples in a single run using barcoded samples. We could identify wheat viruses belonging to six families and eight genera: family *Potyviridae* (Genera *Tritimovirus*, *Poacevirus*, and *Bymovirus*), family *Fimoviridae* (Genus *Emaravirus*), family *Tombusviridae* (Genus *Luteovirus*), family *Solemoviridae (*Genus *Polerovirus*), family *Virgaviridae* (Genus F*urovirus*), family *Bromoviridae* (Genus *Bromovirus*) and with mono, bi, tri, and octa-partite positive to negative-sense RNA viruses in a single sample. The information of diverse wheat virus populations helps design sustainable management strategies of wheat viruses through genetic resistance.

References

Adams, I., Braidwood, L., Stomeo, F., Phiri, N., Uwumukiza, B., Feyissa, B., Mahuku, G., Wangi, A., Smith, J., Mumford, R., others, 2017. Characterizing maize viruses associated with maize lethal necrosis symptoms in sub-Saharan Africa. bioRxiv 161489.

Appel, J., Dewolf, E., Todd, T., Bockus, W., 2015. Preliminary 2015 Kansas Wheat Disease Loss Estimates; Kansas Cooperative Plant Disease Survey Report; Kansas Department of Agriculture: Manhattan, KS, USA, 2015.

Ayub, M., Hardwick, S.W., Luisi, B.F., Bayley, H., 2013. Nanopore-based identification of individual nucleotides for direct RNA sequencing. Nano letters 13, 6144–6150.

Brakke, M., 1987. Virus diseases of wheat. Wheat and wheat improvement 13, 585–624.

Bronzato Badial, A., Sherman, D., Stone, A., Gopakumar, A., Wilson, V., Schneider, W., King, J., 2018. Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. Plant disease 102, 1648–1652.

Brown, B.L., Watson, M., Minot, S.S., Rivera, M.C., Franklin, R.B., 2017. MinION^{TM} nanopore sequencing of environmental metagenomes: a synthetic approach. Gigascience 6, gix007.

Burrows, M., Franc, G., Rush, C., Blunt, T., Ito, D., Kinzer, K., Olson, J., O'Mara, J., Price, J., Tande, C., others, 2009. Occurrence of viruses in wheat in the Great Plains region, 2008. Plant Health Progress 10, 14.

Byamukama, E., Seifers, D., Hein, G., De Wolf, E., Tisserat, N., Langham, M., Osborne, L., Timmerman, A., Wegulo, S., 2013. Occurrence and distribution of Triticum mosaic virus in the central Great Plains. Plant disease 97, 21–29.

Byamukama, E., Tatineni, S., Hein, G.L., Graybosch, R.A., Baenziger, P.S., French, R., Wegulo, S.N., 2012. Effects of Single and Double Infections of Winter Wheat by *Triticum mosaic virus* and *Wheat streak mosaic virus* on Yield Determinants. Plant Disease 96, 859–864.

Byamukama, E., Wegulo, S., Tatineni, S., Hein, G., Graybosch, R., Baenziger, P.S., French, R., 2014. Quantification of yield loss caused by Triticum mosaic virus and Wheat streak mosaic virus in winter wheat under field conditions. Plant disease 98, 127–133.

Cadle-Davidson, L., Gray, S., 2006. Soilborne wheat mosaic virus. The Plant Health Instructor.

Campbell, L., Heyne, E., Gronau, D., Niblett, C., others, 1975. Effect of soilborne wheat mosaic virus on wheat yield. Plant Disease Reporter 59, 472–476.

Cervera, M.T., Riechmann, J.L., Martín, M.T., García, J.A., 1993. 3′-Terminal sequence of the plum pox virus PS and ŏ6 isolates: evidence for RNA recombination within the potyvirus group. Journal of General Virology 74, 329–334.

Choudhury, S., Larkin, P., Meinke, H., Hasanuzzaman, M., Johnson, P., Zhou, M., 2019. Barley yellow dwarf virus infection affects physiology, morphology, grain yield and flour pasting properties of wheat. Crop and Pasture Science 70, 16–25.

Cunfer, B.M., Demski, J.W., Bays, D.C., others, 1988. Reduction in plant development, yield, and grain quality associated with wheat spindle streak mosaic virus. Phytopathology 78, 198–204.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C., 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34, 2666–2669.

Della Bartola, M., Byrne, S., Mullins, E., 2020. Characterization of potato virus Y isolates and assessment of nanopore sequencing to detect and genotype potato viruses. Viruses 12, 478.

Desbiez, C., Lecoq, H., 2008. Evidence for multiple intraspecific recombinants in natural populations of Watermelon mosaic virus (WMV, Potyvirus). Archives of Virology 153, 1749–1754.

Fellers, J., Webb, C., Fellers, M., Shoup Rupp, J., De Wolf, E., 2019. Wheat virus identification within infected tissue using nanopore sequencing technology. Plant Disease 103, 2199-2203.

Filloux, D., Fernandez, E., Loire, E., Claude, L., Galzi, S., Candresse, T., Winter, S., Jeeva, M., Makeshkumar, T., Martin, D.P., others, 2018. Nanopore-based detection and characterization of yam viruses. Scientific reports 8, 1–11.

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

Glais, L., Tribodet, M., Kerlan, C., 2002. Genomic variability in Potato potyvirus Y (PVY): evidence that PVYNW and PVYNTN variants are single to multiple recombinants between PVYO and PVYN isolates. Archives of virology 147, 363–378

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology 52, 696–704.

Hodge, B., Paul, P., Stewart, L.R., 2020. Occurrence and High-Throughput Sequencing of Viruses in Ohio Wheat. Plant disease 104, 1789–1800.

Hodge, B., Salgado, J., Paul, P., Stewart, L., 2019. Characterization of an Ohio isolate of Brome mosaic virus and its impact on the development and yield of soft red winter wheat. Plant disease 103, 1101–1111.

Hollandbeck, F. G., Dewolf, E., and Todd, T. 2017. Kansas Cooperative Plant Disease Survey Report, Preliminary 2017 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Hollandbeck, F. G., Dewolf, E., and Todd, T. 2020. Kansas Cooperative Plant Disease Survey Report, Preliminary 2020 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Hollandbeck, F. G., Onofre, K. A., Dewolf, E., and Todd, T. 2021. Kansas Cooperative Plant Disease Survey Report, Preliminary 2021 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

KS wheat. 2017. Kansas wheat. Wheat streak mosaic is timely issue at wheat. http://kswheat.com/news/2017/08/09/wheat-streak-mosaic-is-timely-issue-at-wheatu-event (accessed on 2/12/2022)

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution 35, 1547.

Kumssa, T., Rupp, J., Fellers, M., Fellers, J., Zhang, G., 2019. An isolate of Wheat streak mosaic virus from foxtail overcomes Wsm2 resistance in wheat. Plant Pathology 68, 783–789.

Lam, H.M., Ratmann, O., Boni, M.F., 2018. Improved algorithmic complexity for the 3SEQ recombination detection algorithm. Molecular biology and evolution 35, 247–251.

Liefting, L.W., Waite, D.W., Thompson, J.R., 2021. Application of Oxford Nanopore Technology to Plant Virus Detection. Viruses 13, 1424

Martin, D., Posada, D., Crandall, K., Williamson, C., 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Research & Human Retroviruses 21, 98–102.

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

Martin, D.P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab, T., Kelz, Z., Kumar, V., Murrell, B., 2021. RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus evolution 7, veaa087.

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford university press.

Ordon, F., Habekuss, A., Kastirr, U., Rabenstein, F., Kühne, T., 2009. Virus resistance in cereals: sources of resistance, genetics and breeding. Journal of phytopathology 157, 535–545.

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

Phannareth, T., Nunziata, S.O., Stulberg, M.J., Galvez, M.E., Rivera, Y., 2020. Comparison of Nanopore Sequencing Protocols and Real-Time Analysis for Phytopathogen Diagnostics. Plant Health Progress 22, 31–36.

Piombo, E., Abdelfattah, A., Droby, S., Wisniewski, M., Spadaro, D., Schena, L., 2021. Metagenomics approaches for the detection and surveillance of emerging and recurrent plant pathogens. Microorganisms 9, 188.

Pita, J., Fondong, V., Sangare, A., Otim-Nape, G., Ogwal, S., Fauquet, C., 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. Journal of General Virology 82, 655–665.

Posada, D., Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proceedings of the National Academy of Sciences 98, 13757–13762

Rabenstein, F., Seifers, D.L., Schubert, J., French, R., Stenger, D.C., 2002. Phylogenetic relationships, strain diversity and biogeography of Tritimoviruses. Journal of General Virology 83, 895–906.

Redila, C.D., Phipps, S., Nouri, S., 2021. Full Genome Evolutionary Studies of Wheat Streak Mosaic-Associated Viruses Using High-Throughput Sequencing. Frontiers in Microbiology 12, 1998.

Redinbaugh, M.G., Stewart, L.R., 2018. Maize Lethal Necrosis: An Emerging, Synergistic Viral Disease. Annual review of virology 5, 301–322

Revers, F., García, J.A., 2015. Molecular biology of potyviruses, in: Advances in Virus Research. Elsevier, pp. 101–199.

Robinson, M.D., Murray, T.D., 2013. Genetic variation of Wheat streak mosaic virus in the United States Pacific Northwest. Phytopathology 103, 98–104.

Rotenberg, D., Bockus, W.W., Whitfield, A.E., Hervey, K., Baker, K.D., Ou, Z., Laney, A.G., De Wolf, E.D., Appel, J.A., 2016. Occurrence of viruses and associated grain yields of paired symptomatic and nonsymptomatic tillers in Kansas winter wheat fields. Phytopathology 106, 202–210.

Roux, S., Emerson, J.B., Eloe-Fadrosh, E.A., Sullivan, M.B., 2017. Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. PeerJ 5, e3817.

Sahlin, K., Medvedev, P., 2021. Error correction enables use of Oxford Nanopore technology for reference-free transcriptome analysis. Nature Communications 12, 1–13.

Sanfaçon, H., 2017. Grand challenge in plant virology: understanding the impact of plant viruses in model plants, in agricultural crops, and in complex ecosystems. Frontiers in microbiology 8, 860.

Seifers, D., Martin, T., Harvey, T., Haber, S., 2007. Temperature-sensitive Wheat streak mosaic virus resistance identified in KS03HW12 wheat. Plant disease 91, 1029–1033.

Seifers, D., Martin, T., Harvey, T., Haber, S., Haley, S., 2006. Temperature sensitivity and efficacy of Wheat streak mosaic virus resistance derived from CO960293 wheat. Plant Disease 90, 623–628.

Seifers, D.L., Harvey, T.L., Louie, R., Gordon, D., Martin, T., 2002. Differential transmission of isolates of the High Plains virus by different sources of wheat curl mites. Plant disease 86, 138–142.

Seifers, D.L., Martin, T., Harvey, T.L., Fellers, J.P., Michaud, J., 2009. Identification of the wheat curl mite as the vector of Triticum mosaic virus. Plant Disease 93, 25–29

Smith, J.M., 1992. Analyzing the mosaic structure of genes. Journal of molecular evolution 34, 126–129.

Stenger, D.C., Seifers, D.L., French, R., 2002. Patterns of polymorphism in Wheat streak mosaic virus: sequence space explored by a clade of closely related viral genotypes rivals that between the most divergent strains. Virology 302, 58–70.

Stewart, L.R., 2016. Sequence diversity of wheat mosaic virus isolates. Virus research 213, 299–303.

Syller, J., 2012. Facilitative and antagonistic interactions between plant viruses in mixed infections. Molecular plant pathology 13, 204–216.

Tatineni, S., Alexander Jeff, Qu Feng, 2021. Differential Synergistic Interactions among Four Different Wheat-infecting Viruses. Frontiers in microbiology, 12 : 800318-800318.

Tatineni, S., Graybosch, R.A., Hein, G.L., Wegulo, S.N., French, R., 2010. Wheat cultivar-specific disease synergism and alteration of virus accumulation during co-infection with Wheat streak mosaic virus and Triticum mosaic virus. Phytopathology 100, 230–238.

Tatineni, S., McMechan, A.J., Wosula, E.N., Wegulo, S.N., Graybosch, R.A., French, R., Hein, G.L., 2014. An eriophyid mite-transmitted plant virus contains eight genomic RNA segments with unusual heterogeneity in the nucleocapsid protein. Journal of virology 88, 11834–11845

Wegulo, S., Byamukama, E., Tatineni, S., Hein, G., Graybosch, R., Baenziger, P., 2012. Effects of single and double infections of winter wheat by Triticum mosaic virus and Wheat streak mosaic virus on grain yield and yield components, in: phytopathology. American Phytopathological Society ST Paul, MN 55121 USA, pp. 134–134.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microbial genomics 3.

Wick, R.R., Judd, L.M., Holt, K.E., 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome biology 20, 1–10.

Wommack, K.E., Bhavsar, J., Ravel, J., 2008. Metagenomics: read length matters. Applied and environmental microbiology 74, 1453–1463.

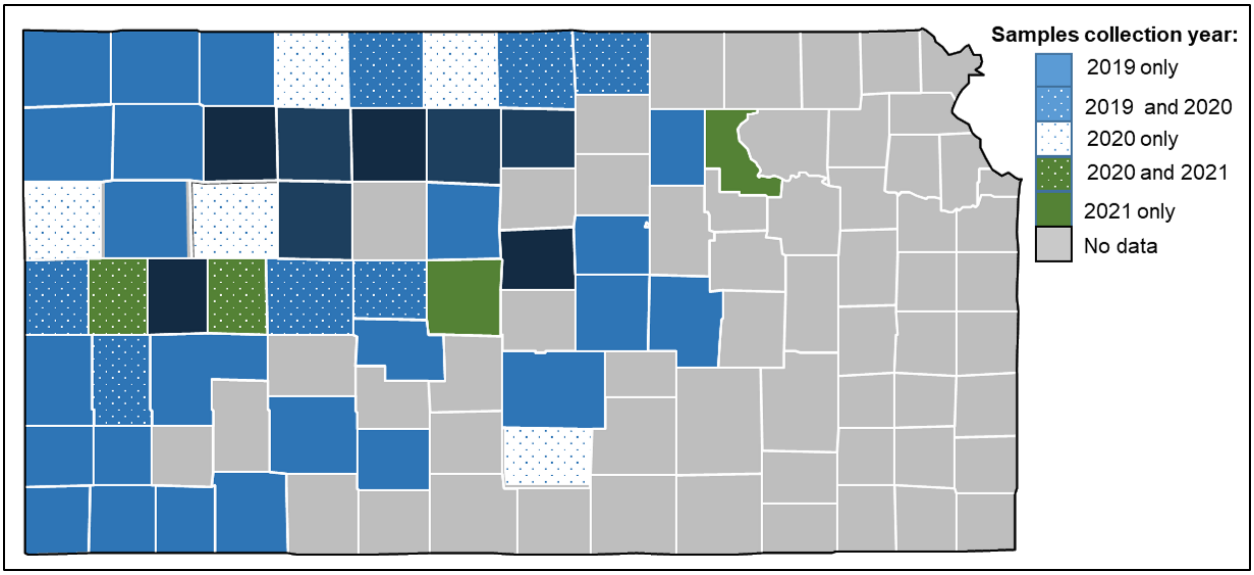


Figure 2.1. Map of Kansas counties where samples were collected in 2019, 2020, and 2021

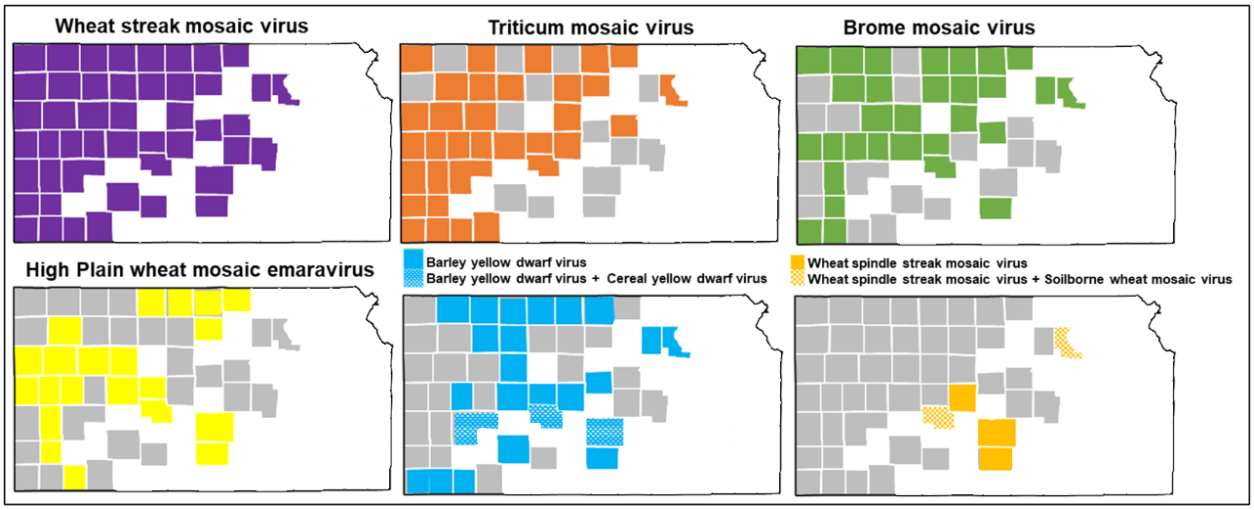


Figure 2.2. Map of Kansas counties with the viruses identified in this study using Oxford Nanopore sequencing. Virus-like symptomatic wheat leaves were collected from the field in 2019, 2020, and 2021. White area of the map indicates never sampled, gray area indicates sampled but negative result of a virus and colored area indicates positive result of a virus.

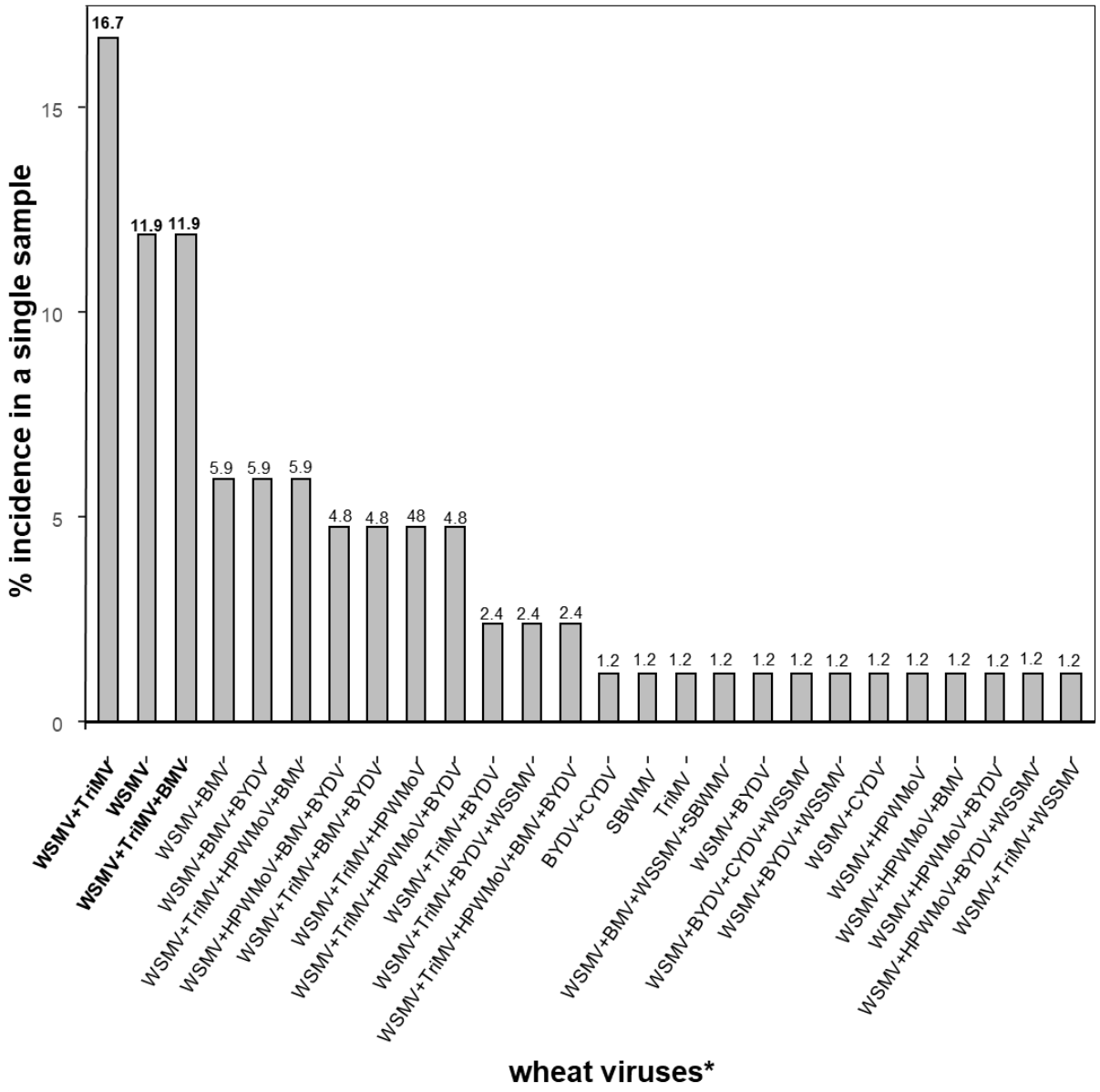


Figure 2.3. Percent incidence of wheat streak mosaic virus (WSMV), triticum mosaic virus (TriMV), High Plains wheat mosaic emaravirus (HPWMoV), brome mosaic virus (BMV), barley yellow dwarf virus (BYDV), wheat spindle streak mosaic virus (WSSMV), cereal yellow dwarf virus (CYDV), soilborne wheat mosaic virus (SBWMV) and virus combinations in samples collected from Kansas wheat fields detected through Oxford Nanopore sequencing. Virus-like symptomatic wheat leaves were collected from the field in 2019, 2020, and 2021.

Table 2.1. Individual virus detected through Oxford Nanopore sequencing. The percentage of incidence was measured in total samples if that individual virus was identified as a single infection or co-infected with other viruses. Virus-like symptomatic wheat leaves were collected from the field in 2019, 2020, and 2021.

Wheat virus detected in samples*	% sample incidence in a single sample (n = 84)
WSMV	96.43
TriMV	58.33
BMV	44.05
HPWMoV	27.38
BYDV	34.52
WSSMV	8.33
CYDV	3.57
SBWMV	2.38

*WSMV= wheat streak mosaic virus, TriMV = triticum mosaic virus, HPWMoV = High Plains wheat mosaic emaravirus, BYDV = barley yellow dwarf virus, BMV = brome mosaic virus, WSSMV = wheat spindle streak mosaic virus, CYDV = cereal yellow dwarf virus, SBWMV =soilborne wheat mosaic virus

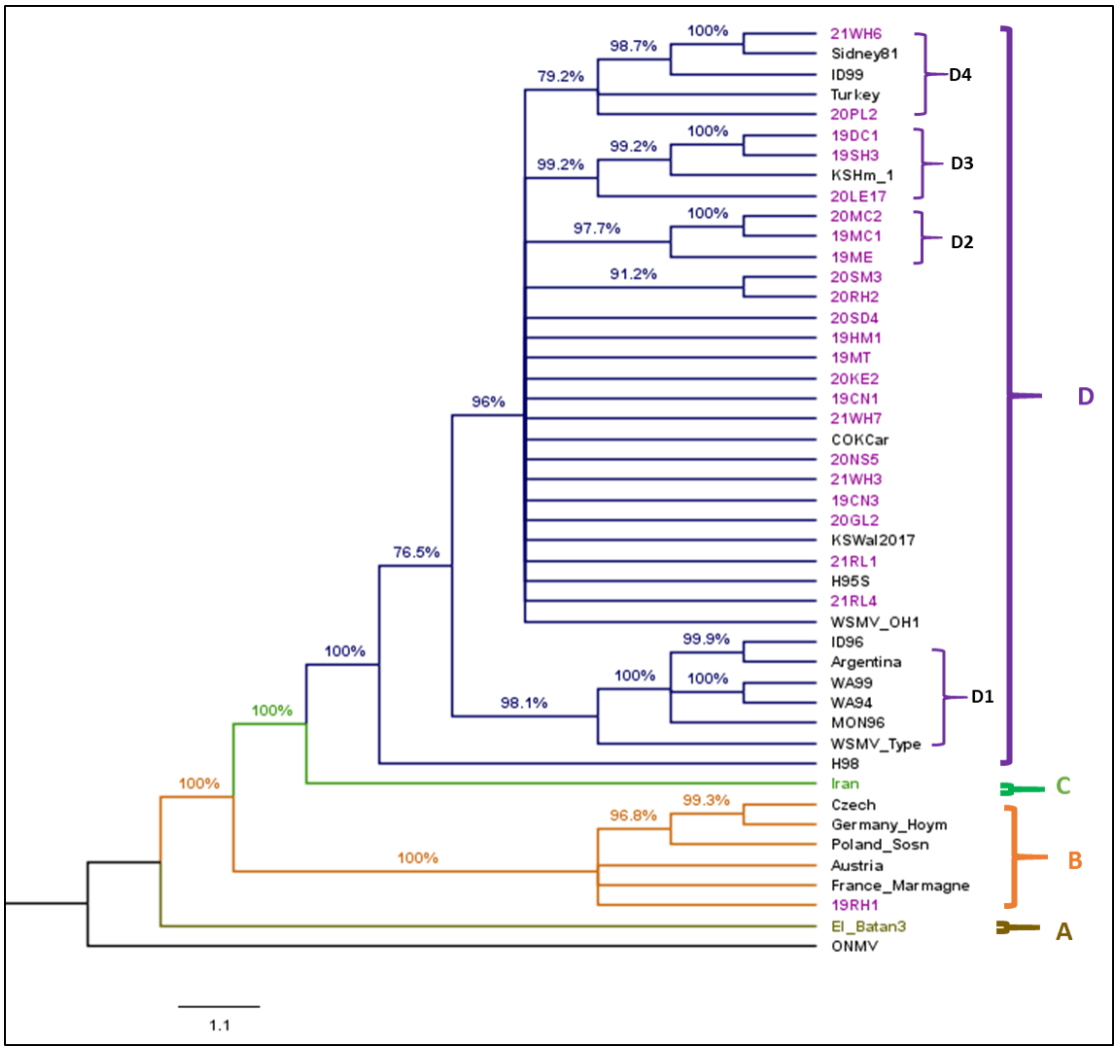


Figure 2.4. Phylogenetic tree of wheat streak mosaic virus (WSMV) isolates sequenced in this study (heightened in purple text) and selected strains. The phylogenetic tree was made with maximum likelihood analysis with a GTR + G + I substitution model of nucleoprotein sequence with 1000 bootstrap. The tree with the highest log likelihood (-56327.64) is shown. The percentage of trees in which the associated taxa were clustered is shown next to the branches. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Oat necrotic mottle virus was used as an outgroup in the analysis. Brackets on the right side indicate the taxa clustered in WSMV clades A to D. Clade D is further divided into subclades D1 to D4.

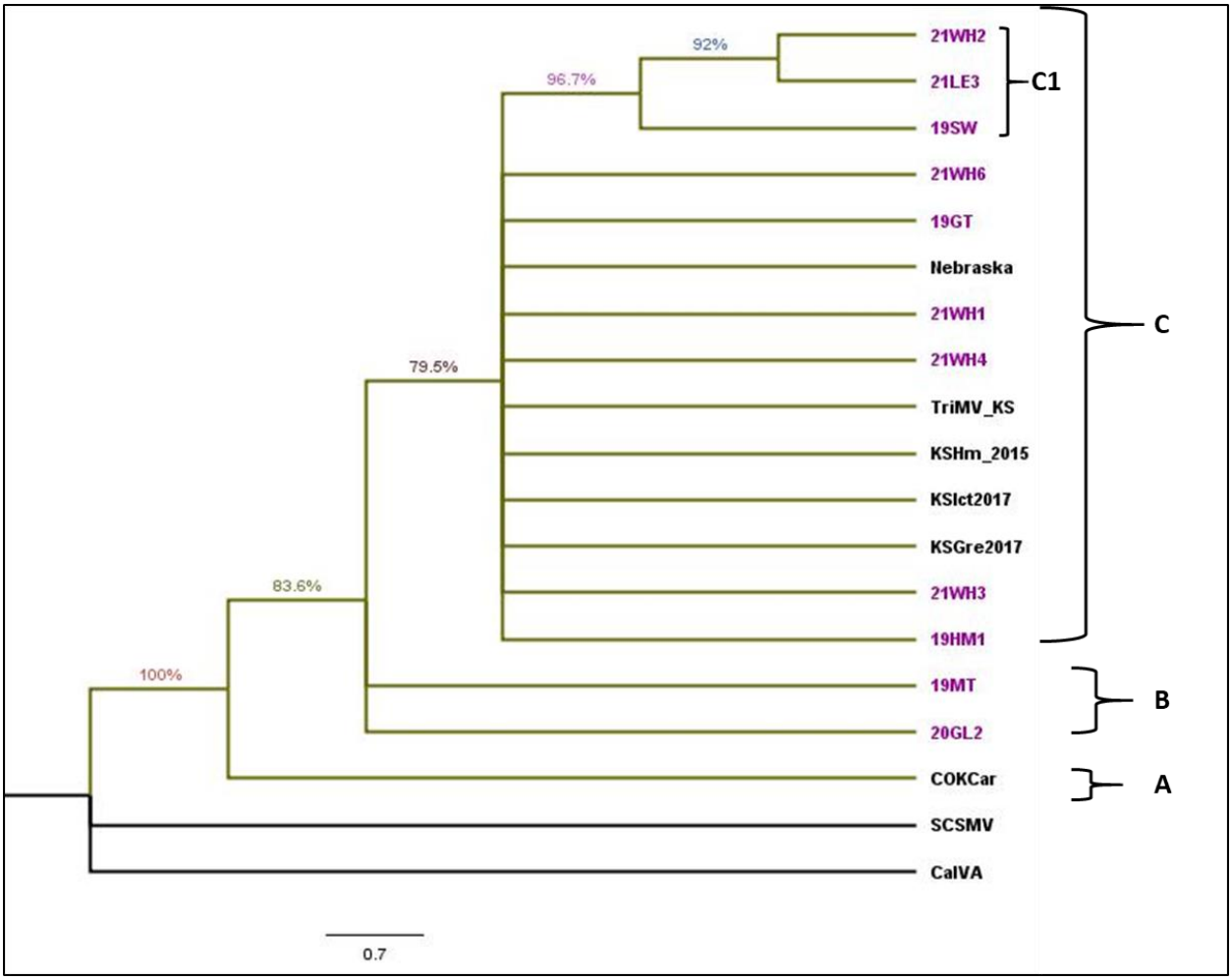


Figure 2.5**.** Phylogenetic tree of Triticum mosaic virus (TriMV) isolates sequenced in this study (heightened in purple text) and selected strains. TriMV isolates are divided into four clades A to C and clade C with C1 sub-clade. The phylogenetic tree was made with maximum likelihood analysis with a GTR + G substitution model of nucleoprotein sequence with 1000 bootstrap. The tree with the highest log likelihood (-25213.21) is shown. The percentage of trees in which the associated taxa were clustered is shown next to the branches. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Sugarcane streak mosaic virus and Caladenia virus A were used as outgroups in the analysis.

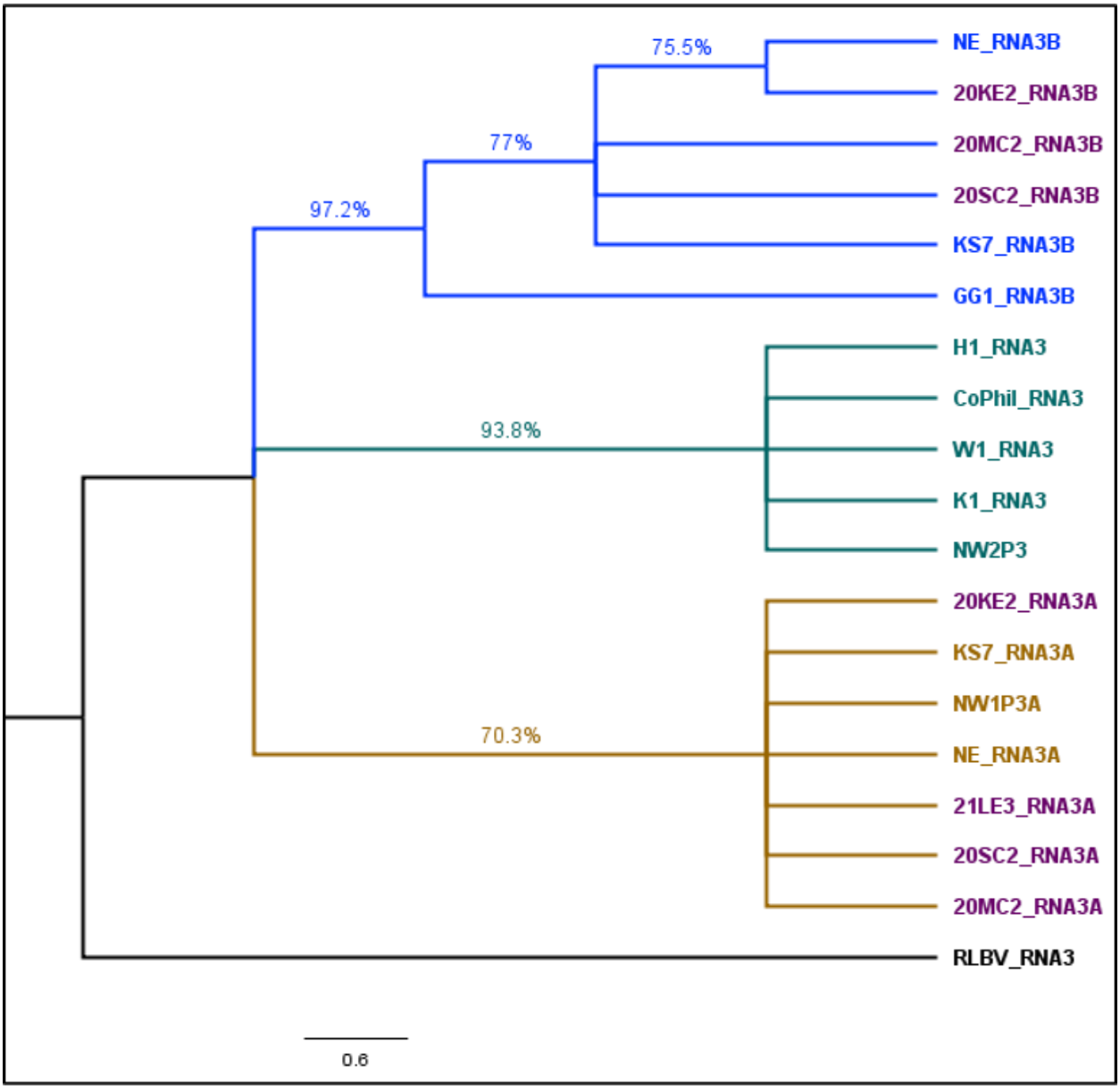


Figure 2.6. Phylogenetic tree of High Plains wheat mosaic emaravirus (HPWMoV) isolates sequenced in this study (heightened in purple text) and selected strains. The phylogenetic tree was made with maximum likelihood analysis with a T92 + G substitution model of nucleoprotein sequence with 1000 bootstrap. The tree with the highest log likelihood (-3324.14.21) is shown. The percentage of trees in which the associated taxa were clustered is shown next to the branches. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Raspberry leaf blotch virus (RLBV) was used as an outgroup in the analysis.

Chapter 3 - Genetic and evolutionary characterization of brome mosaic virus isolates and their association with other wheat viruses in Kansas

Abstract

Brome mosaic virus (BMV) was recently identified from the Kansas wheat fields in association with other wheat viruses. In this study, we report the details of BMV co-infection with other wheat viruses and compare the genetic variability and evolutionary characteristics of the complete genomes of the BMV isolates obtained using Nanopore sequencing. BMV was found co-infected with wheat streak mosaic virus (WSMV) and Triticum mosaic virus (TriMV) (27.8%), followed by co-infection with WSMV only (13.9%) and WSMV + TriMV + HPWMoV (High Plains wheat mosaic emaravirus) (13.9%). BMV isolates obtained in this study had greater nucleotide and amino acid sequence identity (99.4 to 100%) to each other. Clustering of BMV isolates varied while constructing phylogenetic trees based on nucleotide sequences of coding regions of each RNA genome. On average, US BMV isolates showed low divergence. RNA2a possessed relatively high divergence ($\pi = 0.01$) compared to RNA1a and RNA3a ($\pi = 0.004$). Coding regions of all RNAs of BMV were considered negative or purifying selection pressure as nonsynonymous, and synonymous nucleotide ratio was less than one (dNs/dS >1). The molecular characterization of Kansas BMV isolates aids in designing appropriate diagnostic tools, understanding genetic variation, phylogenetic relationships with the other US and non-US isolates, and evolutionary mechanisms employed by genome subunits of BMV would support the sustainable management of wheat viruses through genetic resistance.

Introduction

Brome mosaic virus (Genus: *Bromovirus*, family: *Bromoviridae*) is the type member of a group of icosahedral, positive-strand ssRNA viruses with a tripartite linear genome. The genome is comprised of RNA1, RNA2, and RNA3 (Ahlquist et al., 1984; Kao and Sivakumaran, 2000). RNA1 encodes protein 1a (containing capping and RNA helicase activities), RNA2 encodes protein 2a (putative RNA-dependent RNA polymerase), and RNA3 encodes two proteins: 3a (movement protein, MP and coat protein, CP). The CP is translated or derived from RNA3 as sub-genomic RNA and recognized as RNA4 (Kao and Sivakumaran, 2000; Rao, 2006). Brome mosaic virion encapsidate RNA1 and RNA2 separately, whereas RNA3 is encapsidated with sub-genomic RNA4 in a single virion (Rao, 2006), The capsid of all three particles contains 180 subunits arranged in icosahedral symmetry (Lucas et al., 2002).

Brome mosaic virus (BMV) distributed worldwide, as it has been reported in the United States (Hodge et al., 2019; Mian et al., 2005; Srivatsavai, 2005), Canada (Díaz-Cruz et al., 2018), South Africa (Von Wechmar and Rybicki, 1985), Estonia (Sõmera et al., 2016), Poland (Trzmiel et al., 2015), Lithuania (Urbanavičienė et al., 2012), Serbia (Tošič, 1971), Hungary (Pocsai et al., 1991), Great Britain (Gibson and Kenten, 1978), Brazil (Caetano et al., 1990), and Russia (Lane, 1974). BMV has wide host range and mainly infects grasses from the Poaceae family, including major crops such as wheat, barley, oats, corn, and sorghum, as well as dicot plants including soybean, common beans, faba beans, cowpea, tobacco, and *Nicotiana benthamiana* or *Chenopodium* species (Hodge et al., 2019; Kao and Sivakumaran, 2000; Lane, 1974; Trzmiel et al., 2016). The symptoms caused by this virus vary by plant species. In grasses, the distinct symptoms of BMV include yellow mosaic with light and dark green streaks similar to the symptoms caused by many cereal viruses.

The primary transmission route of BMV in plants is still ambiguous (He et al., 2021). However, the spread of the virus has been reported by several modes of transmission. Mechanical inoculation using plant sap prepared from BMV-infected leaves and purified virions or virion RNAs or *in vitro* transcripts from cloned viral cDNAs was commonly used in greenhouse transmission (He et al., 2021; Hodge et al., 2019; Srivatsavai, 2005). Farm machinery can also transmit the virus in the field mechanically (Lane, 1974; Mian et al., 2005). A low rate of transmission by vectors, including flea beetle (*Altica foliaceae*) (Srivatsavai, 2005), Russian wheat aphid (*Diuraphis noxia*), nematodes: *Longidorus breviannulatus* and *Xiphinema spp*. (Huff et al., 1987; Schmidt et al., 1963), and bird cherry-oat aphid (*Rhopalosiphum padi*) (Damsteegt et al., 1992; Rybicki and Von Wechmar, 1982) were recorded in greenhouse experiments with wheat or barley as a host. BMV transmission was associated with uredospores of wheat stem rust and claimed that BMV attached to the uredospores surface and spread in the fields (Erasmus et al., 1983). Infectious BMV was detected in water resources surrounding cereals fields, demonstrating that the virus can survive without its host and vector (Jeżewska et al., 2019).

BMV was studied as a model for RNA virus biology and as infectious cDNAs in recombinant DNA technology (He et al., 2021; Kao and Sivakumaran, 2000). However, despite the extensive studies on its use as cDNA clones and its basic biology, only a few studies have been conducted to evaluate the BMV incidence on economically important crops. For example, BMV was the dominant virus in wheat fields, among other wheat viruses in Hungary in 1994 – 95 (Papp et al., 1996), and an average 13% incidence of BMV samples collected from wheat fields in Alabama in 2004 (Srivatsavai, 2005). Greenhouse studies showed that BMV reduced wheat height, weight, and yields (Pocsai et al., 1991). Hodge et al., 2019 reported up to 61%

yield loss on soft red winter wheat when inoculated at early growth stages. BMV was detected with a high prevalence in wheat and showed a potentially high risk to wheat production in Ohio (Hodge et al., 2020). Mixed infection of multiple wheat viruses in a single plant compounds the risk, resulting in a synergistic yield reduction (Lane, 1974).

BMV co-infection has been reported with other cereal viruses from Hungary (Papp et al., 1996) and Ohio wheat fields (Hodge et al., 2020). As Kansas is one of the top wheat-producing states in the USA (USDA, National Agricultural Statistics Service, 2020 and 2021), any new threat to wheat production could lead to millions of dollars in lost productivity (Hollandbeck et al 2020 and 2021). Therefore, the wide distribution of BMV in Kansas wheat fields demands further study.

This study presents the detection and characterization of BMV isolates obtained from Kansas wheat using Oxford nanopore sequencing techniques (ONT). The information of BMV co-infection with other viruses provides the foundation for accurate diagnosis in mixed infection of multiple viruses and use of ONT for the dual purpose of surveillance and in-depth genetic and evolutionary characterization.

Materials and methods

Field survey

A field survey was conducted during the wheat growing season from May to July 2019 to 2021 in major wheat growing counties of Kansas. A total of 84 wheat leaves with yellow discoloration or mosaic patterns were sampled from 47 different counties of Kansas. Leaf tissue of each sample was stored at - 20°C until the tissue could be processed for RNA extraction.

RNA extraction

A mirVana miRNA extraction kit (Ambion Catalog number: AM1560, Thermo Fisher Scientific, MA, USA) was used to extract the total RNA from the sample (200 mg of tissue) following the company's instructions. After extraction, the concentration of RNA was measured by NanoDrop spectrophotometer (NanoDrop Technologies, Rockland, DE, USA). Seven µg of total RNA was treated with 1µl of DNase using Turbo DNase-Free ^{TM} kit (AM 1907, Ambion®, Thermo Fisher, MA, USA) in a 50 µl reaction volume according to the manufacturer's instruction to get rid of host genomic DNA.

Nanopore sequencing

An Oxford PCR-cDNA Barcoding kit (SQK-PCB109) was used to prepare the MinION cDNA library following the manufacturer's instruction (Oxford Nanopore Technologies, Oxford, U.K.) and with the following modifications. A total reaction volume of 11 µl was prepared with 1µl (100 to 120 ng/µl) total RNA, 1 µl of 2 µM VN primers (oligo dT VNP, SQK-PCB109, ONT), 1 µl of 10 mM dNTPs (Invitrogen, catalog number: 1875160) and 8 µl of nuclease-free water (NFW) and incubated at 65°C for 5 minutes. Then a total volume of 8 µl was obtained from mixing 4 µl 5x RT buffer (Invitrogen, catalog number 18090200), 1 µl of RNaseOUT (Invitrogen, catalog number 10777019), 2 µl of 10 µM Strand-Switching primer (SSP, SQK-PCB109, ONT), and 1 µl of NFW was added to the incubated mixture making total reaction volume of 19 µl and incubated at 42°C for 2 minutes. 1 µl of Maxima H minus Reserve Transcriptase (Invitrogen, catalog number 18090200), was added to make a total volume of 20 µl and incubated for 90 minutes at 42°C and followed by heat inactivation for 5 minutes at 85°C.

For PCR and barcoding, a 5 µl of reverse-transcribed RNA out of 20 µl was used and 1.5µl Barcode Primers (BP01 to BP12, SQK-PCB109, ONT) each barcode for each sample up to

12 samples, 18.5 µl of NFW, and 25 µl of 2x LongAmp Taq Master (New England Biolabs, catalog number #M0287) was added to make total reaction volume. The thermal cycler program was 95°C for 30secs for initial denaturation, 15 cycles of 95°C for 15secs, 62°C for 15secs, 65°C for 12 minutes, and final extension of 6 minutes at 65°C to amplify ~10kb products. 1 µl of NEB exonuclease 1 (New England Biolabs, catalog number #M0293) was added after completing PCR and incubated for 15 minutes at 37°C, followed by 80°C for 15 minutes. After completing the incubation and heating, 40 µl of AMPure XP beads (Beckman Coulter, #A63881) was added to the reaction and incubated for 5 minutes in a rotator mixer. Then the beads were washed with freshly prepared 70% ethanol following the manufacturer's instructions (Oxford Nanopore Technologies, Oxford, U.K.). The cDNA library was eluted in 12 µl of Elution buffer (EB, SQK-PCB109, ONT). After measuring the concentration of the cDNA library, the barcoded samples were pooled to a final volume of 11 µl and a 1 µl of Rapid Adaptor (RAP, SQK-PCB109, ONT) was added. Then the 12 µl total volume was incubated at 22°C for 15 minutes. Then the prepared library was loaded on the MinION R 9.4.1 flow cell (Oxford Nanopore) following the manufacturer's priming and loading instruction (Oxford Nanopore Technologies, Oxford, U.K.).

Bioinformatics

MinKNOW software (version: 21.06.13, Oxford Nanopore Technologies, Oxford, U.K.) provided the fast5 format data which was the raw electrical signal to the nucleotide sequence passing through the nanopore. These signals were translated to nucleotide bases by using the guppy basecaller high accuracy option (version: 3.1.5, dna_r8.4.1_450bps_hac.cfg, R 9.4.1, Oxford Nanopore) (Wick et al., 2019) to get fastq data. The fastq data obtained after basecalling were separated by barcode and the adapters were trimmed using porechop v0.2.3 (Wick et al., 2017). The small sequence (< 75 bp) and potential larger sequence length (max-length 30000)

were filtered by using Nanofilt v2.3.0 (De Coster et al., 2018). Thus, obtained reads were mapped against the brome mosaic virus reference genomes (Accession numbers: X02380.1, X01678.1, and J02042.1) to get the consensus sequence using CLC Genomic Workbench® v21.0.4 (Qiagen, MD, United States). The following parameters were used in CLC workbench, reads were mapped to reference using the parameters following resequencing analysis with masking mode = no masking, match score = 1, mismatch cost = 2, cost of insertions and deletions = Linear gap cost, insertion cost = 3, deletion cost = 3, length fraction = 0.5, similarity fraction = 0.8, global alignment = no, non-specific match handling = map randomly, output mode = create stand-alone read mappings, create report = yes, collect unmapped reads = no. The consensus sequences with low coverage were blasted in the NCBI nucleotide (blastn) database to confirm the presence of BMV. The minimum sequence read length was 1000 bp were considered as lower limit to count.

Sequence alignment, percent identity, and similarity

The coding region of each protein (RNA1a, RNA2a, MP, and CP) of BMV isolates sequenced in this study and selected isolates from GenBank (Supplementary Table B1) were aligned using Muscle alignment in Mega X (Kumar et al., 2018) with default parameters (Max Iterations 16, Cluster Method) separately. Aligned sequences were analyzed to obtain percent identity using an Multiple Sequence Alignment ‘MUSCLE’ online program supported by EMBL-EBI (Edgar, 2004). The amino acid sequence alignments obtained from Mega X were analyzed using the SMS (Sequence Manipulation Suite) online program available through bioinformatics.org (Stothard, 2000a) to get amino acid percent identity and similarity of the coding regions of each protein of BMV.

Phylogenetic analysis

Alignment of coding regions of RNA1a, RNA2a, RNA3 (MP and CP) sequences of BMV was performed with a muscle program in Mega X (Kumar et al., 2018) including outgroups. The Cassia yellow blotch virus (CYBV) and Olive latent virus 2 (OLV) were used as outgroups (Supplementary Table B1). To construct the phylogenetic tree, the best fitting nucleotide substitution models were determined by the maximum likelihood fits (Kumar et al., 2018; Nei and Kumar, 2000). The best-fitted model was selected based on the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) scores (Guindon and Gascuel, 2003). These models were TN93 + G (Tamura-Nei substitution model with Gamma distributed rate) for RNA1a, HKY + G (Hasegawa-Kishino-Yano model with Gamma distributed rate) for RNA2a, K2 + I (Kimura-2-parameter model with Invariant sites) for MP and CP of BMV. Maximum likelihood phylogenetic trees were constructed using Mega X with parameters as follows: the number of Bootstrap replications of 1000, nucleotide substitution model as mentioned above for coding regions of different RNAs, and the number of threads of 4.

Population genetics analysis

The population genetic analysis was done with the BMV US isolate sequences obtained from this study and isolates with complete coding sequence available in GenBank. The program DnaSP version 5.10 (Librado and Rozas, 2009) was used to analyzed the population genetic parameters including the number of segregating sites (s), the total number of mutations (η), nucleotide diversity (π), and mutation rate (θw) were calculated on the protein-coding sequence of RNA1a (five isolates: 20SM3, 19RP1, BMV_OK, BMV_M1, and BMV_M2), of RNA2a and RNA4 (seven isolates: 20SM3, 19RP1, BMV_OK, BMV_OH, BMV_OH2, BMV_M1, and BMV_M2) and for RNA3a (12 isolates; 7 isolates obtained from this study and five BMV_OK,

BMV_OH, BMV_OH2, BMV_M1, and BMV_M2). The program MEGA X (Kumar et al., 2018) was used to estimate the non-synonymous substitutions (dN), and synonymous substitutions (dS) and their ratio (dN/dS = ω) using the bootstrap variance estimation method with 1000 replicates under the model of Kumar method (Kimura 2-para) for each encoded protein.

Results

Co-infection of brome mosaic virus

Brome mosaic virus was identified from 29 different counties of Kansas out of 48 counties sampled (Figure 3.1). 44% (37 out of 84 total samples) of the samples processed using ONT were positive for BMV. A chord interaction diagram shows the co-infection of BMV with six other wheat viruses (Figure 3.2). Out of these 37 positive samples, BMV was found co-infected mostly with WSMV and TriMV (27.8%), followed by co-infected with WSMV only (13.9%) and co-infected with WSMV + TriMV + HPWMoV (13.9%) (Figure 3.3). BMV was found co-infected with BYDV along with WSMV, TriMV, and HPWMoV (11.1%). Approximately 3% of the BMV positive samples were identified co-infected with wheat spindle streak mosaic virus (WSSMV) and soilborne wheat mosaic virus (SBWMV) (Figure 3.3).

Sequencing BMV genome

We obtained an average of 4.52×10^5 raw reads from 37 samples obtained from Nanopore sequencing (Supplementary Table B2). After mapping with brome mosaic virus reference genome (Supplementary Table B3), an average of 1.43×10^5 reads of BMV were obtained with variable coverage by the sample having an average coverage of 40717.19X (Supplementary Table B3). Complete genome and complete nucleoprotein sequence of 11 isolates (Supplementary Table B3) were obtained and deposited in GenBank. The remaining

samples, lacking adequate coverage and/or missing few nucleotides in the coding regions of the genome, were excluded from further study. Complete genomes of RNA1, RNA2, and RNA3 of BMV were obtained from Smith and Republic counties of Kansas. However, the complete sequence of movement protein was obtained from Cheyenne, Decatur, Ness, and Jewell counties (Supplementary Table B3).

Sequence alignment, percent identity, and similarity

The complete coding sequences of RNA1a and RNA2a (Table 3.1 and Table 3.2 respectively) as well RNA3a (movement protein, MP) and sub-genomic RNA4 (coat protein, CP) (Table 3.3 and Table 3.4 respectively) were aligned with the sequences obtained from the GenBank.

RNA1: RNA1 of BMV encodes for methyltransferase and helicase. The nucleotide sequences of RNA1a (ORF1a) were > 99% identical and 100% amino acid sequence identity between the isolates 20SM3 and 19RP1 from Smith and Republic counties (Table 3.1). The nucleotide sequence and amino acid sequence of both isolates were > 99% identical with the other US isolates from OH, OK, and WI. However, the nucleotide sequence of BMV isolates from the Czech Republic (BMV_CZ) and Estonia (BMV_Estonia) were 98.2% and 98.5% identical with 20SM4 and were 97.9 and 98.2 identical with 19RP1 respectively. The nucleotide sequence of the BMV_CZ isolate was 97.8% to 98.3% identical with other US isolates (OH, OK, and WI).

In ORF1a, the amino acid substitutions were unique among isolates. Two Ohio isolates (BMV_OH, and BMV_OH2) share two amino acid substitutions (Q278R and D569A) out of three. BMV_OH2 and BMV_OK shared one amino acid substitution (K536I). The isolate from Estonia had eight amino acid substitutions out of 10 isolates compared (Supplementary Figure

B1), including three consecutive amino acid substitutions and deletion from 21 to 23 (T21H, T22del, and N23H). It also shared three amino acid substitutions (A257T, D573G, and K827Q) with the BMV_CZ isolate. Both Czech and Estonian isolates showed higher variability than US isolates and isolates from UK and Germany in RNA1a.

RNA2: RNA2 of BMV encodes RNA-dependent RNA polymerase. The complete genome coding sequence of RNA2 was 99.6% identical and 100% amino acid sequence identity between the two isolates 20SM3 and 19RP1 (Table 3.2). The nucleotide sequence and amino acid sequence of both isolates were > 99% identical with other US isolates. However, the nucleotide sequence of both isolates was about 98% and 97.5% identical to the isolate from the Czech Republic and Estonia respectively. Two isolates from OH were 99.9% nucleotide and 100% amino acid sequence identical to each other. Notably, the 20SH3 and 19RP1 isolates share one amino acid substitution (E667K) in ORF2a or RNA2a which is unique to only these Kansas isolates (Supplementary Figure B2). Among other US isolates, there were four amino acid substitutions in BMV_OK isolate (L606W, D627E, T717M, K776R), three amino acid substitutions in BMV_M1 isolate (I609V, M655T, and l746I), two amino acid substitutions in BMV_M2 isolate (K567R and T717M), and BMV_UK isolate also had three amino acid substitutions (S132T, R277K, and T784A) (Supplementary Figure B2). There were six amino acid substitutions in the Czech isolate (H199R, R277K, K281R, K621R, L766S, and L809V) and five amino acid substitutions in Estonian isolate (A134V, A135D, D148E, D162V, and A677S).

RNA3: RNA3 of BMV encodes two proteins: the Movement protein (MP) and the coat protein (CP). The nucleotide and amino acid sequences were 100% identical for ORF 3a (MP) among the three isolates obtained in this study (20SM3, 19CN1, and 19CN3). 19JW1 and 19 RP1 were also 100% identical for nucleotide and amino acid sequences to each other. 19NS2

was >99% nucleotide sequence and 100% amino acid sequence identity for ORF 3a with 20SM3, 19CN1, 19CN3, and 19DC1 (Table 3.3). The nucleotide and amino acid sequences of isolates 19JW1 were >99% identical for ORF 3a (MP) among isolates obtained in this study. The nucleotide sequence and amino acid sequence identity were >99% among isolates from OH, OK, WI, and Estonia. However, nucleotide sequence and amino acid sequence of ORF 3a of the Czech isolate were >98% identical among all isolates analyzed except BMV_OH2 isolate (>97% identity). The 19RP1 and 19JW1 isolates share one amino acid substitution (L275F) with each other. The isolates 19NS2 has one amino acid substation (D166N) (Supplementary Figure B3). BMV_OH, BMV_OH2, BMV_OK, and BMV_M2 shared one amino acid substitution (T299S). Two OH isolates of BMV shared one more isolate (V162I). BMV_M1 and BMV_UK isolates shared one amino acid substitution (P81S) and BMV_UK also shared one more amino acid substitution (D166H) to BMV_Germany. Czech isolate has five amino acid substitutions on S57P, I99V, Q225R, L275P, and G276D (Supplementary Figure B3).

In CP ORF (coat protein), the nucleotide and amino acid sequences were 100% identical among two isolates obtained in this study (20SM3 and 19RP1) as well as with BMV_OH isolate (Table 3.4). The nucleotide sequences were 99.8% and 100% identical with 20SM3 and 19RP1 to BMV_OH2 and BMV_Estonia. However, the nucleotide sequences were about 96.8% and 97.4% amino acid identity between BMV_CZ isolate compared with 20SM3 and 19RP1. BMV_CZ had 95.8% to 96.8% nucleotide sequence and 97.4% to 97.9% amino acid sequence identity with other isolates analyzed in this study. Czech isolate has five amino acid substitutions including R22P, T24A, A25V, R26K, and A124V (Supplementary Figure B4). BMV_OK (R26T and L35F) and BMV_UK (R26G and V100I) have two amino acid substitutions and BMV_M1 (R23W) and BMV_M2 (A25T) have one amino acid substitution (Supplementary Figure B4).

Phylogenetic analysis

The sequence of coding regions of RNA1a, RNA2a, and RNA4 of 11 BMV isolates, two isolates from this study, and nine isolates obtained from GenBank (Supplementary Table B3) were used to construct the three phylogenetic trees separately (Figure 3.4, 3.5, and 3.7). However, we obtained complete sequences of RNA3a of seven isolates from this study and the phylogenetic tree constructed of a total of 16 BMV isolates (Figure 3.6), nine isolates obtained from GenBank (Supplementary Table B1). Cassia yellow blotch virus (CYBV) and Olive latent virus2 (OLV) were used as outgroups in the analysis.

The eleven BMV isolates used in this study were grouped into separate clades depending upon the RNA genome. The coding sequence of RNA1a of 11 BMV isolates consists of three clades (Figure 3.4). Clade A was represented by three isolates: Czech, Estonian, and a Wisconsin isolate from the US. Clade B was represented by a single isolate 20SM3 from Smith County, KS. Clade C included isolates from the US, UK, and Germany. In clade C, two isolates from Ohio and isolates from UK and Germany form separate sister taxa groups (Figure 3.4).

Based on the coding sequence of RNA2 of 11 BMV isolates, the BMV RNA2a topology grouped into a single clade (Figure 3.5). However, similar two isolates from Ohio and isolates from UK and Germany form a sub-clade with separate sister taxa groups (Figure 3.5).

The topology constructed using the coding sequence of the movement protein of BMV consisted of two clades (Figure 3.6). Clade A included the isolates from the Czech Republic and Estonia. Clade B polytomies included isolates from the UK, Germany, and other US isolates. The isolates collected from Jewell (19JW1) and Republic (19RP1) Kansas counties form a sister taxa group and similarly two isolates also form a sister taxa group (Figure 3.6). The other three

isolates collected from Smith (20SM3), Ness (19NS2), Cheyenne (19CN1 and 19CN3), and Decatur (19DC1) counties form the polytomy in clade B.

Similarly, the Phylogenetic tree constructed using coding sequences of RNA4 (Coat protein) also consists of two clades (Figure 3.7). Clade A was represented a single BMV Czech isolate. Clade B consisted of all US isolates and isolate from Estonia, the UK, and Germany. In Clade B, a sister taxa group was formed by isolates from Germany and UK.

Population genetic parameters and neutrality tests

The population genetic parameters including nucleotide diversity, mutation, and mutation rate per segregating site of BMV US isolates were calculated using DnaSP 5.10 (Table 3.5). Among four proteins, RNA2a exhibited the highest diversity ($\pi = 0.01$), while RNA1a and RNA3a showed the lowest diversity ($\pi = 0.004$). The degree of constrains for amino acid changes measured by the dN/dS for each encoded region showed that RNA3a was the least tolerant region with the order of tolerance RNA1a > RNA2a > RNA3b > RNA4 compared to RNA3a and RNA4.

The dN/dS ratio of the number of nonsynonymous substitutions to the number of synonymous substitutions for all proteins coding genes was < 1 for US isolates of BMV (Table 3.5). The selection pressure was measured from the three different algorithms (FEL, FUBAR, and SLAC) and the purifying or negative selection was supported by all three methods. No positive selection pressure was significantly reported at least by two methods (data not shown).

Discussion

One of the main factors that affect the durability of resistance is the dynamics of genetic variability of a pathogen (García-Arenal and McDonald, 2003). The results of the genetic characterization of BMV isolates we recently identified in Kansas wheat fields provide

information on the further study of wheat virus evolution, designing appropriate diagnostic tools, and developing durable viral disease management strategies through the breeding program. In this study, we reported the details of BMV co-infection with other wheat viruses in Kansas that were not previously identified and compared genetic variability and evolutionary characteristics with other BMV isolates obtained from this study and retrieved from the GenBank. Here, we have established the phylogenetic relationship and determined the major evolutionary mechanisms of KS BMV isolates based on coding sequences of all four RNAs.

Our results identifying BMV from 29 different counties of Kansas suggested that BMV has the potential to cause significant economic losses in Kansas wheat production. Previous studies showed that BMV reduced the wheat kernel weight and number of Kernel per spike (Pocsai et al., 1991; Tošič, 1971) and total grain yield up to 61% at early stage inoculation in Ohio wheat fields (Hodge et al., 2019). Our finding of this virus coinfected with common yield-reducing wheat viruses in Kansas demands future studies to examine the bi-, tri-, quadri, or multipartite interaction of these viruses in wheat and their impact on production. A recent study reported the quadripartite infection of wheat by BMV, WSMV, TriMV, and Barley stripe mosaic virus (BSMV) resulted in severe disease synergism with the death of most infected plants (Tatineni et al., 2021). The authors also reported the titer of the viruses depends upon the types of multipartite infection. Therefore, future research should endeavor to measure the impact of multipartite infection on Kansas-adapted wheat cultivars to estimate the differential synergistic impact of viruses on Kansas wheat production.

In this study, BMV was only identified on the samples that were infected with WSMV. After WSMV, the order of common co-infection of BMV was TriMV > BYDV > HPWMoV. Paliwal, 1972 reported wheat curl mite (WCM), a vector of WSMV, TriMV, and HPWMoV

couldn't transmit BMV although BMV was reported from the body cavity, they found no virus particle in salivary glands. Future *in vitro* feeding study of purified virus isolates tagged with a fluorescent protein in artificial stretched parafilm membranes or *in vivo* feeding of the fluorescent-tagged virus would help to track the movement of the virus in the body of WCM. Suppose WCM cannot transmit BMV, in that case, the potential another vector could be an aphid that transmits BYDV, or it also could be transmitted by different methods, including mechanically or by other vectors. A study of adaptive evolutionary response with a serial passage of BMV isolates on WCM from infected plants helps to estimate the virus population structure change. In one sample BMV was co-infected with two soilborne wheat viruses, WSSMV and SBWMV as well. It is difficult to determine the potential BMV vector. Vector determination is crucial for plant virus management as one of the important control strategies of viruses is vector control (Jeger et al., 2004; Perring et al., 1999). Therefore, a comprehensive future study to determine field transmission of BMV, including a survey of wheat fields, would answer this ambiguous transmission of BMV.

Sequence alignment analysis of the coding regions of BMV RNAs showed that they were closely related to each other as they shared a high nucleotide sequence identity (>95%). However, US isolates showed lower similarity to Czech and Estonian isolates. Similar results were also reported by Jeżewska et al., 2019 comparing BMV isolates from Poland with BMV isolates from other European and US isolates. Gadiou and Kundu, 2013; Jeżewska et al., 2019 reported the most divergence in coat protein RNA4 region. Our results also showed higher similarities in both nucleotide and amino acid sequence of RNA3a movement protein, and US isolates had the least nucleotide sequence identity with Czech isolate in coat protein region (96.5%).

The two conserved domains of RNA1a are RNA capping enzyme domain (L52, H80, D106, and R136) helicase-like domains (K691, D755, and G781) and a polymerase-related domain of RNA2a (451 to 484) required for BMV RNA replication and mutation in or near these domains abolishes or decrease BMV RNA synthesis (Ahola et al., 2000; Koonin et al., 1991; Kroner et al., 1989). Amino acid sequence alignments of the BMV isolates in this study also maintained the integrity of the conserved domains of RNA1a and RNA2a (Supplementary Figure B1 and B2).

The amino acid substitutions were variable among isolates, and past studies showed that host adoption played an important role. De Jong and Ahlquist, 1995 described viral RNA accumulation in systemic infection closely associated with interaction with virus-host. They reported that both BMV_M1 and M2 strains systemically infected the monocot host barley but the dicot host cowpea was infected systemically only by BMV_M2. RNA3a movement protein and C-terminus of coat protein control the BMV movement from cell to cell and systemic movement (De Jong et al., 1995; Okinaka et al., 2001) and movement protein also played a role in host specificity (De Jong and Ahlquist, 1992; Mise et al., 1993; Mise and Ahlquist, 1995). De Jong et al., 1995 showed that quadruple substitution in BMV_M2 movement protein (E59Q, S81P, S297G, and T229S) were required to infect cowpea dicot host systemically. None of these substitutions were found in BMV_M1, the monocot-adapted isolate. Our results of BMV isolates obtained from the wheat host in this study had only one amino acid (S81P) change similar to BMV_M2 (Supplementary Figure B3). However, BMV_OH, BMV_OH2, and BMV_OK shared only two amino acid substitutions with BMV_M2 (S81P and T299S). Hodge et al., 2019 reported that cowpea and soybean were systematically infected by BMV_OH, which means only two amino acid substitution in MP were sufficient to infect dicot hosts. The isolates obtained in

this study had only one amino acid substitution out of four essential substitutions. If these isolates can infect the dicot host with one amino acid substitution, that threatens the crop production of dicot and monocot crops due to host expansion. Therefore, future studies to determine the host ranges of Kanas BMV isolates, or survey of corn, soybean, fescue from Kansas for detection of BMV will help to develop the management strategies of crop rotation or management of the BMV reservoir or alternative hosts.

Sacher and Ahlquist, 1989 reported that the deletion of the first 25 amino acids of BMV coat protein failed on the packaging of RNA and systemic infection. BMV isolates analyzed in this study have no change in the first 23 N-terminal amino acids in coat protein showing the conserved N-terminal region required for packaging and systemic infection. Additionally, Rao and Grantham, 1995 revealed that amino-terminal residues of 1 to 7 are required for chlorotic local lesions, and systemic infection in *Chenopodium quinoa* however did not affect barley plant infections. Therefore 1-7 N-terminal residues play important role in virus-host interactions. Deletion first 11, 14, and 18 N-terminal amino acids, especially arginine-rich motif, played a role in modulating symptom expression and movement in dicot and monocot hosts (Rao and Grantham, 1996). Okinaka et al., 2001 investigated the 19 alanine-scanning mutant, the results indicated that the C-terminal region (mainly from 178 to 187 residues) played an essential role in virus encapsidation and movement as alanine mutant on this region failed to produce virion and cell to cell movement. Yi et al., 2009 showed that three residues (D139, R142, and D 148) in the C-terminal of CP required for the BMV RNA accumulation as a mutation on these residues impact CP-associated activities. We also found no variations in the first 23 N-terminal and last 64 C-terminal amino acid residues in CP of BMV isolates analyzed in this study (Supplementary Table B4).

Phylogenetic relations showed modest variation among isolates with slightly different clustering based on nucleotide sequences of RNA genomes (Figure 3.4 to 3.7) and indicated different evolutionary constraints in the different coding regions of RNAs. As only limited complete genome of BMV available in the GenBank, the phylogenetic trees indicated no clear grouping of isolates by geographical areas. The isolates from the Czech Republic and Estonia, UK and German, and two isolates from Ohio were most closely related to each other as they formed sister taxa in all four trees. Also, the BMV_M1 (US isolates) clustered with isolates from Estonia and the Czech Republic and 20SM3 singly formed clade B on phylogenetic tree constructed based on ORF 1a. These results suggested that although, 20SM3 was 100% similar in protein identity with 19RP1 it was sufficiently different in the nucleotide sequence, suggesting synonymous substitutions. Jeżewska et al., 2019 reported a similar clustering of the US isolates, Czech and Estonian isolates based on the CP region that we reported in this study. The slight variability among BMV isolates in different coding regions might be associated with interaction with the specific hosts and the genetic requirements to perform successful cell to cell and systemic movement (De Jong et al., 1995; De Jong and Ahlquist, 1995).

The selection pressure in protein-coding genes was calculated by nonsynonymous to synonymous substitution (dNs/Ds) ratio. The values of the dNs/Ds used to identify protein sites that experience neutral selection (dNs/Ds ≈ 1), negative or purifying selection (dNs/Ds >1), and experience positive or adaptive/diversifying selection (dNs/Ds >1) (Kosakovsky Pond and Frost, 2005; Yang et al., 2000). The ratio of the nonsynonymous and synonymous positions (dNs/Ds) estimation of all ORF of BMV genome was below one that indicates the negative or purifying selection. This result is similar to the common negative, or purifying selection of other plant RNA viruses (García-Arenal et al., 2001) as genetic stability is common. Only few complete

coding sequences of US BMV isolates available in the GenBank, future analysis of many BMV isolates would be needed for more accurate assessment of population parameters. The high genetic stability found for all proteins of BMV could be attributed to negative or purifying selection to maintain the functional integrity of the viral genome as described by other RNA viruses (Moreno et al., 2004). The low genetic diversity and purifying selection as common selection pressure was also reported for populations of other RNA viruses including WSMV, TriMV, cucumber mosaic virus, and Citrus psorosis virus (Martín et al., 2006; Nouri et al., 2014; Redila et al., 2021; Stenger et al., 2002). As plant RNA viruses have a short genome, each amino acid sequence contributed to encode protein. However, there is some room for the variation in RNA viruses through mutation and recombination (Drake and Holland, 1999; García-Arenal et al., 2001). The linear model of replication, preventing mutational meltdown, and genetic bottleneck are the other reasons for low variability or neutral and purifying selection in RNA viruses (French and Stenger, 2003; Stent, 1963).

Overall, the present study characterized the newly identified BMV isolates from Kansas wheat fields. This study showed the significance of nanopore sequencing in detection, diagnosis, and molecular characterization based on the whole genome sequence of undetected plant pathogens. Information of genetic variation, phylogenetic relationship with the other US and non-US isolates, and evolutionary mechanism employed by four different RNA genomes of BMV would support the sustainable management of wheat viruses through genetic resistance.

References

Ahlquist, P., French, R., Janda, M., Loesch-Fries, L.S., 1984. Multicomponent RNA plant virus infection derived from cloned viral cDNA. Proceedings of the national academy of sciences 81, 7066–7070.

Ahola, T., Den Boon, J.A., Ahlquist, P., 2000. Helicase and capping enzyme active site mutations in brome mosaic virus protein 1a cause defects in template recruitment,

negative-strand RNA synthesis, and viral RNA capping. Journal of virology 74, 8803–8811.

Caetano, V., Marinho, V., Lin, M., Formiga, L., Kitajima, E., 1990. Ocorrencia do virus do mosaico do capim bromo (Brome mosaic virus) em trigo, no estado do Rio Grande do Sul. Fitopatol. Bras 15, 363–365.

Damsteegt, V., Gildow, F., Hewings, A., Carroll, T., 1992. A clone of the Russian wheat aphid (*Diuraphis noxia*) as a vector of barley yellow dwarf, barley stripe mosaic, and brome mosaic viruses. Plant Disease 76, 1155–1160.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., and Van Broeckhoven, C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34: 2666–2669

De Jong, W., Ahlquist, P., 1995. Host-specific alterations in viral RNA accumulation and infection spread in a brome mosaic virus isolate with an expanded host range. Journal of virology 69, 1485–1492.

De Jong, W., Ahlquist, P., 1992. A hybrid plant RNA virus made by transferring the noncapsid movement protein from a rod-shaped to an icosahedral virus is competent for systemic infection. Proceedings of the National Academy of Sciences 89, 6808–6812.

De Jong, W., Chu, A., Ahlquist, P., 1995. Coding changes in the 3a cell-to-cell movement gene can extend the host range of brome mosaic virus systemic infection. Virology 214, 464–474.

Díaz-Cruz, G., Smith, C., Wiebe, K., Charette, J., Cassone, B., 2018. First report of brome mosaic virus infecting soybean, isolated in Manitoba, Canada. Plant Disease 102, 460.

Drake, J.W., Holland, J.J., 1999. Mutation rates among RNA viruses. Proceedings of the National Academy of Sciences 96, 13910–13913.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32, 1792–1797.

Erasmus, D., Rybicki, E., Von Wechmar, M., 1983. The association of brome mosaic virus and wheat rusts. II. Detection of BMV in/on uredospores of wheat stem rust. Journal of Phytopathology 108, 34–40.

French, R., Stenger, D.C., 2003. Evolution of Wheat streak mosaic virus: dynamics of population growth within plants may explain limited variation. Annual review of phytopathology 41, 199–214.

Gadiou, S., Kundu, J.K., 2013. Complete genome sequence of a brome mosaic virus isolate from the Czech Republic. Czech Journal of Genetics and Plant Breeding 46, 178–182.

García-Arenal, F., Fraile, A., Malpica, J.M., 2001. Variability and genetic structure of plant virus populations. Annual review of phytopathology 39, 157–186.

García-Arenal, F., McDonald, B.A., 2003. An analysis of the durability of resistance to plant viruses. Phytopathology 93, 941–952.

Gibson, R., Kenten, R., 1978. The occurrence of brome mosaic virus in Britain. Plant Pathology 27, 66–67.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology 52, 696–704.

He, G., Zhang, Z., Sathanantham, P., Diaz, A., Wang, X., 2021. Brome Mosaic Virus (Bromoviridae). Encyclopedia of Virology 252.

Hodge, B., Paul, P., Stewart, L.R., 2020. Occurrence and High-Throughput Sequencing of Viruses in Ohio Wheat. Plant Disease 104, 1789–1800.

Hodge, B., Salgado, J., Paul, P., Stewart, L., 2019. Characterization of an Ohio isolate of brome mosaic virus and its impact on the development and yield of soft red winter wheat. Plant Disease 103, 1101–1111.

Hollandbeck, F. G., Dewolf, E., and Todd, T. 2020. Kansas Cooperative Plant Disease Survey Report, Preliminary 2020 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Hollandbeck, F. G., Onofre, K. A., Dewolf, E., and Todd, T. 2021. Kansas Cooperative Plant Disease Survey Report, Preliminary 2021 Kansas wheat disease loss estimates, Kansas Department of Agriculture.

Huff, D.E., Davis, R.F., Myers, R.F., 1987. *Longidorus breviannulatus* as a vector for brome mosaic virus. Journal of nematology 19, 143.

Jeger, M., Holt, J., Van Den Bosch, F., Madden, L., 2004. Epidemiology of insect-transmitted plant viruses: modelling disease dynamics and control interventions. Physiological Entomology 29, 291–304.

Jeżewska, M., Trzmiel, K., Zarzyńska-Nowak, A., 2019. Detection of infectious brome mosaic virus in irrigation ditches and draining strands in Poland. European journal of plant pathology 153, 285–292.

Kao, C.C., Sivakumaran, K., 2000. Brome mosaic virus, good for an RNA virologist's basic needs. Molecular plant pathology 1, 91–97.

Koonin, E.V., Mushegian, A.R., Ryabov, E.V., Dolja, V.V., 1991. Diverse Groups of Plant RNA and DNA Viruses Share Related Movement Proteins that may Possess Chaperone-like Activity. Journal of General Virology 72, 2895–2903

Kosakovsky Pond, S.L., Frost, S.D., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Molecular biology and evolution 22, 1208–1222.

Kroner, P., Richards, D., Traynor, P., Ahlquist, P., 1989. Defined mutations in a small region of the brome mosaic virus 2 gene cause diverse temperature-sensitive RNA replication phenotypes. Journal of Virology 63, 5302–5309.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution 35, 1547.

Lane, L.C., 1974. The bromoviruses. Advances in virus research 19, 151–220.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Lucas, R.W., Larson, S.B., McPherson, A., 2002. The crystallographic structure of brome mosaic virus. Journal of molecular biology 317, 95–108.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R., Potter, S.C., Finn, R.D., others, 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic acids research 47, W636–W641

Martín, S., García, M.L., Troisi, A., Rubio, L., Legarreta, G., Grau, O., Alioto, D., Moreno, P., Guerri, J., 2006. Genetic variation of populations of Citrus psorosis virus. Journal of general virology 87, 3097–3102.

Mian, M.R., Zwonitzer, J., Hopkins, A., Ding, X., Nelson, R., 2005. Response of tall fescue genotypes to a new strain of brome mosaic virus. Plant disease 89, 224–227.

Mise, K., Ahlquist, P., 1995. Host-specificity restriction by Bromovirus cell-to-cell movement protein occurs after initial cell-to-cell spread of infection in non-host plants. Virology 206, 276–286.

Mise, K., Allison, R., Janda, M., Ahlquist, P., 1993. Bromovirus movement protein genes play a crucial role in host specificity. Journal of virology 67, 2815–2823.

Moreno, I., Thompson, J., Garcia-Arenal, F., 2004. Analysis of the systemic colonization of cucumber plants by Cucumber green mottle mosaic virus. Journal of General Virology 85, 749–759.

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford university press.

Nouri, S., Arevalo, R., Falk, B.W., Groves, R.L., 2014. Genetic structure and molecular variability of Cucumber mosaic virus isolates in the United States. PLoS One 9, e96582.

Okinaka, Y., Mise, K., Suzuki, E., Okuno, T., Furusawa, I., 2001. The C terminus of brome mosaic virus coat protein controls viral cell-to-cell and long-distance movement. Journal of virology 75, 5385–5390.

Paliwal, Y., 1972. Brome mosaic virus infection in the wheat curl mite *Aceria tulipae*, a nonvector of the virus. Journal of Invertebrate Pathology 20, 288–302.

Papp, M., Mesterházy, Á., Vasdinyei, R., Gáborjányi, R., 1996. Mixed virus infection of wheat in South-East Hungary in 1994 and 1995. Cereal Research Communications 179–182.

Perring, T.M., Gruenhagen, N.M., Farrar, C.A., 1999. Management of plant viral diseases through chemical control of insect vectors. Annual review of entomology 44, 457–481.

Pocsai, E., Kobza, S., Murányi, I., Szunics, L., 1991. Brome mosaic virus infection in different cereal breeding materials. Acta Phytopathologica et Entomologica Hungarica 26, 207–212.

Rao, A., 2006. Genome packaging by spherical plant RNA viruses. Annual Review of Phytopathology. 44, 61–87.

Rao, A., Grantham, G.L., 1996. Molecular studies on Bromovirus capsid protein: II. Functional analysis of the amino-terminal arginine-rich motif and its role in encapsidation, movement, and pathology. Virology 226, 294–305.

Rao, A., Grantham, G.L., 1995. Biological significance of the seven amino-terminal basic residues of brome mosaic virus coat protein. Virology 211, 42–52.

Redila, C.D., Phipps, S., Nouri, S., 2021. Full Genome Evolutionary Studies of Wheat Streak Mosaic-Associated Viruses Using High-Throughput Sequencing. Frontiers in Microbiology 1998.

Rybicki, E., Von Wechmar, M., 1982. Characterization of an Aphid-Transmitted Virus Disease of Small Grains: Isolation and Partial Characterization of Three Viruses. Journal of Phytopathology 103, 306–322.

Sacher, R., Ahlquist, P., 1989. Effects of deletions in the N-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. Journal of virology 63, 4545–4552.

Schmidt, H., Fritzsche, R., Lehmann, W., 1963. Die übertragung des Weidelgrasmosaik-virus durch Nematoden. Naturwissenschaften 50, 386–386.

Sõmera, M., Gantsovski, M., Truve, E., Sooväli, P., 2016. First report of brome mosaic virus in wheat in Estonia. Plant Disease 100, 2175–2175.

Srivatsavai, V., 2005. Identification, distribution, and vector biology of brome mosaic virus of wheat in Alabama (PhD Thesis). University of Alabama.

Stenger, D.C., Seifers, D.L., French, R., 2002. Patterns of polymorphism in Wheat streak mosaic virus: sequence space explored by a clade of closely related viral genotypes rivals that between the most divergent strains. Virology 302, 58–70.

Stent, G.S., 1963. Molecular biology of bacterial viruses. Molecular biology of bacterial viruses.

Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28, 1102–1104.

Tatineni, S., Alexander Jeff, Qu Feng, 2021. Differential Synergistic Interactions among Four Different Wheat-infecting Viruses. Frontiers in microbiology.

Tošič, M., 1971. Virus Diseases of Wheat in Serbia: I. Isolation and determination of the Wheat streak mosaic virus and brome mosaic virus 1. Journal of Phytopathology 70, 145–162.

Trzmiel, K., Szyd\lo, W., Zarzyńska-Nowak, A., Jeżewska, M., 2015. First report of brome mosaic virus (BMV) and Wheat streak mosaic virus (WSMV) co-infection in triticale plants in Poland. Plant Disease 99, 1290–1290.

Trzmiel, K., Zarzyńska-Nowak, A., Lewandowska, M., Szyd\lo, W., 2016. Identification of new brome mosaic virus (BMV) isolates systemically infecting Vigna unguiculata L. European Journal of Plant Pathology 145, 233–238.

Urbanavičienė, L., and Žižytė, M. 2012. Identification of brome mosaic virus in cocksfoot (*Dactylis glomerata* L.) and meadow fescue (*Festuca pratensis* Huds.) in Lithuania. Agric 99, 167–172.

Von Wechmar, M., Rybicki, E., 1985. Brome Mosaic Virus Infection Mimics, Barley Yellow Dwarf Virus Disease Symptoms in Small Grains. Journal of Phytopathology 114, 332–337.

Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microbial genomics **3**. Microbiology Society.

Wick, R.R., Judd, L.M., and Holt, K.E. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome biology **20**: 1–10.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.-M.K., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Yi, G., Vaughan, R.C., Yarbrough, I., Dharmaiah, S., Kao, C.C., 2009. RNA binding by the brome mosaic virus capsid protein and the regulation of viral RNA accumulation. Journal of molecular biology 391, 314–326.

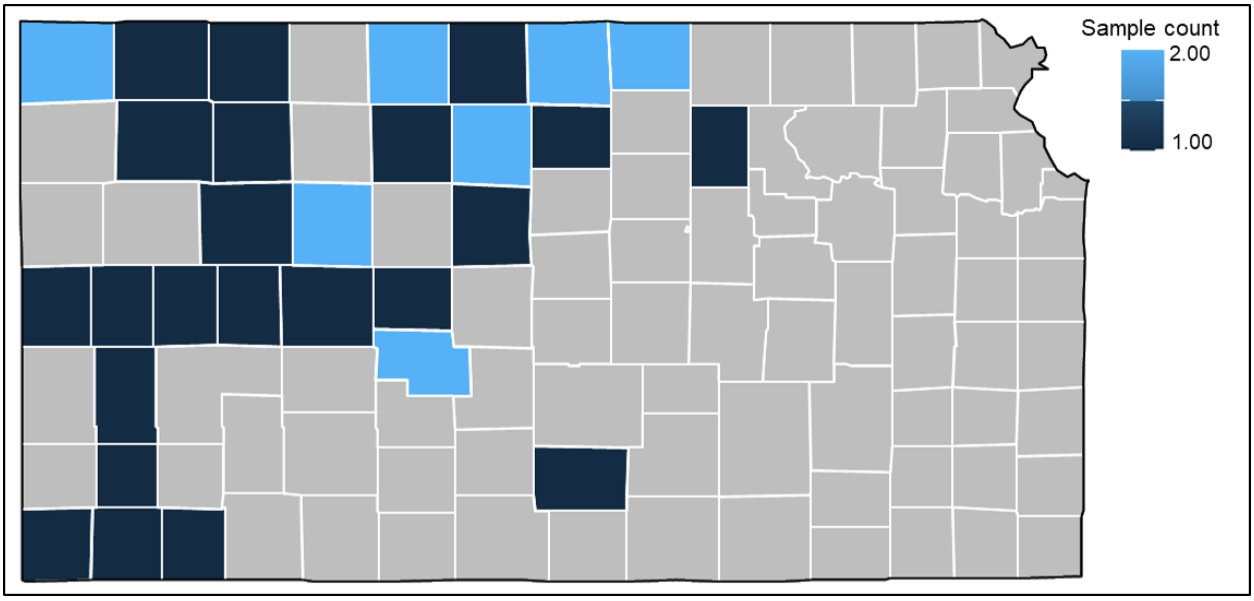


Figure 3.1. The map of Kansas with counties where brome mosaic virus was identified using Nanopore sequencing.

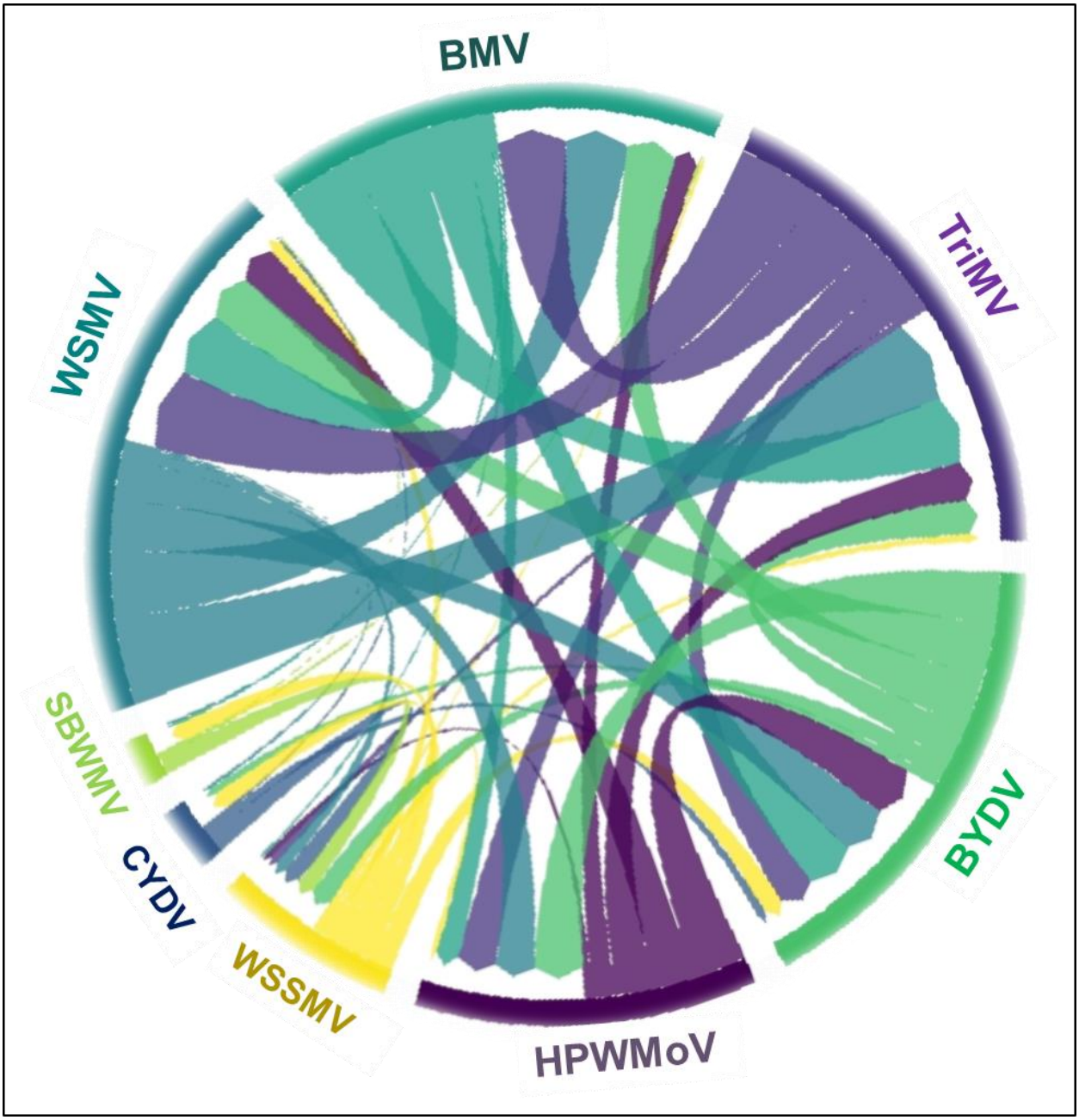


Figure 3.2. A chord interaction diagram of co-infection of wheat streak mosaic virus (WSMV), triticum mosaic virus (TriMV), High Plains wheat mosaic emaravirus (HPWMoV), brome mosaic virus (BMV), barley yellow dwarf virus (BYDV), wheat spindle streak mosaic virus (WSSMV), cereal yellow dwarf virus (CYDV), soilborne wheat mosaic virus (SBWMV) in symptomatic wheat leaves detected using Nanopore sequencing (n = 84). The arch shows the connection and size of the total percentage of virus co-infected in a single sample.

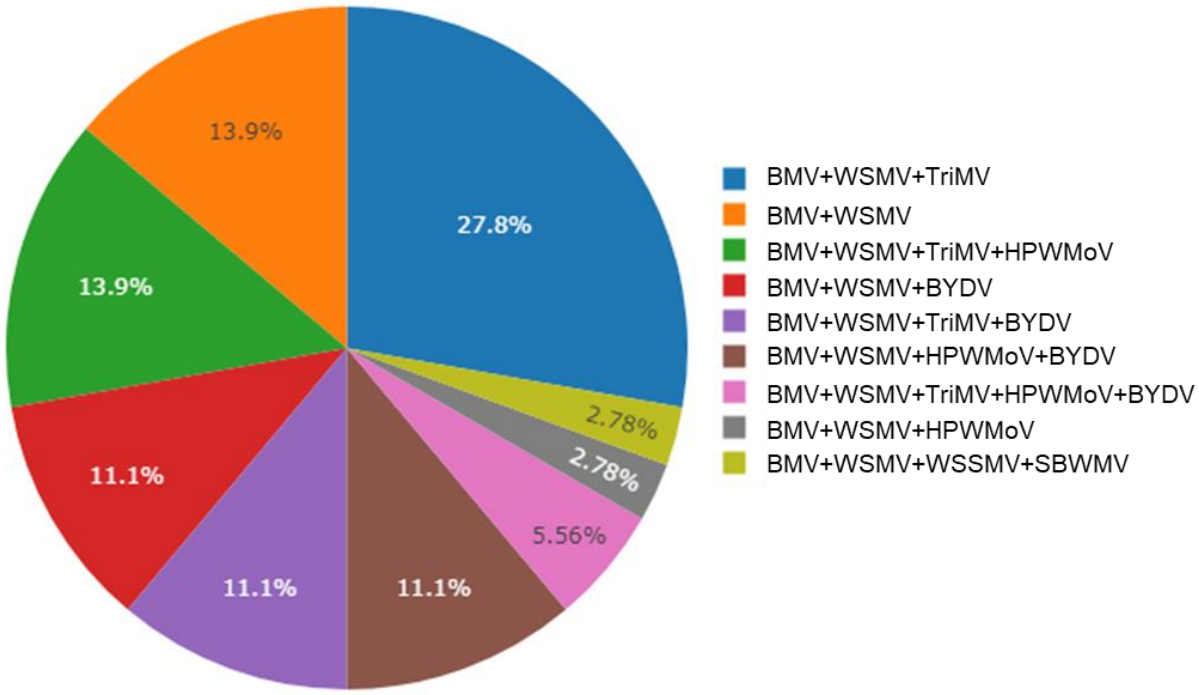


Figure 3.3. Percent incidence of brome mosaic virus co-infected with of wheat streak mosaic virus (WSMV), triticum mosaic virus (TriMV), High Plains wheat mosaic emaravirus (HPWMoV), brome mosaic virus (BMV), barley yellow dwarf virus (BYDV), wheat spindle streak mosaic virus (WSSMV), cereal yellow dwarf virus (CYDV), soilborne wheat mosaic virus (SBWMV)in leaf samples collected from Kansas wheat fields detected by Nanopore sequencing. Virus-like symptomatic wheat leaves were collected from winter wheat field in 2019, 2020, and 2021**.**

Table 3.1. Nucleotides/amino acid sequence identity (similarity) of RNA1/RNA1a genome of rome mosaic virus (BMV) isolates from Kansas and other known BMV isolates retrieved from GeneBank. The nucleotide percentage identity was calculated from ClustalW alignment online tools (Clustal 2.1) (Madeira et al., 2019), and amino acid sequence identity and similarity were calculated from a sequence manipulation suite (Stothard, 2000).

BMV isolates*	20SM3	19RP1	BMV-OH	BMV-OH2	BMV-OK	BMV-M1
20SM3	-					
19RP1	99.5/100(100)	-				
BMV-OH	99.3/99.6(99.7)	99.2/99.6(99.7)	-			
BMV-OH2	99.3/99.7(99.7)	99.3/99.7(99.7)	99.7/99.8(99.9)	-		
BMV-OK	99.5/99.8(99.8)	99.5/99.8(99.8)	99.2/99.5(99.6)	99.5/99.7(99.7)	-	
BMV-M1	99.3/99.5(99.7)	99.1/99.5(99.7)	98.8/99.1(99.4)	98.8/99.2(99.4)	99.1/99.3(99.5)	-
BMV-M2	99.3/99.7(99.9)	99.5/99.7(99.9)	99.1/99.3(99.6)	99.3/99.4(99.6)	99.5/99.5(99.7)	98.9/99.2(99.6)
BMV-UK	99.5/99.7(99.8)	99.5/99.7(99.8)	99.2/99.3(99.5)	99.3/99.4(99.5)	99.5/99.5(99.6)	99.1/99.3(99.5)
BMV-Germany	99.6/99.7(99.8)	99.5/99.7(99.8)	99.2/99.3(99.5)	99.3/99.4(99.5)	99.6/99.5(99.6)	99.1/99.3(99.5)
BMV-CZ	98.2/98.9(99.2)	97.9/98.9(99.1)	97.7/98.5(98.9)	97.7/98.7(98.9)	97.8/98.8(98.9)	98.2/99.1(99.3)
BMV-Estonia	98.5/99.5(99.5)	98.2/99.2(99.2)	98.1/99.1(99.2)	98.2/99.2(99.2)	98.1/99.3(99.3)	98.3/99.3(99.6)
	BMV_M2	BMV_UK	BMV_Germany	BMV_CZ		
BMV_UK	99.4/99.4(99.7)	-				
BMV_Germany	99.5/99.4(99.7)	99.9/100(100)	-			
BMV_CZ	97.6/98.7(99.1)	97.9/98.8(98.9)	97.9/98.8(98.9)	-		
BMV_Estonia	98.1/99.2(99.4)	98.2/99.3(99.3)	98.3/99.3(99.3)	97.8/99.2(99.5)		

Table 3.2. Nucleotides/amino acid sequence identity (similarity) of RNA2/RNA2a genome of brome mosaic virus (BMV) isolates from Kansas and other known BMV isolates retrieved from GeneBank. The nucleotide percentage identity was calculated from ClustalW alignment online tools (Clustal 2.1) (Madeira et al., 2019), and amino acid sequence identity and similarity were calculated from a sequence manipulation suite (Stothard, 2000).

BMV isolates*	20SM3	19RP1	BMV_OH	BMV_OH2	BMV_OK	BMV_M1
20SM3	-					
19RP1	99.6/100(100)	-				
BMV-OH	99.2/99.4(99.6)	99.3/99.9(99.9)	-			
BMV-OH2	99.4/99.9(99.9)	99.5/99.9(99.9)	99.9/100(100)	-		
BMV-OK	99.2/99.4(99.6)	99.3/99.4(99.6)	99.1/99.5(99.8)	99.2/99.5(99.8)	-	
BMV-M1	98.3/99.0(99.5)	98.1/99.0(99.5)	97.9/99.2(99.6)	98.0/99.2(99.6)	97.9/98.7(99.4)	-
BMV-M2	99.3/99.6(99.8)	99.3/99.6(99.8)	99.2/99.8(99.9)	99.3/99.8(99.9)	99.2/99.5(99.9)	98.1/98.9(99.5)
BMV-UK	98.3/99.5(99.8)	99.5/99.5(99.8)	99.3/99.6(99.9)	99.4/99.6(99.9)	99.2/99.2(99.6)	98.1/98.8(99.5)
BMV-Germany	99.4/99.6(99.8)	99.6/99.6(99.8)	99.5/99.8(99.9)	99.6/99.8(99.9)	99.3/99.3(99.6)	98.2/98.9(99.5)
BMV-CZ	98.4/98.8(99.5)	98.2/98.8(99.5)	98.0/98.9(99.6)	98.1/98.9(99.6)	98.0/98.4(99.4)	98.3/99.1(99.8)
BMV-Estonia	97.5/98.5(99.0)	97.4/98.5(99.1)	97.2/98.7(99.2)	97.3/98.7(99.2)	97.2/98.2(98.9)	97.5/98.8(99.3)
	BMV_M2	BMV_UK	BMV_Germany	BMV_CZ		
BMV_UK	99.2/99.4(99.8)	-				
BMV_Germany	99.4/99.5(99.8)	99.9/99.9(100)	-			
BMV_CZ	98.2/98.8(99.5)	98.1/98.7(99.5)	98.3/98.7(99.5)	-		
BMV_Estonia	97.3/98.3(99.0)	97.4/98.4(99.0)	97.4/98.4(99.0)	97.2/98.5(99.3)		

Table 3.3. Nucleotides/amino acid sequence identity (similarity) of RNA3/RNA3a (Movement protein, MP) genome of brome mosaic virus (BMV) isolates from Kansas and other known BMV isolates retrieved from GeneBank. The nucleotide percentage identity was calculated from ClustalW alignment online tools (Clustal 2.1) (Madeira et al., 2019), and amino acid sequence identity and similarity were calculated from a sequence manipulation suite (Stothard, 2000).

BMV isolates*	20SM3	19CN1	19CN3	19DC1	19NS2	19JW1
19CN1	100/100(100)	-				
19CN3	100/100(100)	100/100(100)	-			
19DC1	100/100(100	100/100(100)	100/100(100)	-		
19NS2	99.8/99.7(100)	99.8/99.7(100)	99.8/99.7(100)	99.8/99.7(100)	-	
19JW1	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.3(99.7)	-
19RP1	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.3(99.7)	100/100(100)
BMV-OH2	99.3/99.3(100)	99.3/99.3(100)	99.3/99.3(100)	99.3/99.3(100)	99.3/99.0(100)	99.3/99.0(99.7)
BMV-OH	99.5/99.3(100)	99.5/99.3(100)	99.5/99.3(100)	99.5/99.3(100)	99.5/99.0(100)	99.5/99.0(99.7)
BMV_M1	99.3/99.7(99.7)	99.3/99.7(99.7)	99.3/99.7(99.7)	99.3/99.7(99.7)	99.3/99.3(99.7)	99.3/99.3(99.3)
BMV-M2	99.6/99.0(99.7)	99.6/99.0(99.7)	99.6/99.0(99.7)	99.6/99.0(99.7)	99.6/98.7(99.7)	99.6/98.7(99.3)
BMV-OK	99.8/99.7(100)	99.8/99.7(100)	99.8/99.7(100)	99.8/99.7(100)	99.8/99.3(100)	99.8/99.3(99.7)
BMV-UK	99.6/99.0(99.3)	99.6/99.0(99.3)	99.6/99.0(99.3)	99.6/99.0(99.3)	99.7/99.0(99.3)	99.6/98.7(99.0)
BMV-Germany	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.7(99.7)	99.8/99.3(99.3)
BMV-Estonia	99.6/100(100)	99.6/100(100)	99.6/100(100)	99.6/100(100)	99.3/99.7(100)	99.3/99.7(99.7)
BMV-CZ	98.6/98.4(98.7)	99.6/98.4(98.7)	98.6/98.4(98.7)	99.6/98.4(98.7)	98.4/98.0(98.7)	98.4/98.4(98.7)
BMV isolates*	19RP1	BMV-OH2	BMV_OH	BMV-M1	BMV-M2	BMV-OK
BMV-OH2	99.3/99.0(99.7)	-				
BMV-OH	99.5/99.0(99.7)	99.9/100(100)	-			
BMV_M1	99.3/99.3(99.3)	98.9/99.0(99.7)	99.0/99.0(99.7)	-		
BMV-M2	99.6/98.7(99.3)	99.3/99.0(99.7)	99.5/99.0(99.7)	99.1/98.7(99.3)	-	
BMV-OK	99.8/99.3(99.7)	99.6/99.7(100)	99.7/99.7(100)	99.3/99.3(99.7)	99.8/99.3(99.7)	-
BMV-UK	99.6/98.7(99.0)	99.1/98.4(99.3)	99.2/98.4(99.3)	99.3/98.3(99.3)	99.3/98.0(99.0)	99.6/98.7(99.3)
BMV-Germany	99.8/99.3(99.3)	99.3/99.0(99.7)	99.5/99.0(99.7)	99.3/99.3(99.3)	99.6/98.7(99.3)	99.8/99.3(99.7)
BMV-Estonia	99.3/99.7(99.7)	98.9/99.3(100)	99.0/99.3(100)	99.1/99.7(99.7)	99.1/99.0(99.7)	99.3/99.7(100)

BMV-CZ	98.4/98.4(98.7)	97.9/97.7(98.7)	98.0/97.7(98.7)	98.1/98.0(98.4)	98.1/97.4(98.4)	98.4/98.0(98.7)
	BMV_UK	BMV_Germany	BMV_Estonia			
BMV_Germany	99.8/99.3(99.7)	-				
BMV_Estonia	99.1/99.0(99.3)	99.3/99.7(99.7)	-			
BMV_CZ	98.1/97.4(98.0)	98.4/98.0(98.4)	98.8/98.4(98.7)			

Table 3.4. Nucleotides/amino acid sequence identity (similarity) of RNA4 (coat protein, CP) genome of brome mosaic virus (BMV) isolates from Kansas and other known BMV isolates retrieved from GeneBank. The nucleotide percentage identity was calculated from ClustalW alignment online tools (Clustal 2.1) (Madeira et al., 2019), and amino acid sequence identity and similarity were calculated from a sequence manipulation suite (Stothard, 2000).

BMV isolates*	20SM3	19RP1	BMV-OH	BMV-OH2	BMV-OK	BMV-M1
19RP1	100/100 (100)	-				
BMV-OH	100/100(100)	100/100(100)	-			
BMV-OH2	99.8/100(100)	99.8/100(100	99.8/100(100)	-		
BMV-OK	99.6/98.9(98.9)	99.6/98.9(98.9)	99.7/98.9(98.9)	99.5/98.9(98.9)	-	
BMV-M1	98.9/99.5(99.5)	98.9/99.5(99.5)	98.9/99.5(99.5)	99.8/99.5(99.5)	98.6/98.4(98.4)	-
BMV-M2	99.5/99.5(99.5)	99.5/99.5(99.5)	99.5/99.5(99.5)	99.3/99.5(99.5)	99.1/98.4(98.4)	98.4/98.9(98.9)
BMV-UK	98.9/98.4(99.5)	98.9/98.4(99.5)	98.9/98.4(99.5)	98.8/98.4(99.5)	98.8/98.4(98.4)	98.3/97.9(98.9)
BMV-Germany	99.6/100(100)	99.6/100(100)	99.6/100(100)	99.5/100(100)	99.3/98.9(98.9)	98.9/99.5(99.5)
BMV-CZ	96.8/97.4(98.9)	96.8/97.4(98.9)	96.8/97.4(98.9)	96.7/97.4(98.9)	96.7/96.8(97.9)	96.5/96.8(98.4)
BMV-Estonia	98.9/100(100)	98.9/100(100)	98.9/100(100)	98.8/100(100)	98.6/98.9(98.9)	98.6/99.5(99.5)
	BMV_M2	BMV_UK	BMV_Germany	BMV_CZ		
BMV_UK	98.4/97.9(98.9)	-				
BMV_Germany	99.1/99.5(99.5)	99.3/98.4(99.5)	-			
BMV_CZ	96.3/97.4(98.4)	95.8/96.3(98.4)	96.5/97.4(98.9)	-		
BMV_Estonia	98.4/99.5(99.5)	97.9/98.4(99.5)	98.6/100(100)	96.8/97.4(98.9)		

Table 3.5. Population genetics parameters for encoded region of selected United States brome mosaic virus isolates calculated using DnaSP (Librado and Rozas, 2009) and MEGA X (Kumar et al., 2018)

Genomic region	Number of isolates	S*	η†	π‡	θw#	dS§	dN♯	dN/Ds (ω)Ψ
RNA1a	5	30	30	0.0048 ± 0.001	0.0049 ± 0.003	0.013 ± 0.003	0.001 ± 0.0004	0.077
RNA2a	7	79	79	0.01 ± 0.003	0.013 ± 0.006	0.027 ± 0.003	0.002 ± 0.001	0.074
RNA3a (MP)	12	15	15	0.004 ± 0.0008	0.005 ± 0.002	0.007 ± 0.003	0.003 ± 0.001	0.428
RNA3a (CP)	7	12	12	0.006 ± 0.002	0.008 ± 0.004	0.014 ± 0.005	0.003 ± 0.001	0.214

*Total number of segregating sites
†Total number of mutations
‡Overall mean diversity with the standard deviation calculated by DnaSP
#Estimated mutation rate using segregation sites
§Number of synonymous substitutions per site from the overall mean of sequence pairs
♯Number of non-synonymous substitutions per site from the overall mean of sequence pairs
ΨRatio of dN/Ds used to determine the selective pressure for coding regions

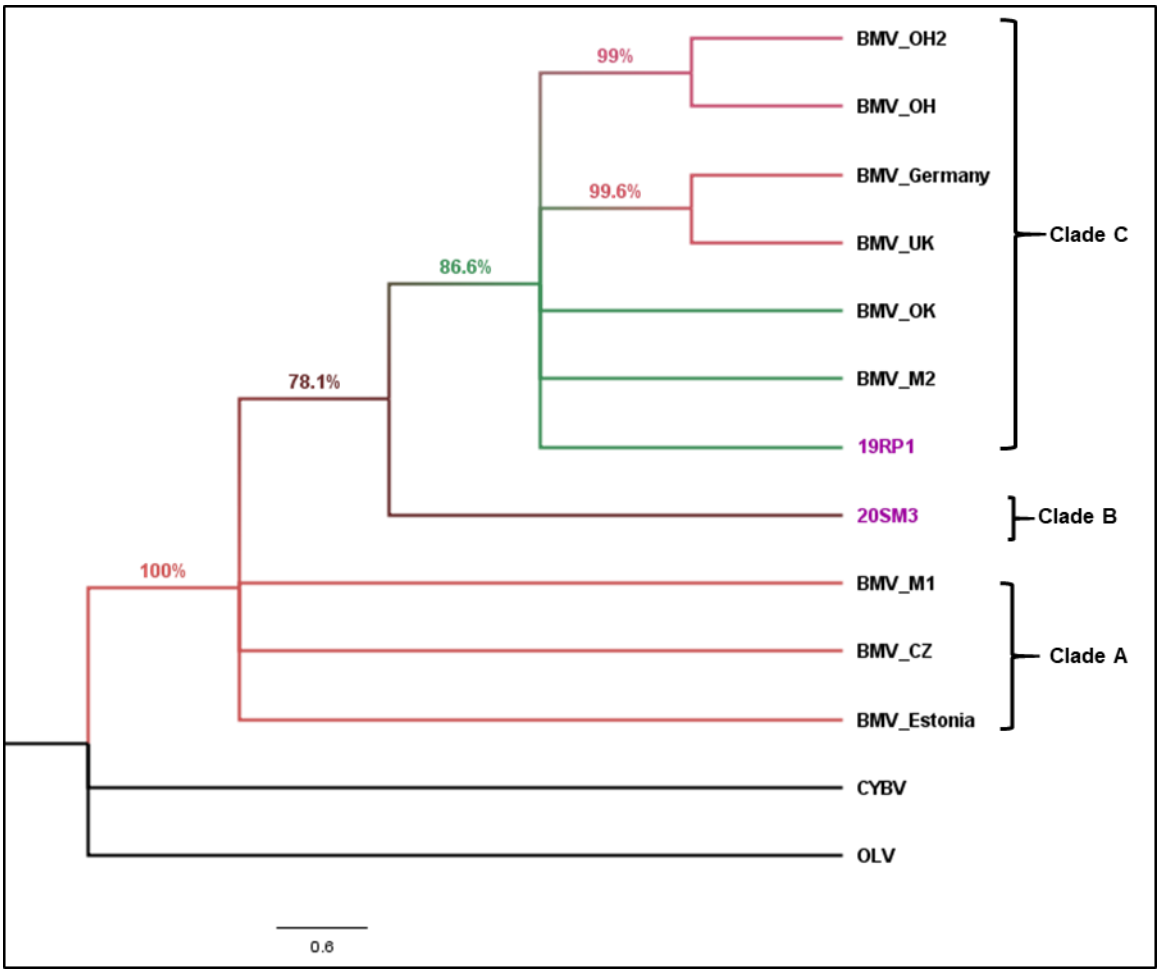


Figure 3.4. Phylogenetic tree of brome mosaic virus (BMV) isolates based on the coding sequence alignment of RNA1a sequenced in this study (highlighted in purple text) and selected strains retrieved from GenBank. The phylogenetic tree was made using the maximum likelihood analysis with a TN93 + G substitution model conducted in MEGA X (Kumar et al., 2018). The tree with the highest log likelihood (-11086.86) is shown. The percentage of replicate trees in which the associated taxa clustered together based on 1000 bootstrap replicates is presented. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Cassia yellow blotch virus (CYBV) and Olive latent virus2 (OLV) were used as outgroups in the analysis. Brackets on the right side indicate the taxa clustered in BMV clades A to C.

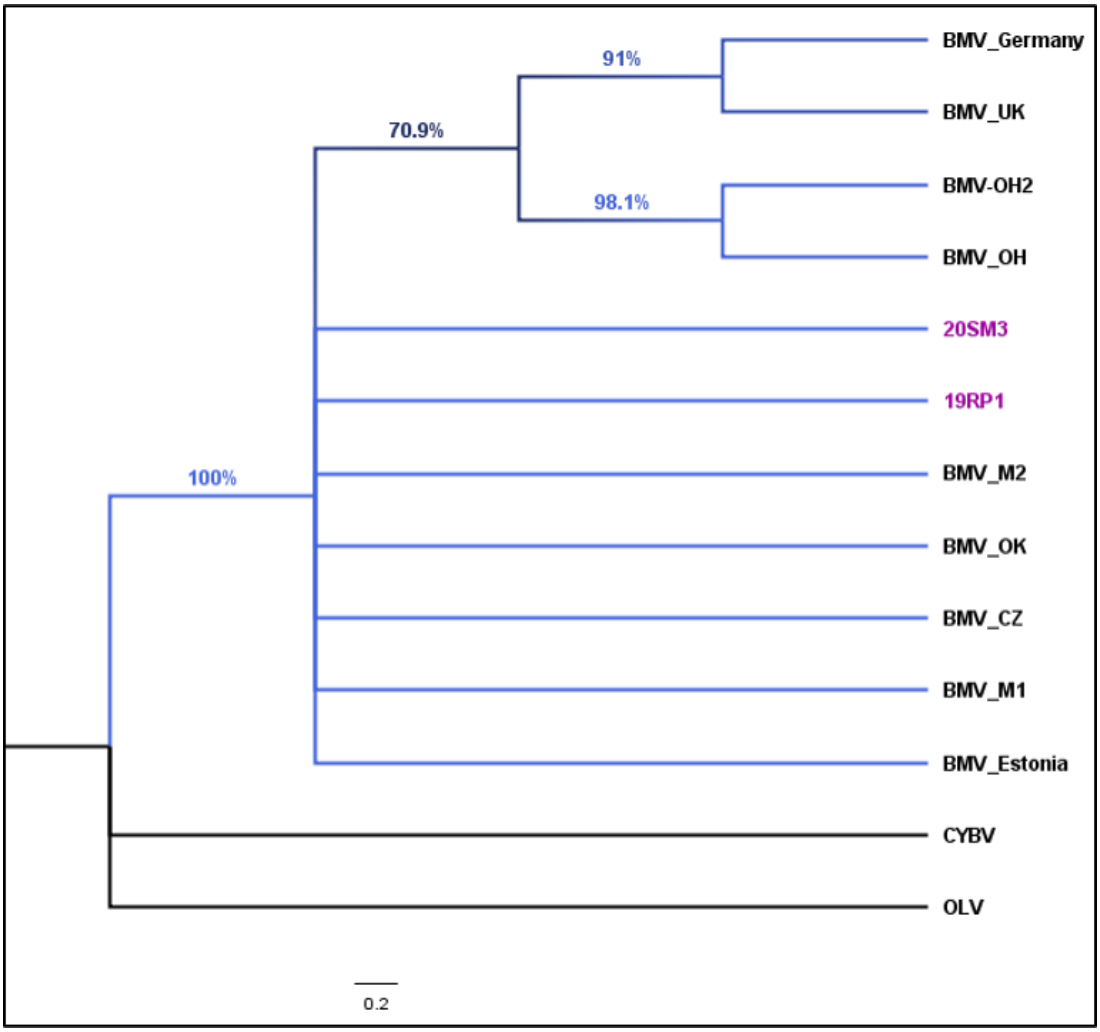


Figure 3.5. Phylogenetic tree of brome mosaic virus (BMV) isolates based on the coding sequence alignment of RNA2a sequenced in this study (highlighted in purple text) and selected strains retrieved from GenBank. The phylogenetic tree was made using the maximum likelihood analysis with a TN93 + G substitution model conducted in MEGA X (Kumar et al., 2018). The tree with the highest log likelihood (-9746.52) is shown. The percentage of replicate trees in which the associated taxa clustered together based on 1000 bootstrap replicates is presented. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Cassia yellow blotch virus (CYBV) and Olive latent virus 2 (OLV) were used as outgroups in the analysis.

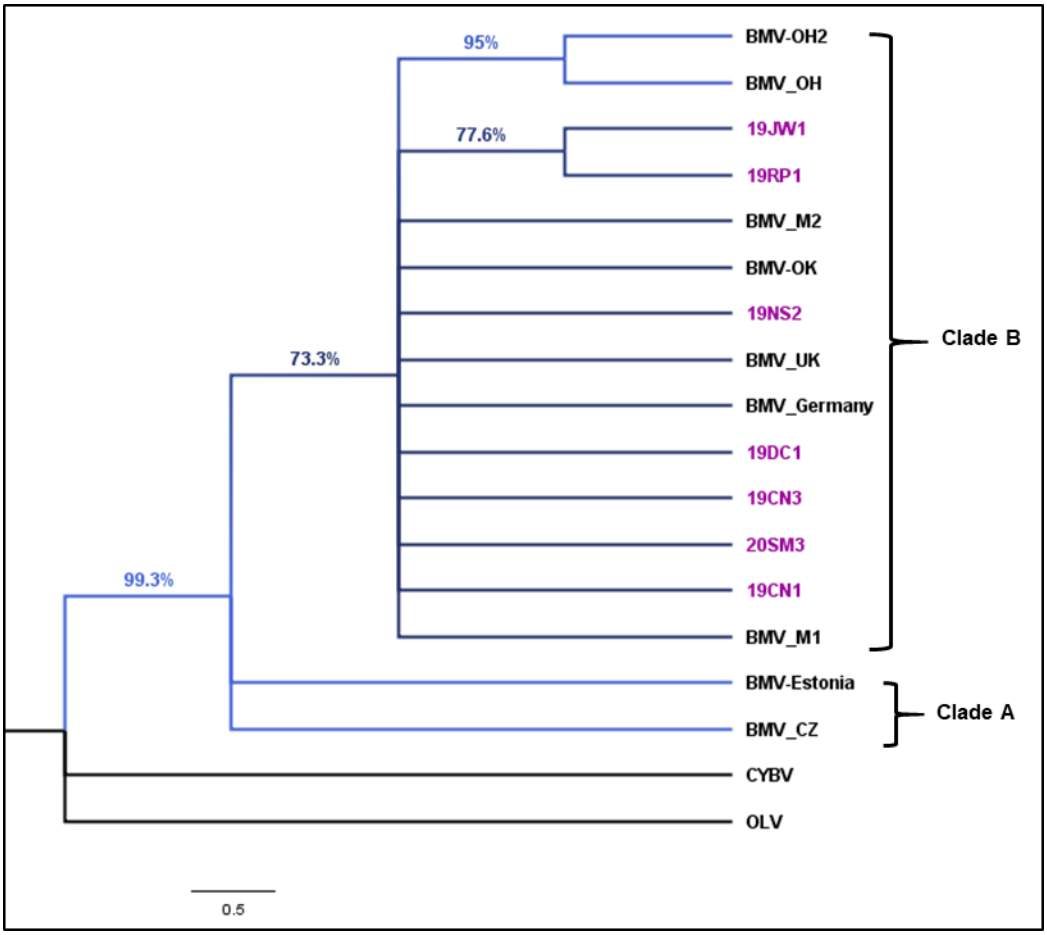


Figure 3.6. Phylogenetic tree of brome mosaic virus (BMV) isolates based on the coding sequence alignment of RNA3a (movement protein) sequenced in this study (highlighted in purple text) and selected strains retrieved from GenBank. The phylogenetic tree was made using the maximum likelihood analysis with a K2 + I substitution model conducted in MEGA X (Kumar et al., 2018). The tree with the highest log likelihood (-4171.53) is shown. The percentage of replicate trees in which the associated taxa clustered together based on 1000 bootstrap replicates is presented. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Cassia yellow blotch virus (CYBV) and Olive latent virus 2 (OLV) were used as outgroups in the analysis. Brackets on the right side indicate the taxa clustered in BMV clades A and B.

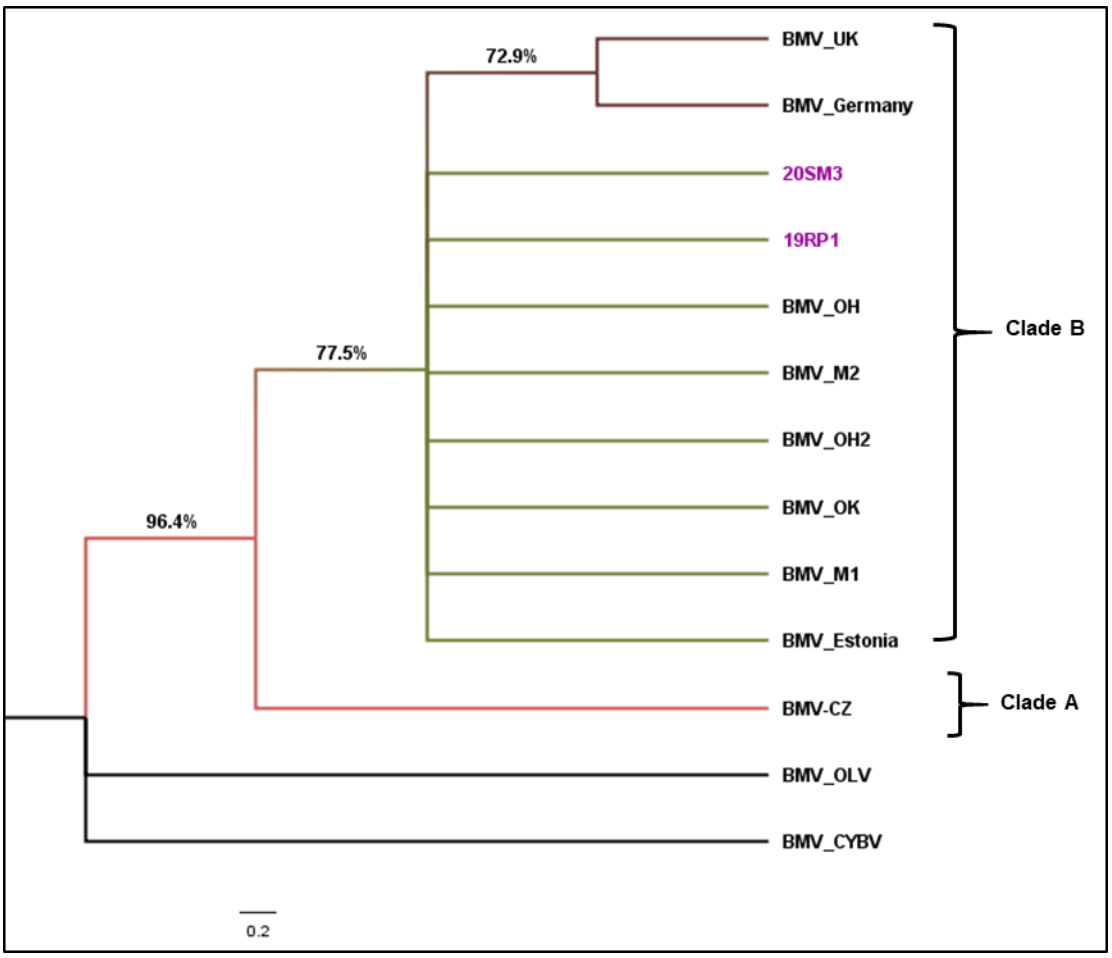


Figure 3.7. Phylogenetic tree of brome mosaic virus (BMV) isolates based on the coding sequence alignment of RNA4 (coat protein) sequenced in this study (highlighted in purple text) and selected strains retrieved from GenBank. The phylogenetic tree was made using the maximum likelihood analysis with a K2 + I substitution model conducted in MEGA X (Kumar et al., 2018). The tree with the highest log likelihood (-2359.86) is shown. The percentage of replicate trees in which the associated taxa clustered together based on 1000 bootstrap replicates is presented. The posterior probability of 70% was the cutoff value and branches not supported were collapsed. Cassia yellow blotch virus (CYBV) and Olive latent virus 2 (OLV) were used as outgroups in the analysis. Brackets on the right side indicate the taxa clustered in BMV clades A and B.

Chapter 4 - A reproducible methodology for absolute viral quantification and viability determination in mechanical inoculations of wheat streak mosaic virus

Nar B. Ranabhat, Myron A. Bruce, John P. Fellers, and Jessica L. Shoup Rupp

Received: 20 December 2021/Accepted 25 March 2022

(Ranabhat et al. 2022, Tropical Plant Pathology, DOI: https://doi.org/10.1007/s40858-022-00507-y.)

Abstract

Wheat streak mosaic virus (WSMV) is a common wheat virus causing economic losses to production in the Great Plains of North America. Reproducible inoculation of WSMV by mechanical methods is essential to evaluate the resistance in breeding lines and relies on successful inoculation and infectivity of the virus particles. We used reserve transcription-quantitative PCR (RT-qPCR) for absolute quantification of viral genome copy numbers in both WSMV inoculum and in infected wheat leaves. A time-course study was designed to determine the viability of WSMV in inoculum over time as well as the copy number related to the phenotypic rating scale. In the phosphate inoculation buffer, WSMV was stable with average viral genome copy number $1.86 \times 10^6 \pm 4.85 \times 10^5$. Plants inoculated with this inoculation buffer using finger rub mechanical inoculation contained WSMV genome copy numbers in the infected leaves ranging between 2.66×10^4 and 4.69×10^6 at 21 to 28 days post-inoculation. Viral copy numbers were statistically similar between leaves inoculated immediately and those inoculated at later time points. There was a weak linear relationship between phenotypic rating score and copy number in infected leaves with the linear model explaining 40 % of the variability ($R^2 = 0.40$)

indicating the difficulty in disease assessment based solely on phenotypic symptoms. This work describes an accurate methodology to quantify virus concentration in the inoculum and infected plants, as well as emphasizes the demand for absolute measurement of virus load to validate the subjective assessment for unbiased viral disease assessment.

Introduction

Wheat streak mosaic virus (Family: *Potyviridae*; Genus: *Tritimovirus*) is one of the most common cereal viruses infecting wheat (*Triticum aestivum* L.) in the North American Great Plains (Burrows et al., 2009). Wheat streak mosaic virus (WSMV) has also been reported across the globe, causing sporadic epidemics on wheat (Ellis et al., 2003; French and Stenger, 2003; Navia et al., 2013). In the Great Plains, WSMV infections can cause wheat yield losses that range from 1 to 5% annually, however total loss in localized fields is common (Burrows et al. 2009; Appel et al. 2015; Ranabhat, N. Kansas State University, USA, 2019 personal observation). In 2017, Kansas wheat producers lost 19.2 million bushels of wheat worth $76.8 million due to WSMV (KSwheat 2017). WSMV symptomology includes yellow streaked leaves with a mosaic pattern, stunted growth, reduced root mass, and decreased yield (Price et al., 2010b; Rahman et al., 1974).

WSMV, along with two other viruses triticum mosaic virus (Family: *Potyviridae*; Genus: *Poacevirus*) and High Plains wheat mosaic emaravirus (Family: *Fimoviridae; Genus: Emaravirus*) constitute the wheat streak mosaic (WSM) complex. All three viruses are primarily transmitted by the eriophyid wheat curl mite (*Aceria tosichella* Kiefer, (Seifers et al., 1997; Slykhuis, 1973; Tatineni et al., 2009). WSMV and Triticum mosaic virus (TriMV) can also be transmitted mechanically (Byamukama et al., 2014; Wosula et al., 2018). Commercial wheat cultivars with genetic resistance to WSMV are available; however, resistance is limited to three

major resistance genes; *Wsm1*, *Wsm2*, and *Wsm3* (Friebe et al., 2011; Graybosch et al., 2009; Lu et al., 2011b). Therefore, reliable identification of new germplasm with WSMV resistance and the subsequent development of new resistant cultivars through breeding programs is essential for the sustainable management of WSMV.

Plant cultivars that are screened for virus resistance are most often inoculated mechanically (Lu et al., 2011b; Seifers et al., 2007b, 2006; Wosula et al., 2018) except those viruses which can't be mechanically inoculated. Reproducible inoculation of the virus by mechanical methods is necessary to evaluate the level of resistance in breeding lines. Inoculation requires appropriate delivery of the virus to wheat and success largely depends on the infectivity of the virus particle. The effectiveness of the inoculum depends on concentration and its viability to systemically infect plants. Traditionally, the resistance level of cultivars inoculated with the virus is based on phenotypic symptom assessment (DeWolf et al. 2019; Johnson et al. 2019; Marburger et al. 2018; Rupp et al. 2014). There are currently no recognized standard area diagrams for the assessment of wheat streak mosaic, adding greater subjectivity phenotypic scoring. This is further complicated by the difficulty of accurately capturing symptoms to provide a reference. However, rating based only on symptom assessment during virus resistance screening cannot provide an accurate assessment as mild or no symptoms can have a relatively higher virus titer or vice versa (Lecoq et al., 2004; Ranieri et al., 1993).

Previous work has shown that WSMV has differing infection rates (Byamukama et al., 2012; Oliveira-Hofman et al., 2015; Tatineni et al., 2010b). The relative level of viral genomes was calculated by comparing the threshold cycle (Ct) values of virus-infected samples with the normalized expression of Ct values of wheat 18S RNA(Tatineni et al., 2019, 2010b). Absolute viral genome copy number is more accurate quantification of viral load in both viral inoculum

and infected tissue. Proper quantification of viral load provides actual host-virus interaction and helps in making an unbiased decision during the breeding program. The following work was designed with the overall objective of developing a reproducible method of absolute quantification of WSMV and determination of the inoculum viability and infectivity over time. We hypothesized the initial viral genome copy number in inoculum influences the final virus titer *in planta* and the virus titer in inoculum is unaffected by the time measuring the titer immediately and after a few hours post inoculum preparations. In this study, we used sensitive and robust SYBR green RT-qPCR assay for absolute quantification of viral genome copy number in the inoculum and inoculated wheat, as well as its relationship to the phenotypic rating scale. Accurate quantification of virus load will provide an unbiased plant-virus interaction evaluation for those breeding lines having no or mild symptoms but have higher virus titer as these lines contribute to virus spreads in the fields.

Materials and methods

Source of viral inoculum

WSMV (KSMHK isolate, GenBank MK318280.1 Fellers et al. 2019) was maintained in the susceptible winter wheat cultivar 'Tomahawk' (AgriPro, 1990) in greenhouse conditions (16 to 20°C and 16/8 light/dark cycles) in the Applied Wheat Pathology Laboratory, Kansas State University, Manhattan, KS. Wheat plants were inoculated by using the finger rubbing technique (Rupp et al. 2014). Tomahawk seeds were placed in cone-tainers (3.8 cm diameter, 21 cm depth, Ray Leach, Model SC10, Stuewe & Sons, OR, USA) containing premium potting soil (Baccto®, sphagnum peat bog, Michigan Peat Co, TX, USA). Wheat plants were grown in a growth chamber at 18℃ - 20℃ (day) to 16℃ (night) temperature, with 16:8 hours, light: dark regimen.

Virus inoculum was prepared by macerating 0.45 g tissue per 15 ml ice-cold 0.01 M sodium phosphate buffer, pH 7.2 (1:33.3 [wt/vol] tissue: buffer) using a chilled ceramic mortar and pestle. Inoculum viability was tested in a time-course study where plants were inoculated at 0, 2, 6, 10, 24, 48, 72, and 96 hours after inoculum preparation. The inoculum was stored at 4℃ after preparation for successive inoculation at different times. Wheat plants inoculated after 10 h were of different ages by day(s). Forty plants were arranged in a completely randomized design with four subsamples (individual plants) per treatment per biological replication. Each time point consisted of four inoculated plants and one mock-inoculated (buffer only) plant as a negative control. The entire experiment was repeated twice for a sum total of 80 plants in two independent biological replicates. In each biological replicate, the inoculum was prepared independently as described. For each time point, the second leaf of plants at the three leaves-stage was dusted lightly with carborundum (Fisher Scientific, MA, USA) and 40 µl of inoculum was placed over the carborundum. The leaf was pinched between the thumb and forefinger and the virus inoculum was pulled down the length of the leaf three to four times to increase the chances of infection (Rupp et al. 2014). For each time point, 500 µl of inoculum was flash-frozen in liquid nitrogen for reserve transcription-quantitative PCR (RT-qPCR). The youngest Feekes stage 4 leaf was sampled from each plant after 21 days of inoculation for RT-qPCR analysis and phenotypic rating. Leaf samples were flash-frozen in liquid nitrogen and stored at -80℃ until RNA extraction. The phenotypic rating score of the sample was done following the 1-9 Kansas State University rating scale (Supplementary Figure C1) (Rupp 2015).

RNA extraction and reserve transcription

Leaf tissue samples from each plant (total 4 plants per biological replication) were pooled and the pooled sample was considered one sample per time-point for each biological replication.

That means four plants of each treatment per replication were pooled for RT-qPCR analysis (n = 8 per treatment, two such pools for each treatment). The leaf sample was subjected to RNA extraction. For RNA extraction from inoculum, 50 µl of inoculum solution was used as one sample per time-point for each biological replication (two replications per treatment). Total RNA from each sample was extracted with the *mir*Vana RNA extraction kit (Ambion Catalog number: AM1560, ThermoFisher, MA, USA) according to the manufacturer's instructions for total RNA. The RNA concentration was measured spectrophotometrically by a NanoDrop spectrophotometer (NanoDrop Technologies, Rockland, DE, USA). Extracted total RNA (7 µg) was cleaned with the Turbo DNase-Free ^{TM} kit (AM 1907, Ambion®, ThermoFisher, MA, USA) according to the manufacturer's instructions in a 50 µL reaction.

DNase treated total RNA (1 µg) was used to synthesize the first-strand cDNA in 20 µL reaction volume by using a 5x iScript supermix (BioRad, La Jolla, CA, USA) according to the manufacturer's instructions and reaction conditions. Primers for RT-qPCR were designed from the gene sequence of WSMV KSMHK (GenBank MK318280.1) using Integrated DNA Technologies, Primer Quest (www.idtdna.com) with the parameters of Tm 58 °C, 18 to 22 nt, and 100 - 200 bp amplicon size. Primer efficiencies were tested before use and primers with efficiencies between 90 to 110% were used in the experiment on a four-point dilution series.

Cloning and sequencing of WSMV genome fragment

The first-strand cDNA obtained from RNA extracted from symptomatic leaf sample was used as a template for PCR. To obtain a large amplicon of the WSMV genome, WSMV-MHK3-F (forward) and WSMV-MHK1-R (reverse) primers were used to produce a 2,264 bp product covering the CI and VPg-NIa (3398 to 5644 bp) region of the WSMV genome (Table 4.1). PCR was performed using a 2x PCR master mix (Promega Corp., Madison, WI, USA) in a 25 µl of

volume according to the manufacturer's instructions. The temperature profile for PCR included an initial denaturation at 94°C for 3 min followed by 35 cycles of 1 min denaturation at 94°C, 30 seconds annealing at 55°C, and 2 min extension at 72°C with final extension for 7 min at 72°C. Amplified products were analyzed by agarose gel electrophoresis.

The PCR product was purified using a QIAquick gel extraction kit (QIAGEN, USA) according to the manufacturer's instructions. The purified amplicon was cloned into a topoisomerase-activated vector, pCR® 2.1-TOPO (Invitrogen, Life Technologies, CA, USA) according to the manufacturer's instructions. Transformation of the vector was performed using TOP10 competent cells (One Shot®, Invitrogen, CA, USA). Plasmid DNA was extracted from cultures using a plasmid purification kit (QIAGEN spin miniprep kit, QIAGEN, USA) according to the company's instructions. The plasmid was sequenced in MCLAB, CA, USA by using universal sequencing primer M13 forward (-20) and M13 reverse. Plasmid sequenced was blasted by using the NCBI BLAST tool (Altschul et al., 1990) after being obtained from MCLAB, CA.

Plasmid dilution, standard curve, and reverse transcription qPCR

A standard curve method was used to determine the absolute WSMV copy number in samples. A serial dilution of plasmid DNA carrying the previously cloned WSMV fragment was used as a template in a 5 point, 10-fold dilution series. The mass of plasmid and concentration of the plasmid solution was used to calculate the desired copy number and the dilution series was calculated as described in Table 4.2. The plasmid solution with calculated copy number 3,000,000 was serially diluted (Table 4.2) and used as a template to establish a standard curve. The RT-qPCR was performed in a BioRad CFX96 Real-Time System (BioRad, La Jolla, CA, USA) with a total reaction volume of 25 µl. Each reaction included 12.5 µl of iQ SYBR Green

Supermix (BioRad, La Jolla, CA, USA), 50 ng of cDNA, 10 pmol of MHK3-F and MHK3-R primer (Table 4.1), and water added to a total volume of 25 µl according to Neugebauer et al. 2018. Technical replicates were prepared in triplicate. The thermocycler conditions were 95°C for 3 min for initial denaturation, followed by 35 cycles of 10 seconds at 95°C denaturation, and 30 seconds anneal/extend at 60°C. The run was completed with a melt curve of 65°C to 95°C heating for 0.5°C increments for 5 seconds. For each biological replication of the experiment, the plasmid dilution series reactions were loaded on the same plate to build the appropriate standard curve. Data of each reaction was analyzed by BioRad Real-Time System (CFX96, version 3.01215.0601) to determine the quantification cycle (Cq) values, PCR efficiency, and the standard curve.

Statistical analysis

The WSMV copy number in each sample was calculated by plotting the quantification cycle (Cq) values of the samples into the standard curve generated from the plasmid dilution series by BioRad CFX96 software. Plotted copy numbers of samples within the range of the dilution series standard curve descending from 3×10^6 to 3×10^2 were included for further analysis. In this experiment, data below the lower boundary of the standard curve were excluded from the analysis to avoid extrapolation. Data of mock-inoculated control Cq values were below the lower boundary of the standard curve were excluded from the analysis. The absolute copy number of the samples from two biological replicates was analyzed using the SAS PROC mixed procedure (SAS Institute Inc., Cary, NC, USA). Treatment was used as a fixed effect and replication was used as a random effect. Tukey's multiple comparison test (Tukey, 1949) was used to compare the mean difference in the WSMV copy number in inoculum among treatments. The mean comparison of WSMV copy number in infected wheat leaves was done by using

Dunnett's method (Dunnett, 1955) to compare the treatment means against treatment, 0 h was used as a baseline group. The average phenotypic rating scores from each treatment were regressed with the viral copy number of the infected leaf. Log transformed copy number data and square root transformed rating data were used for analysis to satisfy the constant variance and normality. An influence observation was detected from Cook's distance and removed from the analysis (Ramsey and Schafer, 2012).

Results

Plasmid and standard curve determination

Cloned plasmid fragments were verified as the sequence aligned to the genome sequence of WSMV (isolates KSMHK, Accession number MK318280.1) of the desired WSMV- CI-NIa fragment in the plasmid (3483-5573 nt in the reference sequence) (Supplementary Table C2). The standard curve was prepared using a 10-fold serial dilution of plasmid copy number descending from 3×10^6 to 3×10^2 molecules and it showed a strong linear relationship with the value of R^2 0.999. The amplification efficiencies were ranging from 96.9% to 103.8% between replicates (Supplementary Table C3 A, B). While the standard curve of each assay was used to calculate the copy number of each sample, the standard curve data from different assays were also plotted to calculate the overall R^2 of 0.982, indicating good reproducibility of the plasmid, primer, and the technique (Supplementary Table C4). A single sharp peak at the melting temperature of 80.0°C (Supplementary Table C3 C, D) shows the specific PCR products amplified with primer set MHK1. Similar results were observed in each replication.

Assessment of the WSMV copy number in the inoculum

The WSMV inoculum genomic RNA copy numbers in the inoculum ranged between 1.15×10^6 and 3.28×10^6 and varied with time post preparation ($F = 4.36$, $P = 0.001$). RNA copy

number at 0 h was similar to numbers at 2 h, 6 h, 24 h, and 48 h ($P > 0.05$, Table 4.3) but it was higher than the RNA copy number of 10 h, 72 h, and 96 h ($P< 0.05$, Table 4.3).

Estimation of the absolute WSMV copy number in inoculated wheat leaves

The absolute genomic RNA copy numbers in the WSMV infected leaves inoculated with inoculum prepared at different times ranged between 2.66×10^4 and 4.69×10^6. Dunnett's test showed that the WSMV RNA copy numbers in wheat leaves inoculated at different time points were statistically similar to the leaf inoculated at 0 h ($P > 0.05$, Table 4.3). As expected, Cq values of mock-inoculated control were outside of the lowest point in the dilution series therefore those data were not included in the table.

The relation between absolute WSMV copy number and the phenotypic rating score

In order to establish the relationship between viral copy number and phenotypic rating score, a linear regression of the average rating scores with the average viral genome copy number of infected leaves at each time point of inoculation post preparation was performed. There was a weak linear trend of phenotypic rating score and WMSV copy number (Figure 4.1). The value of the regression coefficient of determination $R^2 = 0.40$ (Figure 4.1). The typical WSMV symptoms on infected leaves varied with the time of inoculation (Figure 4.2).

Discussion

The results presented in this study describe quantifiable methodologies that can be applied to plant-virus resistance evaluation studies and highlights the necessity of using quantitative methods to facilitate accurate disease assessment in plant breeding programs especially to those breeding line having no or mild symptoms with higher virus titer or vice-versa. Accurate measurement of viral load provides the quantitative host-virus interaction and

helps to make the accurate decision of varietal selection. The information on virus stability helps to make the decision of inoculum preparation time while screening the breeding nursery with mechanical inoculation of plant viruses.

Viral genome copy numbers were determined using RT-qPCR. The RT-qPCR has previously been used for WSMV detection (Price et al., 2010a) and examined the relative quantification of WSMV compared with wheat reference genes (Tatineni et al. 2010, 2019) but not the absolute viral copy number. Here, a quantitative method was used to calculate the absolute viral copy number using a WSMV genome fragment-containing plasmid to simulate viral RNA strands of a known copy number. The dilution series determined that RT-qPCR could detect the plasmid at dilutions of 3×10^2 to 3×10^6 particles, covering a wide dynamic range of concentrations (4 orders of magnitude). A strong linear relationship with the value of coefficients of determination, $R^2 = 0.999$, and the presence of a single fluorescence peak in the melting curve analysis support the primer specificity and indicating good reproducibility of the plasmid as a standard and the RT-qPCR technique to determine viral copy number in the inoculum and *in planta*.

In mechanical infections, viral movement and titer are affected by what cells are infected, how many cells are infected, and how many virions are active. The knowledge of the viability of the virus particles in the homogenized buffer over time will particularly assist researchers and breeders during field or other large-scale inoculations. Generally, it is assumed that inoculum has a viability limit. Our data revealed that WSMV was stable in the phosphate buffer for hours after post inoculum preparation as the inoculum contained more than a million viral genome copies per 50 µl of inoculum solution at each time point tested. Virus copy number at 10h showed deviation from the tendency, which might need further inquiry, but data were consistent over all

the subsamples within the replicate and no outlier was detected statistically. This finding indicates that inoculum can be made in the laboratory at least 6 hours from preparation to inoculation with no loss of viability when considering both viral copy number and adequate ability to phenotype.

The number of virus particles in the initial inoculum influences the final virus titer in inoculated wheat leaves. Our data showed that the viral copy number in wheat leaves inoculated immediately after inoculum preparation was statistically similar to the viral copy number at other time points. This result further supports the stable viability of WSMV KSMHK in the described buffer. The variability of the end viral titer in the inoculated leaves depends on several factors such as inoculation efficiency, the number of wounds made by carborundum during inoculation, and physiological conditions of the plant during infection (Roenhorst et al., 1988). The success of mechanical inoculation is further impeded by heavily damaged cells during inoculation resulting in inducing stress response proteins and release of vacuole contents, further producing antagonists for virus infection mediated by plant hormones (Savatin et al., 2014).

Phenotyping is a common screening method used by plant breeders to develop new plant cultivars. It is time-sensitive and considered a bottleneck for a breeding program (Furbank and Tester, 2011). Visual rating scales are subjective and require highly skilled workers. Our data revealed that the linear model robustness calculated by the coefficient of determination was only 40%. This implies that only 40 percent of the sample variability of copy numbers and phenotypic rating scores was explained by the estimated regression model. This implies that the relationship between viral copy numbers and visual rating has a weak linear relationship. R^2 provides an estimate of the proportion of variability between copy number and rating score and explains the dynamics of the relationship between them.

Although analysis might warrant a larger data set to claim the non-linear relationship between copy number and rating scale, it demands the quantitative methods of viral titer measurement to enable the accurate screening of breeding material. Plants infected with pathogens can show inconsistent in QTL-type resistance phenotypes that need careful phenotypic analysis (Poland et al., 2009). Roossinck, 2012 found that plants infected in a natural ecological setting often lacked symptoms and titer levels didn't correlate. Scoring based only on phenotypic symptom expression could select the lines having high virus titer and low phenotypic symptom expression. Plants with low or no symptoms but having high titer can serve as disease reservoirs and are epidemiologically significant. Therefore, it is important to validate phenotypic assessment with absolute quantification of viral load for at least selected lines during the development of virus-resistant germplasm through a breeding program.

Copy number (Table 4.3) showed comparatively high virus accumulation (high mean copy number) in wheat leaves inoculated after 48 h of post inoculum preparation but did not display a 'severe' phenotype (Figure 4.2). This inconsistency between viral load and symptom severity might be due to several molecular interactions between host and virus, one possible reason could be after 10 h of post inoculum preparation the aggressiveness of the virus particles might decrease. However, in-depth molecular analysis on the interaction between viral proteins and host gene expression might add a clearer picture. The phenomenon of weak symptoms and high virus accumulation and vice-versa is common in plant viruses and viroids (Flores et al., 2016). Therefore, using only phenotypic assessment of cultivars for breeding programs might increase the likelihood of misleading results as high virus titer correlates with high yield loss and vice-versa.

This work is the first to examine the effect of the initial viability (stability of virus in buffer) and the concentration of WSMV particles used in inoculated studies. Using the highly sensitive absolute measurement of viral copy numbers to validate the visual rating score will provide an accurate disease severity assessment thus sets an improved standard in virus resistance breeding.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of molecular biology 215, 403–410.

Appel, J., Dewolf, E., Todd, T., Bockus, W., 2015. Preliminary 2015 Kansas Wheat Disease Loss Estimates; Kansas Cooperative Plant Disease Survey Report; Kansas Department of Agriculture: Manhattan, KS, USA, 2015.

Burrows, M., Franc, G., Rush, C., Blunt, T., Ito, D., Kinzer, K., Olson, J., O'Mara, J., Price, J., Tande, C., others, 2009. Occurrence of viruses in wheat in the Great Plains region, 2008. Plant Health Progress 10, 14.

Byamukama, E., Tatineni, S., Hein, G.L., Graybosch, R.A., Baenziger, P.S., French, R., Wegulo, S.N., 2012. Effects of Single and Double Infections of Winter Wheat by *Triticum mosaic virus* and *Wheat streak mosaic virus* on Yield Determinants. Plant Disease 96, 859–864.

Byamukama, E., Wegulo, S., Tatineni, S., Hein, G., Graybosch, R., Baenziger, P.S., French, R., 2014. Quantification of yield loss caused by Triticum mosaic virus and Wheat streak mosaic virus in winter wheat under field conditions. Plant disease 98, 127–133.

DeWolf, E.D., Lollato, R. and Whitworth, R.J. 2019. Wheat Variety Disease and Insect Ratings 2019. Kansas State University, K-State Research and Extension, MF991. Wheat rating

Dunnett, C.W., 1955. A multiple comparison procedure for comparing several treatments with a control. Journal of the American Statistical Association 50, 1096–1121.

Ellis, M., Rebetzke, G., Mago, R., Chu, P., 2003. First report of Wheat streak mosaic virus in Australia. Australasian Plant Pathology 32, 551–553.

Fellers, J.P., Webb, C., Fellers, M.C., Shoup Rupp, J., De Wolf, E., 2019. Wheat virus identification within infected tissue using nanopore sequencing technology. Plant disease 103, 2199–2203.

Flores, R., Owens, R.A., Taylor, J., 2016. Pathogenesis by subviral agents: viroids and hepatitis delta virus. Current opinion in virology 17, 87–94.

French, R., Stenger, D.C., 2003. Evolution of Wheat streak mosaic virus: dynamics of population growth within plants may explain limited variation. Annual review of phytopathology 41, 199–214.

Friebe, B., Liu, W., Qi, L., Wilson, D., Raupp, W., Poland, J., Bowden, R., Fritz, A., Seifers, D., Gill, B., 2011. Notice of release of KS12WGGRC59 wheat streak mosaic virus-and Triticum mosaic virus-resistance wheat germplasm. Annual Wheat Newsletter 57, 280.

Furbank, R.T., Tester, M., 2011. Phenomics–technologies to relieve the phenotyping bottleneck. Trends in plant science 16, 635–644.

Graybosch, R.A., Peterson, C., Baenziger, P.S., Baltensperger, D.D., Nelson, L.A., Jin, Y., Kolmer, J., Seabourn, B., French, R., Hein, G., others, 2009. Registration of 'Mace'hard red winter wheat. Journal of Plant Registrations 3, 51–56.

Johnson, J., Haley, S., Jones, J., Asfeld, E., Meyer, R., Trujillo, W., Kann, D., Roesch, K., Spring, J., Larson, K., Vigil, M., Pettinger, B. 2019. Colorado Winter Wheat Variety Performance Trials. Tech. Rep. 19-2. Agricultural Experiment Station, Colorado State University, Ft. Collins, Colorado

KS wheat. 2017. Kansas wheat. Wheat streak mosaic is timely issue at wheat. http://kswheat.com/news/2017/08/09/wheat-streak-mosaic-is-timely-issue-at-wheatu-event (accessed on 2/12/2022)

Lecoq, H., Moury, B., Desbiez, C., Palloix, A., Pitrat, M., 2004. Durable virus resistance in plants through conventional approaches: a challenge. Virus research 100, 31–39.

Lu, H., Price, J., Devkota, R., Rush, C., Rudd, J., 2011. A dominant gene for resistance to Wheat streak mosaic virus in winter wheat line CO960293-2. Crop science 51, 5–12.

Marburger, D., Calhoun, R., Carver, B., Hunger, B., Watson, B., and Gillespie, C. 2018. Oklahoma small Grains variety performance tests 2017-2018. OSU cooperative extension service, CR-2141 and CR-2143. Oklahoma State University, Stillwater, Oklahoma

Navia, D., de Mendonça, R.S., Skoracka, A., Szyd\lo, W., Knihinicki, D., Hein, G.L., da Silva Pereira, P.R.V., Truol, G., Lau, D., 2013. Wheat curl mite, Aceria tosichella, and transmitted viruses: an expanding pest complex affecting cereal crops. Experimental and Applied Acarology 59, 95–143.

Neugebauer, K.A., Bruce, M., Todd, T., Trick, H.N., Fellers, J.P., 2018. Wheat differential gene expression induced by different races of Puccinia triticina. PloS one 13.

Oliveira-Hofman, C., Wegulo, S.N., Tatineni, S., Hein, G., 2015. Impact of Wheat streak mosaic virus and Triticum mosaic virus coinfection of wheat on transmission rates by wheat curl mites. Plant disease 99, 1170–1174.

Poland, J.A., Balint-Kurti, P.J., Wisser, R.J., Pratt, R.C., Nelson, R.J., 2009. Shades of gray: the world of quantitative disease resistance. Trends in plant science 14, 21–29.

Price, J., Smith, J., Simmons, A., Fellers, J., Rush, C., 2010a. Multiplex real-time RT-PCR for detection of Wheat streak mosaic virus and Tritcum mosaic virus. Journal of virological methods 165, 198–201.

Price, J., Workneh, F., Evett, S., Jones, D., Arthur, J., Rush, C., 2010b. Effects of Wheat streak mosaic virus on root development and water-use efficiency of hard red winter wheat. Plant Disease 94, 766–770.

Rahman, F., Ross, J.G., Gardner, W.S., 1974. Tolerance to Wheat Streak Mosaic Virus in Spring and Winter Wheat Cultivars 1. Crop Science 14, 178–180.

Ramsey, F., Schafer, D., 2012. The statistical sleuth: a course in methods of data analysis. Cengage Learning.

Ranieri, R., Lister, R.M., Burnett, P.A., 1993. Relationships between Barley Yellow Dwarf Virus Titer and Symptom Expression in Barley. Crop Science 33, 968–973.

Roenhorst, J., Van Lent, J., Verduin, B., 1988. Binding of cowpea chlorotic mottle virus to cowpea protoplasts and relation of binding to virus entry and infection. Virology 164, 91–98.

Roossinck, M.J., 2012. Plant virus metagenomics: biodiversity and ecology. Annual review of genetics 46, 359–369.

Rupp, J.L.S., 2015. RNA interference mediated virus resistance in transgenic wheat (PhD Thesis). Kansas State University.

Rupp, J.L.S., Simon, Z.G., Gillett-Walker, B. and Fellers, J.P. 2014. Resistance to Wheat streak mosaic virus identified in synthetic wheat lines. Euphytica 198, 223–229.

Savatin, D.V., Gramegna, G., Modesti, V., Cervone, F., 2014. Wounding in the plant tissue: the defense of a dangerous passage. Frontiers in plant science 5, 470.

Seifers, D., Martin, T., Harvey, T., Haber, S., 2007. Temperature-sensitive Wheat streak mosaic virus resistance identified in KS03HW12 wheat. Plant disease 91, 1029–1033.

Seifers, D., Martin, T., Harvey, T., Haber, S., Haley, S., 2006. Temperature sensitivity and efficacy of Wheat streak mosaic virus resistance derived from CO960293 wheat. Plant Disease 90, 623–628.

Seifers, D.L., Harvey, T.L., Martin, T., Jensen, S.G., 1997. Identification of the wheat curl mite as the vector of the High Plains virus of corn and wheat. Plant Disease 81, 1161–1166.

Slykhuis, J.T., 1973. *Aceria tulipae* keifer (*Acarina: Eriophyidae*) in relation to the spread of wheat streak mosaic. Readings in Insect-plant Disease Relationships 11, 394.

Tatineni, S., Alexander, J., Gupta, A.K., French, R., 2019. Asymmetry in Synergistic Interaction Between *Wheat streak mosaic virus* and *Triticum mosaic virus* in Wheat. MPMI 32, 336–350.

Tatineni, S., Graybosch, R.A., Hein, G.L., Wegulo, S.N., French, R., 2010. Wheat cultivar-specific disease synergism and alteration of virus accumulation during co-infection with Wheat streak mosaic virus and Triticum mosaic virus. Phytopathology 100, 230–238.

Tatineni, S., Ziems, A.D., Wegulo, S.N., French, R., 2009. Triticum mosaic virus: a distinct member of the family Potyviridae with an unusually long leader sequence. Phytopathology 99, 943–950.

Tukey, J.W., 1949. Comparing individual means in the analysis of variance. Biometrics 99–114.

Wosula, E., McMechan, A.J., Knoell, E., Tatineni, S., Wegulo, S.N., Hein, G.L., 2018. Impact of timing and method of virus inoculation on the severity of wheat streak mosaic disease. Plant disease 102, 645–650.

Table 4.1. Primers used in this study

Primer Name	Sequence 5’ to 3’	Tm (°C)	Amplicon size (bp)	Position in sequence*
WSMV-MHK1-F	TGGACCGATCGGATTAAG	58	107	5555-5573
WSMV-MHK1-R	TAGAAGTGCCAGTAT			5644-5662
WSMV-MHK3-F	CATGAGGCAACACAAGTAG	58	104	3398-3417
WSMV-MHK3-F	CCATCAAGTGGTGCATATC			3483-3502

*Nucleotide positions on the sequence of isolate WSMV KSMHK (GeneBank accession number MK318280.1, Fellers et al. 2019)

Table 4.2. Dilution series of plasmid from initial plasmid concentration and standard copy number*

Copy number	Mass in grams†	6µl of plasmid DNA solution was pipetted into each reaction (grams/µl)
3000000	$300000 \times 6.72 \times 10^{-18} = 2.0156 \times 10^{-11}$	$2.0156 \times 10^{-12} / 6 = 3.359 \times 10^{-12}$
300,000	$300000 \times 6.72 \times 10^{-18} = 2.0156 \times 10^{-12}$	$2.0156 \times 10^{-12} / 6 = 3.359 \times 10^{-13}$
30,000	$30000 \times 6.72 \times 10^{-18} = 2.0156 \times 10^{-13}$	$2.0156 \times 10^{-13} / 6 = 3.359 \times 10^{-14}$
3,000	$3000 \times 6.72 \times 10^{-18} = 2.0156 \times 10^{-14}$	$2.0156 \times 10^{-14} / 6 = 3.359 \times 10^{-15}$
300	$300 \times 6.72 \times 10^{-18} = 2.0156 \times 10^{-15}$	$2.0156 \times 10^{-15} / 6 = 3.359 \times 10^{-16}$

Source of plasmid DNA for dilution	Initial conc. (grams/µl) $\mathbf{C_1}$	Volume of plasmid DNA (µl) $\mathbf{V_1}$	Volume of diluent (µl)	Final volume (µl) $\mathbf{V_2}$	Final conc. (grams/µl) $\mathbf{C_2}$	Resulting copy number of 6µl plasmid
Stock	4.07×10^{-07}	10	990	1000	4.07×10^{-09}	N/A
Dilution 1	4.07×10^{-09}	10	990	1000	4.07×10^{-11}	N/A
Dilution 2	4.07×10^{-11}	8.3	91.7	100	3.359×10^{-12}	3000000
Dilution 3	4.07×10^{-13}	10	90	100	3.359×10^{-13}	300000
Dilution 4	3.359×10^{-13}	10	90	100	3.359×10^{-14}	30000
Dilution 5	3.359×10^{-14}	10	90	100	3.359×10^{-15}	3000
Dilution 6	3.359×10^{-15}	10	90	100	3.359×10^{-16}	300

*Calculation adapted from Applied Biosystems (2003)

† Mass of the 1 bp of the plasmid was calculated by using the average mass of one deoxy-nucleoside monophosphate pair (DNMP = 654 Da = 1.09×10^{-21} g). Therefore, mass of one 6164 bp plasmid = $6164 \times 1.09 \times 10^{-21}$ g = 6.72×10^{-18} g

Table 4.3. Absolute quantification of genomic RNA copies in wheat streak mosaic virus (WSMV) inoculum and the leaf of wheat (Tomahawk) infected with that inoculum during the different times after inoculum prepared obtained by SYBR green quantitative (RT-qPCR) using the standard curve of plasmid DNA

TAIP*	Copy number in inoculum†	Copy number in infected leaf††	Average phenotypic score‡‡
0 h	$3.28 \times 10^6 \pm 4.85 \times 10^5$a	$1.71 \times 10^{5‡} \pm 1.09 \times 10^5$	5.25 ± 5.3
2 h	$2.21 \times 10^6 \pm 4.85 \times 10^5$ab	$4.69 \times 10^5 \pm 1.09 \times 10^5$a	7.25 ± 2.4
6 h	$1.81\times 10^6 \pm 4.85 \times 10^5$ab	$1.34 \times 10^5 \pm 1.09 \times 10^5$ a	4.62 ± 5.1
10 h	$1.15\times 10^6 \pm 4.85 \times 10^5$b	$2.66 \times 10^4 \pm 1.09 \times 10^5$ a	4.25 ± 4.5
24 h	$1.88 \times 10^6 \pm 4.85 \times 10^5$ab	$3.68 \times 10^4 \pm 1.09 \times 10^5$a	1.25 ± 0.3
48 h	$1.91 \times 10^6 \pm 4.85 \times 10^5$ab	$4.46 \times 10^5 \pm 1.09 \times 10^5$ a	2.37 ± 1.9
72 h	$1.34 \times 10^6 \pm 4.85 \times 10^5$b	$1.57 \times 10^5 \pm 1.09 \times 10^5$ a	2.75 ± 0.3
96 h	$1.30\times 10^6 \pm 4.85 \times 10^5$b	$2.10 \times 10^5 \pm 1.09 \times 10^5$a	2.00 ± 1.4

*Time after inoculation preparation.

†Average copy number in inoculum from two replicates. Different letter within the copy number of inoculum column indicates significantly different groups of means based on Tukey's HSD test ($P < 0.05$)

†† Average copy number and standard error in infected wheat leaves from two biological replicates (Four plants of each treatment per biological replication were pooled for RT-qPCR analysis, therefore a total of eight plants were used per treatment). The same letter within the column of copy number in infected leaves indicates no significant difference between 0 h to other times based on Dunnett's method ($P < 0.05$)

‡Baseline control in Dunnett's method of multiple comparisons

‡‡Average phenotypic score obtained by following the Kansas State University standard scoring protocol (Rupp, 2015)

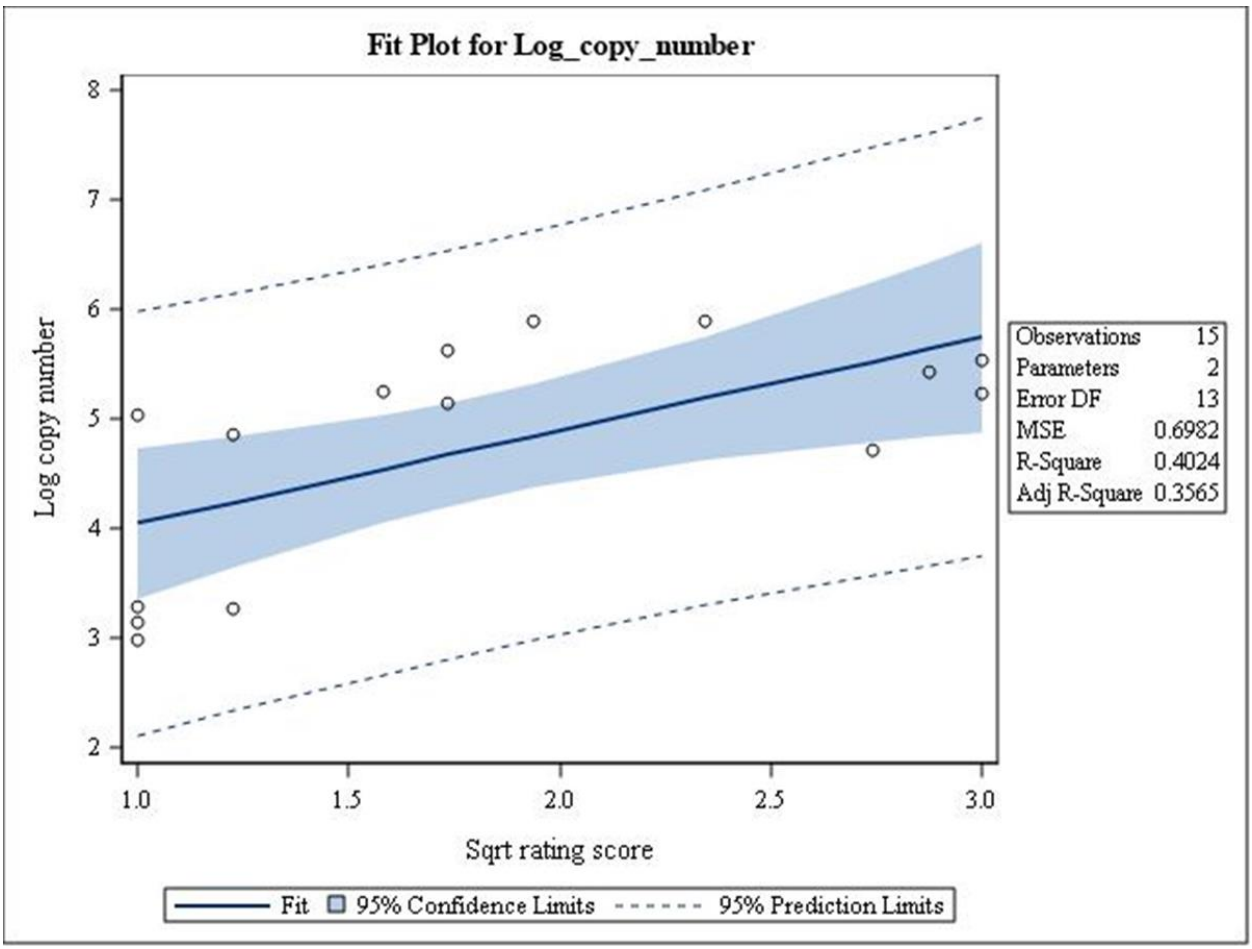


Figure 4.1. Regression of WSMV log copy number in inoculated leaves and phenotypic rating score. Blue shading around the fitted line represents 95% confidence limits. Sqrt = square root.

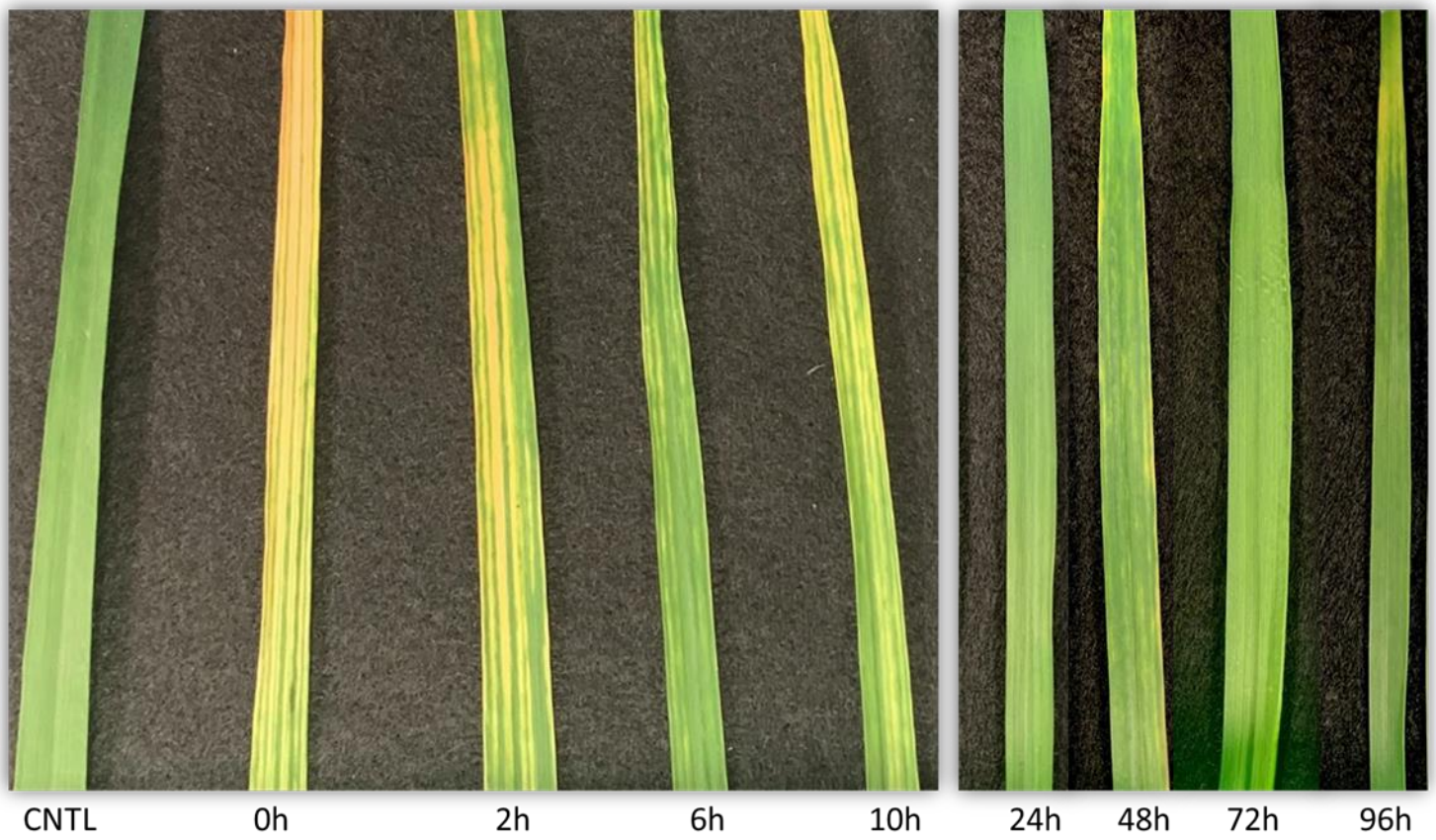


Figure 4.2. Wheat streak mosaic virus (WSMV) infected symptomatic leaves of wheat (Tomahawk) inoculated at a different time of 0 h to 96 h after inoculum preparation. CNTL = mock-inoculated control.

Appendix A - Supplementary Figures and Tables from Chapter 2

Appendix A. Tables and Figures

Supplementary Table A.1. List of sequences of cereal viruses retrieved from GenBank that were used as reference genomes to get consensus sequences

Name of virus	Accession number
Ageratum yellow leaf curl betasatellite	KC305091.1
Agropyron mosaic virus	NC_005903.1
Barley mild mosaic virus	AJ544268.1
Barley stripe mosaic virus RNA1	NC_003469.1
Barley stripe mosaic virus RNA2	NC_003481.1
Barley stripe mosaic virus RNA3	NC_003478.1
Barley yellow dwarf virus-PAS	NC_002160.2
Barley yellow dwarf virus-MAV	NC_003680.1
Barley yellow dwarf virus-GAV	KF523382.1
Barley yellow dwarf virus-PAV	EF043235.1
Barley yellow striate mosaic virus, polymerase (L) gene	FJ665628
Barley yellow striate mosaic virus, glycoprotein (G) gene	KP163565.1
Cereal yellow dwarf virus	EF521830.1
Hordeum mosaic virus	NC_005904.1
Maize streak virus	AF239960.1
Oat golden stripe virus RNA1	NC_002358.1
Oat golden stripe virus RNA2	NC_002357.1
Oat necrotic mottle virus	NC_005136.1
Tobacco mosaic virus Queensland	AF332868
Rice black streaked dwarf virus (S1)	KC134289.1
Rice black streaked dwarf virus (S2)	KC134290.1
Rice black streaked dwarf virus (S3)	KC134291.1
Rice black streaked dwarf virus (S4)	KC134292.1
Rice black streaked dwarf virus (S5)	KC134293.1
Rice black streaked dwarf virus (S6)	KC134294.1
Rice black streaked dwarf virus (S7)	KC134295.1
Rice black streaked dwarf virus (S8)	KC134296.1
Rice black streaked dwarf virus (S9) cds_AFX68415.1_1	KC134297.1
Rice black streaked dwarf virus (S9) cds_AFX68415.1_2	KC134297.1
Rice black streaked dwarf virus (S10)	KC134298.1
Soil borne wheat mosaic virus RNA1	KT736088.1
Soil borne wheat mosaic virus RNA2	KT736089.1
Wheat streak mosaic virus type strain	AF285169
Wheat streak mosaic virus Hoym	HG810954.1
Wheat mosaic virus KS7 RNA1	KT988860.1
Wheat mosaic virus KS7 RNA2	KT988861.1
Wheat mosaic virus KS7 RNA3A	KT988862.1
Wheat mosaic virus KS7 RNA3B	KT988863.1

Name of virus	Accession number
Wheat mosaic virus KS7 RNA4	KT988864.1
Wheat mosaic virus KS7 RNA5	KT988865.1
Wheat mosaic virus KS7 RNA6	KT988866.1
Wheat mosaic virus KS7 RNA7	KT988867.1
Wheat mosaic virus KS7 RNA8	KT988868.1
Wheat dwarf virus	KJ473705.1
Wheat eqlid mosaic virus	NC_009805.1
Wheat rosette stunt virus	AF059602.1
Wheat spindle streak mosaic virus RNA1	NC_040508.1
Wheat spindle streak mosaic virus RNA2	NC_040507.1
Wheat stripe virus RNA2	AY312434.1
Wheat stripe virus RNA3	AY312435.1
Wheat stripe virus RNA4	AY312436.1
Wheat yellow mosaic virus RNA1	AB910332.1
Wheat yellow mosaic virus RNA2	AB910336.1
Foxtail mosaic virus	EF630359.1
Triticum mosaic virus KS	FJ263671.1
Barley virus G	KT962089.1
Barley yellow mosaic virus RNA1	AJ132268.1
Barley yellow mosaic virus RNA2	AJ132269.1
Brome mosaic virus RNA1	NC_002026.1
Brome mosaic virus RNA2	NC_002027.1
Brome mosaic virus RNA3	NC_002028.2
European wheat striate mosaic virus RNA1	MN044342.1
European wheat striate mosaic virus RNA2	MN044343.1
European wheat striate mosaic virus RNA3	MN044344.1
European wheat striate mosaic virus RNA4	MN044345.1
Chinese wheat mosaic virus RNA1	NC_002359.1
Chinese wheat mosaic virus RNA2	NC_002356.1
Johnsongrass mosaic virus	KX897165.1
Maize chlorotic mottle virus	X14736.2
Maize yellow dwarf virus-RMV	KC921392.1
Maize yellow mosaic virus-Morogoro	MW036244.1
Maize yellow striate virus	KY884303.1
Oat dwarf virus	KX533459.1
Panicum mosaic virus	MH885652.1
Ryegrass mosaic virus	MT005828.1
Sitobion miscanthi flavi-like virus	MH778148.1
Soil borne cereal mosaic virus RNA1	NC_002042.1
Soil borne cereal mosaic virus RNA2	NC_002041.1
Sugarcane mosaic virus	AJ297628.1
Wheat leaf yellowing-assocaited virus	KY605226.1
Wheat yellow dwarf virus-GPV	NC_012931.1
Wheat yellow striate virus	MG604920.1
Brome streak mosaic rymovirus	Z48506.1

Supplementary Table A.2. List of 2019 survey samples chosen for Nanopore sequencing

Sample ID	County	Number of total reads
19CN1	Cheyenne	771585
19CN3	Cheyenne	244288
19CY4	Clay	358208
19DC1	Decatur	1016763
19EW2	Ellsworth	299673
19FI	Finney	321968
19FI2	Finney	152662
19FO	Ford	314710
19GH1	Graham	307261
19GH2	Graham	130419
19GL1	Greeley	785329
19GT	Grant	249860
19HM1	Hamilton	304740
19JW1	Jewell	282607
19KE1	Kearny	137360
19KM	Kingman	97922
19KW	Kiowa	378653
19LG2	Logan	682766
19MC1	Mitchell	792618
19ME	Meade	359888
19MN	Marion	145666
19MP	McPherson	284984
19MT	Morton	295365
19NS2	Ness	221786
19NS3	Ness	270071
19OB1	Osborne	221669
19PL1	Phillips	136534
19PL3	Phillips	87385
19PN1	Pawnee	609762
19PN2	Pawnee	130358
19PN3	Pawnee	115807
19RA3	Rawlins	565130
19RH1	Rush	344955
19RN	Reno	119307
19RO1	Rooks	145552
19RP1	Republic	608705
19RS2	Russell	269145
19SA1	Saline	482021
19SH3	Sherman	300490
19SM4	Smith	59895
19ST	Stanton	317525
19SV	Stevens	557096
19SW	Seward	401427
19TH2	Thomas	52369

19TR1	Trego	654870

Supplementary Table A.3. List of 2020 and 2021 survey samples chosen for Nanopore sequencing

Sample ID	County	Year collected	Number of total reads
20 LE1	Lane	2020	90608
20LE10	Lane	2020	35844
20GH2	Graham	2020	1310636
20NT1	Norton	2020	138350
20SD4	Sheridan	2020	1294849
20GL2	Greeley	2020	106591
20TR2	Trego	2020	650465
20GO	Gove	2020	1796859
20NS5	Ness	2020	743623
20SC2	Scott	2020	762741
20WA2	Wallace	2020	419192
20WH	Wichita	2020	60378
20JW3	Jewell	2020	324793
20JW4	Jewell	2020	1963095
20MC2	Mitchell	2020	327715
20OB2	Osborne	2020	315507
20PL2	Phillips	2020	161659
20RP3	Republic	2020	361837
20KE2	Kearny	2020	129680
20KM	Kingman	2020	367816
20EW	Ellsworth	2020	174521
20SM3	Smith	2020	1392535
20SM4	Smith	2020	79933
20RH2	Rush	2020	171903
20LE17	Lane	2020	184671
21RL1	Riley	2021	580019
21RL4	Riley	2021	1348843
21RL5	Riley	2021	537729
21WH1	Wichita	2021	896519
21WH2	Wichita	2021	539305
21WH3	Wichita	2021	555369
21WH4	Wichita	2021	318681
21WH5	Wichita	2021	342432
21WH6	Wichita	2021	282203
21LE3	Lane	2021	524572
21BT1	Barton	2021	1909520
21BT2	Barton	2021	484472
21BT3	Barton	2021	101556

Supplementary Table A.4. List of complete viral genome sequences and characterization of the consensus sequences of wheat streak mosaic virus that identified on wheat samples using Nanopore sequencing.

Sample ID	County	No of reads†	Coverage (X)	Nucleotide identity (%)‡
19CN1	Cheyenne	8232	334.28	98.0
19DC1	Decatur	32941	1141.6	97.2
19CN3	Cheyenne	4504	234.8	97.8
19SH3	Sherman	15240	2820.36	97.3
19SV	Stevens	3929	141.3	90.4
19ST	Stanton	3590	179.39	97.0
19NS2	Ness	12633	826.23	97.5
19RA3	Rawlins	23855	4180.06	97.8
19GH1	Graham	939	113.27	97.9
19SW	Seward	4494	241.62	97.6
19TR1	Trego	8708	383.8	98.1
19MC1	Mitchell	6995	207.72	97.3
19FI	Finney	5133	263.39	98.1
19RH1	Rush	4483	327.94	88.4
19HM1	Hamilton	8637	576.3	97.9
19ME	Meade	4268	199.45	97.8
19MT	Morton	28509	1752.32	98.4
20NS5	Ness	7971	405.12	97.8
20PL2	Phillips	1787	126.21	98.2
20SD4	Sheridan	31250	5432.71	97.6
20GL2	Greeley	16781	1449.28	98.2
20GO	Gove	8357	307.58	94.9
20TR2	Trego	3833	182.9	97.6
20GH2	Graham	32426	7525.89	98.1
20WA2	Wallace	7400	258.44	96.3
20MC2	Mitchell	2608	163.53	97.3
20JW3	Jewell	5501	333.36	98.1
20EW	Ellsworth	1033	93.7	96.9
20SM3	Smith	1454	86.21	97.6
20KE2	Kearny	2459	158.48	98.2
20LE17	Lane	10029	681.0	98.2
20RH2	Rush	4030	227.2	98.0
21WH6	Wichita	11896	834.85	97.1
21WH7	Wichita	7012	167.87	97.6
21WH3	Wichita	3345	162.04	98.2
21RL4	Riley	6224	430.08	97.4
21RL1	Riley	17683	3746.95	97.3

† Number of reads obtained using nanopore sequencing and mapped with reference genome using CLC Genomics Workbench

‡Percent nucleotide identity of Kansas isolates sequence of this study to the type species (WSMV type isolate. AF285169.1) sequence

Supplementary Table A.5. List of complete viral genome sequences and characterization of the consensus sequence of Triticum mosaic virus identified on wheat samples using Nanopore sequencing.

Sample ID	County	No of reads†	Coverage (X)	Nucleotide identity (%)‡
19GT	Grant	3228	109.63	99.7
19SW	Seward	6891	316.79	99.7
19HM1	Hamilton	7776	324.93	99.7
19MT	Morton	3152	160.25	99.7
20GL2	Greeley	6851	542.47	99.7
21WH1	Wichita	3750	129.6	99.6
21WH2	Wichita	14040	892.22	99.7
21WH3	Wichita	3041	110.26	99.7
21LE3	Lane	4939	213.8	99.7
21WH4	Wichita	3881	126.13	99.7
21WH6	Wichita	9071	391.06	99.6

† Number of reads obtained using nanopore sequencing and mapped with reference genome using CLC Genomics Workbench

‡Percent nucleotide identity of Kansas isolates sequence of this study to the reference genome TriMV KS isolate, FJ263671.1) sequence

Supplementary Table A.6. List of sequences of viruses retrieved from GenBank

Sample ID	virus	Origin	Accession Number
Argentina	Wheat streak mosaic virus	Argentina	FJ348359.1
Austria	Wheat streak mosaic virus	Austria	LN624217.1
COKCar	Wheat streak mosaic virus	CO, USA	MT762110.1
Czech	Wheat streak mosaic virus	Czech	AF454454.1
El Batan3	Wheat streak mosaic virus	Mexico	AF285170.1
H95S	Wheat streak mosaic virus	KS, USA	AF5116114.2
H98	Wheat streak mosaic virus	KS, USA	AF511615.2
Germany_Hoym	Wheat streak mosaic virus	Germany	HG810954.1
ID96	Wheat streak mosaic virus	ID, USA	AF511618.2
ID99	Wheat streak mosaic virus	ID, USA	AF511619.2
KSHm1	Wheat streak mosaic virus	KS, USA	MK318276.1
KSWal2017	Wheat streak mosaic virus	KS, USA	MK318281.1
France_Marmagne	Wheat streak mosaic virus	France	HG810953.1
MON96	Wheat streak mosaic virus	MT, USA	AF511630.2
Naghadeh Iran	Wheat streak mosaic virus	Iran	EU914917.1
ONMV*	Oat necrotic mottle virus		NC005136.1
Sidney81	Wheat streak mosaic virus	NE, USA	AF057533.1
Sosn	Wheat streak mosaic virus	Poland	MH939146.1
Turkey1	Wheat streak mosaic virus	Turkey	AF454455.1
WA94	Wheat streak mosaic virus	WA, USA	FJ348358.1
WA99	Wheat streak mosaic virus	WA, USA	AF511643.2
WSMV_OH1	Wheat streak mosaic virus	OH, USA	MK975887.1
WSMV_TYPE	Wheat streak mosaic virus	KS, USA	AF285169.1
KSGre2017	Triticum mosaic virus	KS, USA	MK318272.1
COKCar	Triticum mosaic virus	CO, USA	MT762125.1
KSHm_2015	Triticum mosaic virus	KS, USA	MK318273.1
KSIct2017	Triticum mosaic virus	KS, USA	MK318274.1
NE	Triticum mosaic virus	NE, USA	FJ669487.1
U06-123	Triticum mosaic virus	KS, USA	FJ263671.1
YN-YZ211*	Sugarcane streak mosaic virus		KJ187047.1
CalVA KP1*	Caladenia virus A		JX156425.1

*These viruses were used as outgroups for the phylogenetic analysis

Supplementary Table A.7. List of complete viral genome sequences and characterization of the consensus sequences of High Plains wheat mosaic emaravirus identified on wheat samples using Nanopore sequencing

Sample ID	county	Genome	No of reads†	Coverage (X)‡	Nucleotide identity (%)‡
20SC2	Scott	RNA1	1181	67.4	99.4
20SC2	Scott	RNA2	1291	159.7	99.4
20SC2	Scott	RNA3A	2982	912.5	99.6
20SC2	Scott	RNA3B	67011	18320.7	98.3
20SC2	Scott	RNA4	6970	650.0	99.1
20SC2	Scott	RNA5	198	43.3	99.6
20SC2	Scott	RNA6	869	194.8	99.4
20SC2	Scott	RNA7	2232	757.0	99.1
20SC2	Scott	RNA8	1222	37.7	98.8
20MC2	Mitchell	RNA3A	370	16.1	96.3
20MC2	Mitchell	RNA3B	972	250.6	99.3
20MC2	Mitchell	RNA4	494	12.87	97.7
20MC2	Mitchell	RNA7	210	18.0	96.6
20KE2	Kearny	RNA3A	350	97.0	98.1
20KE2	Kearny	RNA3B	8470	1995.0	99.4
20KE2	Kearny	RNA4	1716	650.0	98.6
20KE2	Kearny	RNA6	121	16.0	99.1
20KE2	Kearny	RNA7	369	88.6	98.3
20KE2	Kearny	RNA8	280	10.0	98.8
21LE3	Lane	RNA3A	634	157.4	98.7
20RH2	Rush*	RNA4	1289	233.4	-
20RH2	Rush	RNA7	655	90.6	85.5

† Number of reads obtained using nanopore sequencing and mapped with reference genome using CLC Genomics Workbench

‡Percent nucleotide identity of Kansas isolates sequence of this study to the reference genomes sequence

*Missing 115 bps at 5’ end and excluded from nucleotide percentage identity analysis

Supplementary Table A.8. List of sequences of High Plains wheat mosaic emaravirus (HPWMoV) isolates retired from GenBank

Sample ID	RNA	virus	Origin	Accession Number
W1	RNA3	HPWMoV	OH, USA	KT970501.1
Cophil	RNA3	HPWMoV	CO, USA	MT762120.1
K1	RNA3	HPWMoV	OH, USA	KT98889.1
H1_RNA3	RNA3	HPWMoV	OH, USA	KT98881.1
NW2P3	RNA3	HPWMoV	OH, USA	MN250347.1
NE_RNA3A*	RNA3A	HPWMoV	NE, USA	KJ939625.1
KS7_3A	RNA3A	HPWMoV	KS, USA	KT988862.1
NW1_3A	RNA3A	HPWMoV	OH, USA	MN250339.1
NE_RNA3B*	RNA3B	HPWMoV	NE, USA	KJ939626.1
KS7_RNA3B	RNA3B	HPWMoV	KS, USA	KT988863.1
GG1_RNA3B	RNA3B	HPWMoV	OH, USA	KT988872.1
NE_RNA1*	RNA1	HPWMoV	NE, USA	NC_029570.1
NE_RNA2*	RNA2	HPWMoV	NE, USA	NC_029549.1
NE_RNA4*	RNA4	HPWMoV	NE, USA	NC_029551.1
NE_RNA5*	RNA5	HPWMoV	NE, USA	NC_029552.1
NE_RNA6*	RNA6	HPWMoV	NE, USA	NC_029553.1
NE_RNA7*	RNA7	HPWMoV	NE, USA	NC_029554.1
NE_RNA8*	RNA8	HPWMoV	NE, USA	NC_029555.1
RLBV**	RNA3	Raspberry leaf blotch virus	-	FR823301.1

*These HPWMoV isolates used as viral reference genomes used for mapping
**These viruses were used as outgroups for the phylogenetic analysis

Supplementary Table A.9. List of complete viral genome sequences and characterization of the consensus sequence of Soilborne wheat mosaic virus identified on wheat samples using Nanopore sequencing

Sample ID	county	Accession numbers	Genome	No of reads†	Coverage (X)	Nucleotide identity (%)‡
19PN2	Pawnee		RNA1	12818	816.55	98.47
19PN2	Pawnee		RNA2	12723	596.12	98.52
21RL5	Riley		RNA1	300	124.94	96.63
21RL5	Riley		RNA2	1982	105.2	99.97

† Number of reads obtained using nanopore sequencing and mapped with reference genome using CLC Genomics Workbench

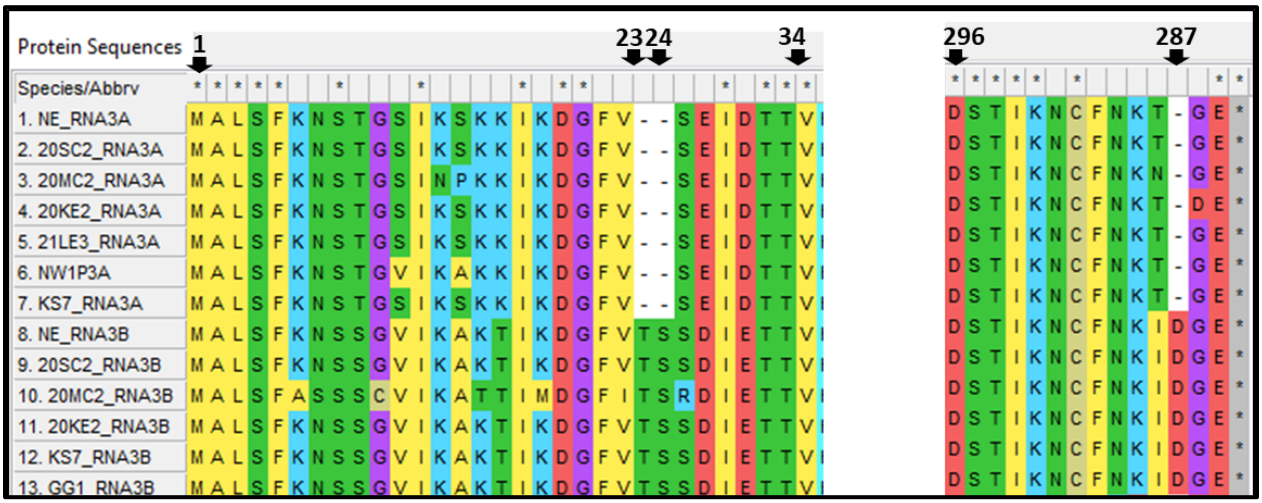
‡The reference genome used to compare the nucleotide identity were (RNA1: NC_002041.1 and RNA2: NC_002042.1)

Supplementary Table A.10. Potential recombinant isolates of wheat streak mosaic virus (WSMV) analyzed by using 7 different algorithms in the RDP5 program

Recombinants	Major parent	Minor parent	RDP*	BootScan*	GENECONV*	MaxChi*	Chimera*	3seq*	SiScan*
19SV	Germany Hoym	WSMV type	1.51×10^{-36}	5.42×10^{-26}	1.85×10^{-24}	1.02×10^{-11}	3.25×10^{-12}	6.98×10^{-07}	4.47×10^{-13}
20GO	H95S	19RH1	4.37×10^{-48}	2.89×10^{-41}	3.6×10^{-39}	1.19×10^{-20}	3.57×10^{-21}	2.84×10^{-37}	2.61×10^{-30}
20WA	20LE1	19RH1	6.16×10^{-63}	9.33×10^{-17}	4.92×10^{-12}	8.56×10^{-06}	4.33×10^{-06}	1.48×10^{-09}	2.12×10^{-03}
19ST	19RH1	Sydney81	1.21×10^{-95}	-	1.38×10^{-04}	6.62×10^{-10}	2.28×10^{-06}	-	2.46×10^{-46}
19RA3	19SH3	19GH1	5.99×10^{-65}	1.46×10^{-03}	-	1.69×10^{-07}	1.06×10^{-05}	5.72×10^{-05}	3.43×10^{-11}
20EW	Sydney81	19SH3	3.13×10^{-22}	1.24×10^{-08}	2.16×10^{-07}	1.72×10^{-10}	2.04×10^{-08}	8.43×10^{-09}	6.52×10^{-11}
20TR2	19MT	19GH2	1.27×10^{-11}	3.42×10^{-04}	2.35×10^{-03}	4.28×10^{-06}	1.26×10^{-05}	2.02×10^{-03}	3.8×10^{-14}
20GH2	21RL1	21WH3	9.28×10^{-04}	2.79×10^{-03}	1.46×10^{-03}	1.93×10^{-03}	2.45×10^{-03}	3.29×10^{-04}	1.5×10^{-06}
20JW3	19DC1	20MC2	-	4.11×10^{-03}	-	1.6×10^{-03}	1.01×10^{-03}	9.61×10^{-04}	6.41×10^{-05}
19TR1	19CN1	20MC2	3.09×10^{-04}	-	-	2.22×10^{-03}	1.42×10^{-03}	3.23×10^{-07}	1.685×10^{-03}
19FI	19GH1	19SH3	4.07×10^{-03}			1.82×10^{-03}	1.42×10^{-03}	3.23×10^{-07}	1.685×10^{-03}
19NS2	20SD4	MON96	1.28×10^{-08}	4.19×10^{-02}	-	3.37×10^{-05}	1.48×10^{-07}	1.8×10^{-12}	-
19GH1	21RL1	21WH3	5.97×10^{-06}	7.48×10^{-05}	2.29×10^{-03}	-	-	2.75×10^{-03}	1.32×10^{-03}
19SW	20MC2	19DC1		2.27×10^{-03}		5.44×10^{-06}	6.08×10^{-06}	2.62×10^{-06}	5.99×10^{-03}

**P*-values

Supplementary Figure A.1. Alignment of nucleocapsid protein sequence encoded by High Plains wheat mosaic emaravirus RNA3 and its two variant RNA3A and 3B isolates obtained from this study as well as isolates sequence retrieved from GenBank. RNA3B showed with a 3 amino-acid insertion at the positions of 23, 24, and the 287 C terminus of the protein.



Appendix B - Supplementary Figures and Tables from Chapter 3

Appendix B Supplementary Figures and Tables

Supplementary Table B.1. List of sequences of brome mosaic virus retrieved from GenBank.

Sample ID	RNA genome	Origin	Accession Number
BMV_OH2	RNA1	OH, USA	MN241035.1
BMV_OH2	RNA2	OH, USA	MN241036.1
BMV_OH2	RNA3	OH, USA	MN241037.1
BMV_OH	RNA1	OH, USA	MH025765.1
BMV_OH	RNA2	OH, USA	MH025766.1
BMV_OH	RNA3	OH, USA	MH025767.1
BMV_OK	RNA1	OK, USA	DQ530423.1
BMV_OK	RNA2	OK, USA	DQ530424.1
BMV_OK	RNA3	OK, USA	DQ530425.1
BMV_M1*	RNA1	WI, USA	X02380.1
BMV_ M1*	RNA2	WI, USA	X01678.1
BMV_ M1*	RNA3	WI, USA	J02042.1
BMV_M2	RNA1	WI, USA	AB183262.1
BMV_M2	RNA2	WI, USA	AB183263.1
BMV_M2	RNA3	WI, USA	AB183261.1
BMV_DSMZ_PV-0194	RNA1	UK	MW582787.1
BMV_DSMZ_PV-0194	RNA2	UK	MW582788.1
BMV_DSMZ_PV-0194	RNA3	UK	MW582789.1
BMV_Germany	RNA3	Germany	MT737803.1
BMV_Estonia	RNA1	Estonia	KU726253.1
BMV_Estonia	RNA2	Estonia	KU726254.1
BMV_Estonia	RNA3	Estonia	KU726255.1
BMV_CZ	RNA1	Czech	GU584131.1
BMV_CZ	RNA2	Czech	GU584130.1
BMV_CZ	RNA3	Czech	GU584129.1
CYBV**	RNA1	-	NC_006999.2
CYBV**	RNA2	-	NC_007000.2
CYBV**	RNA3	-	NC_007001.1
OLV**	RNA1	-	X94346.1
OLV**	RNA2	-	X94347.1
OLV**	RNA3	-	X76993.1

*These viruses were used as the reference genomes

**These viruses were used as outgroups for the phylogenetic analysis

Supplementary Table B.2. List of survey samples positive to brome mosaic virus and number of raw reads obtained from Nanopore sequencing

Sample ID	County	Year collected	Number of total reads
19CN1	Cheyenne	2019	771585
19CN3	Cheyenne	2019	244288
19CY4	Clay	2019	358208
19DC1	Decatur	2019	1016763
19GT	Grant	2019	249860
19JW1	Jewell	2019	282607
19MT	Morton	2019	295365
19NS2	Ness	2019	221786
19OB1	Osborne	2019	221669
19PL1	Phillips	2019	136534
19PN1	Pawnee	2019	609762
19PN2	Pawnee	2019	130358
19RA3	Rawlins	2019	565130
19RH1	Rush	2019	344955
19RO1	Rooks	2019	145552
19RP1	Republic	2019	608705
19RS2	Russell	2019	269145
19SV	Stevens	2019	557096
19SW	Seward	2019	401427
19TH2	Thomas	2019	52369
19TR1	Trego	2019	654870
20 LE1	Lane	2020	90608
20SD4	Sheridan	2020	1294849
20GL2	Greeley	2020	106591
20TR2	Trego	2020	650465
20GO	Gove	2020	1796859
20SC2	Scott	2020	762741
20WH	Wichita	2020	60378
20JW3	Jewell	2020	324793
20MC2	Mitchell	2020	327715
20OB2	Osborne	2020	315507
20PL2	Phillips	2020	161659
20RP3	Republic	2020	361837
20KE2	Kearny	2020	129680
20KM	Kingman	2020	367816
20SM3	Smith	2020	1392535
20SM4	Smith	2020	79933

Supplementary Table B.3. A list of complete viral genome sequences and characterization of the consensus sequences of brome mosaic virus identified on wheat samples using Nanopore sequencing

Sample ID	county	Genome	No of reads†	Coverage (%)
20SM3*	Smith	RNA3	942377	243202.96
20SM3θ	Smith	RNA2	64478	6437.2
20SM3#	Smith	RNA1	30245	3126.65
19CN1δ	Cheyenne	RNA3	236	100.72
19CN3 δ	Cheyenne	RNA3	92	40.65
19DC1 δ	Decatur	RNA3	241	113.76
19NS2 δ	Ness	RNA3	51	20.14
19JW1 δ	Jewell	RNA3	4597	1413.4
19RP1*	Republic	RNA3	522697	191560.5
19RP1θ	Republic	RNA2	7513	1274.27
19RP1#	Republic	RNA1	4731	598.79

† Number of reads obtained using nanopore sequencing and mapped with reference genome using CLC Genomics Workbench

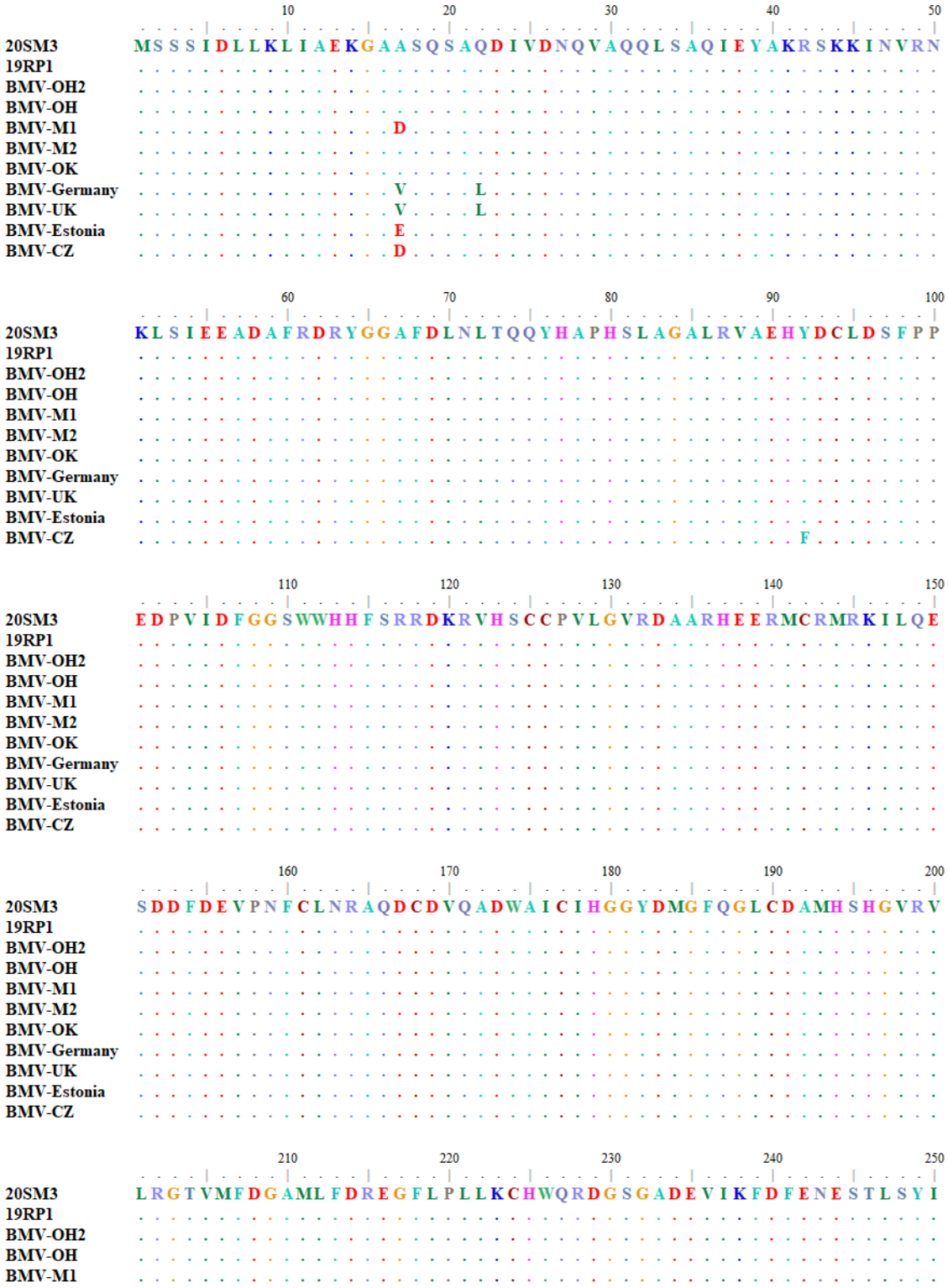
*Complete RNA3 sequence with both movement protein and coat protein-coding and non-coding regions

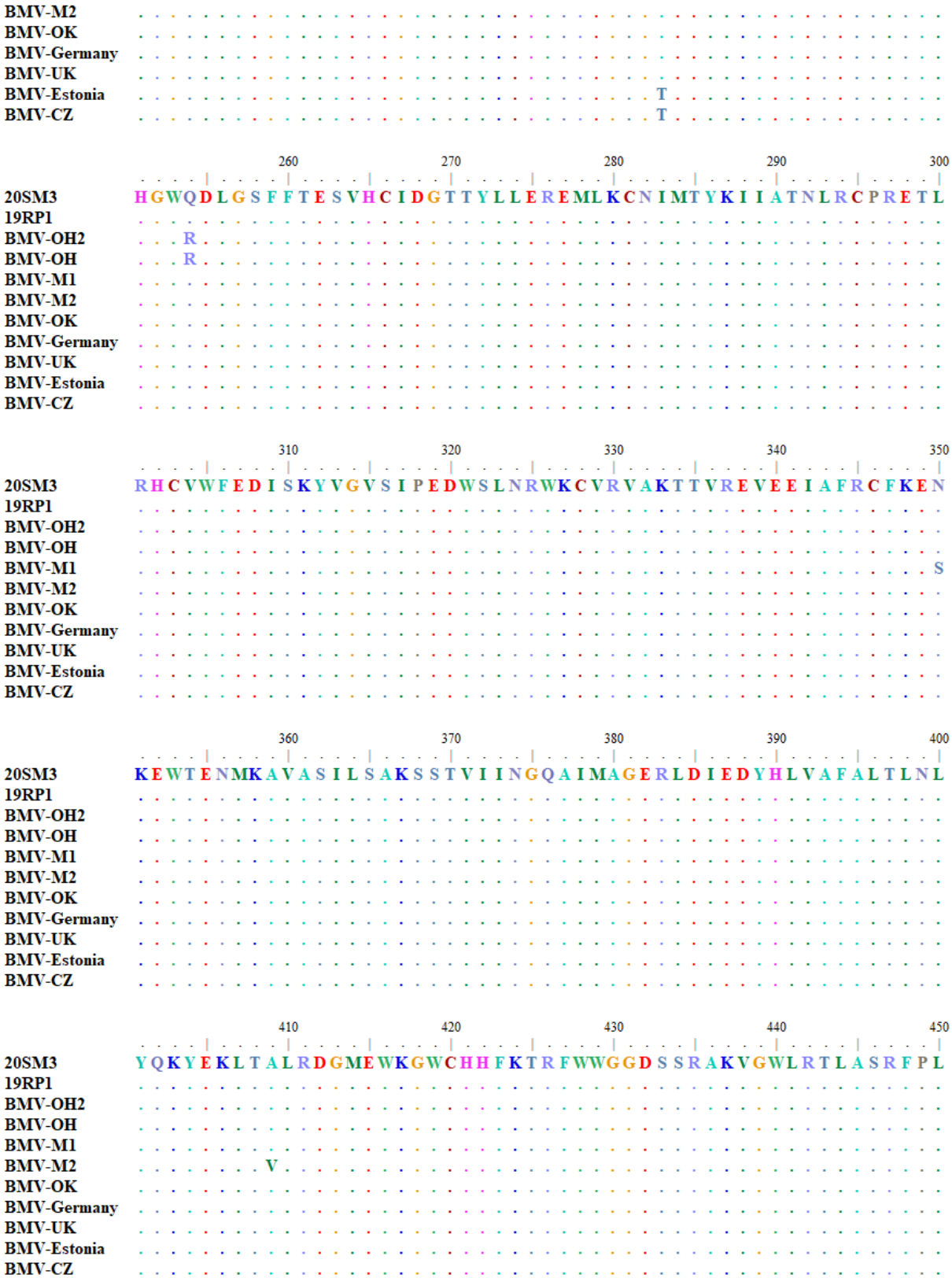
δIncomplete RNA3 sequence with only movement protein-coding region

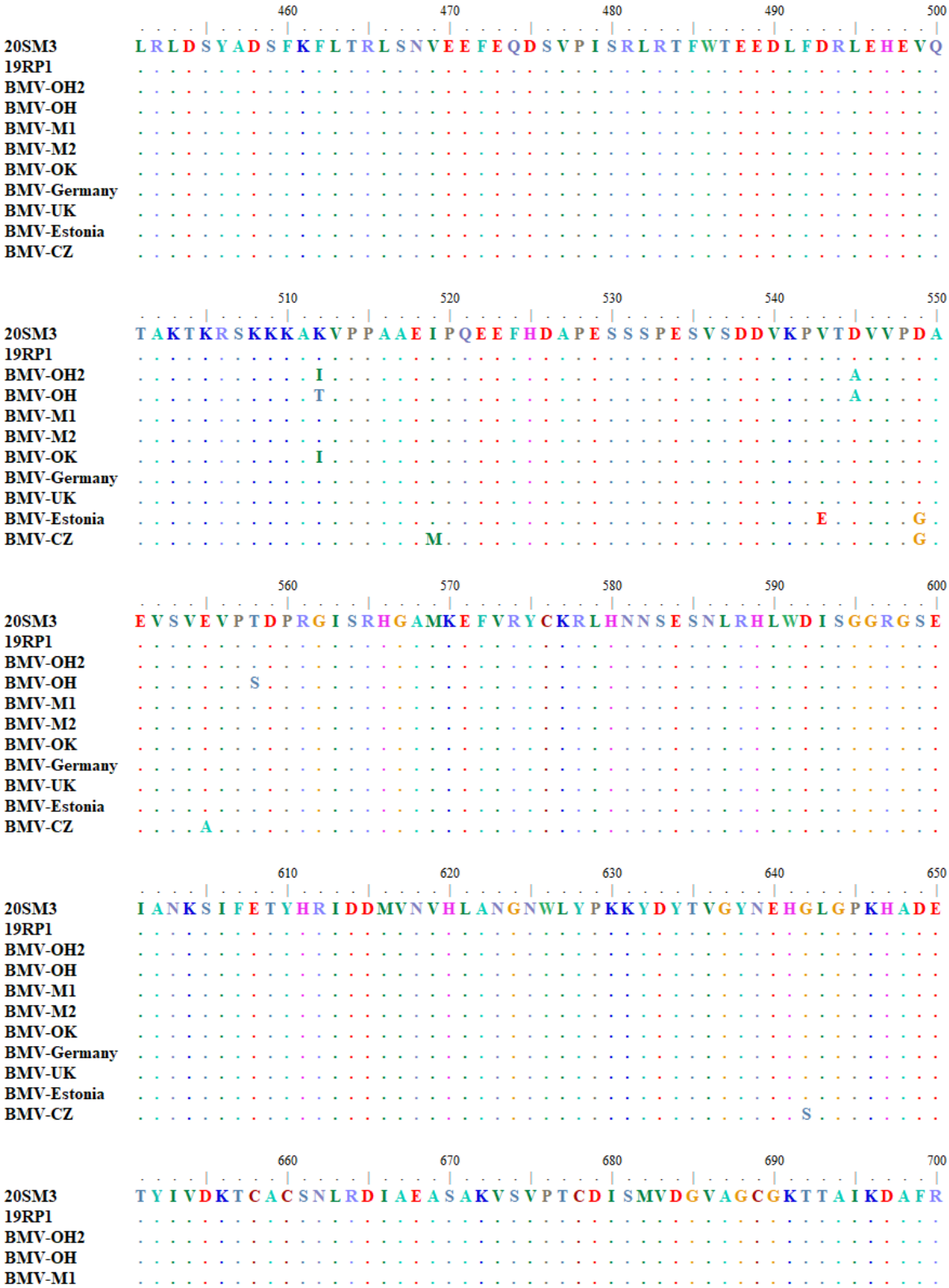
θComplete RNA2 sequence with coding and non-coding regions

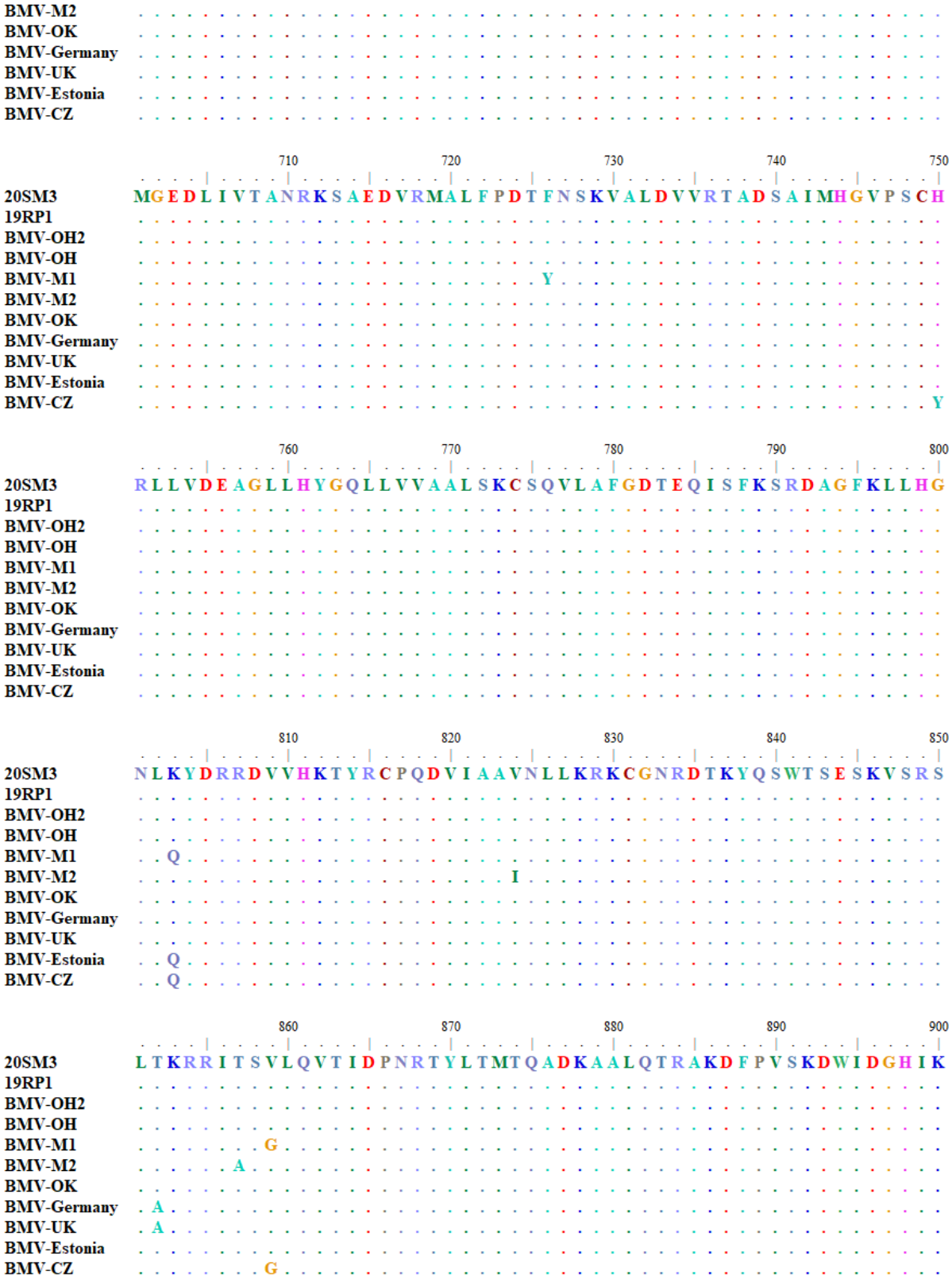
Complete RNA1 sequence with coding and non-coding regions

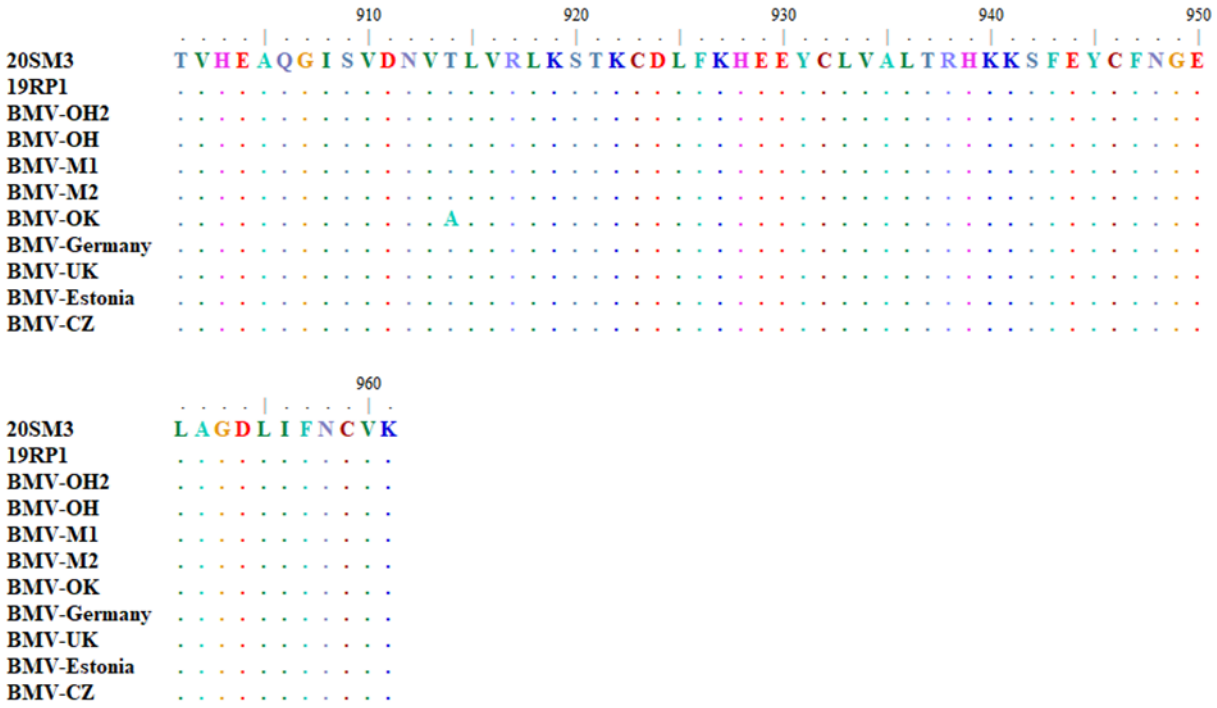
Supplementary Figure B.1. Amino acid sequence alignment diagram for RNA1a of brome mosaic virus. The alignment was obtained by using muscle alignment in Mega X. Identical amino acids in all isolates are indicated by dots. The substitution of amino acid in some isolates appears as a letter



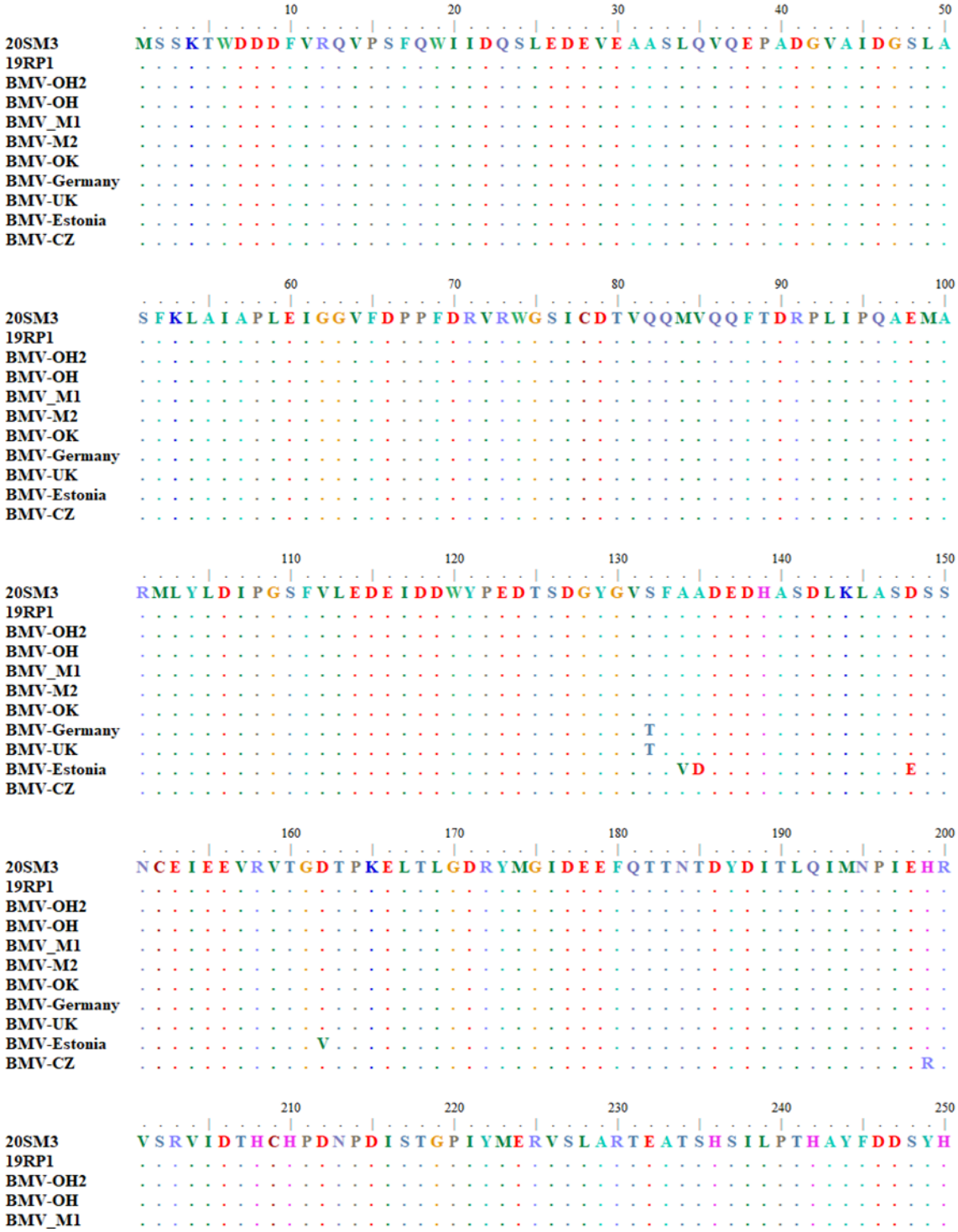

BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
T
BMV-CZ
T
260
270
280
290
300
20SM3
H G W Q D L G S F F T E S V H C I D G T T Y L L E R E M L K C N I M T Y K I I A T N L R C P R E T L
19RP1
BMV-OH2
R
BMV-OH
R
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
310
320
330
340
350
20SM3
R H C V W F E D I S K Y V G V S I P E D W S L N R W K C V R V A K T T V R E V E E I A F R C F K E N
19RP1
BMV-OH2
BMV-OH
BMV-M1
S
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
360
370
380
390
400
20SM3
K E W T E N M K A V A S I L S A K S S T V I I N G Q A I M A G E R L D I E D Y H L V A F A L T L N L
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
410
420
430
440
450
20SM3
Y Q K Y E K L T A L R D G M E W K G W C H H F K T R F W W G G D S S R A K V G W L R T L A S R F P L
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
V
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ

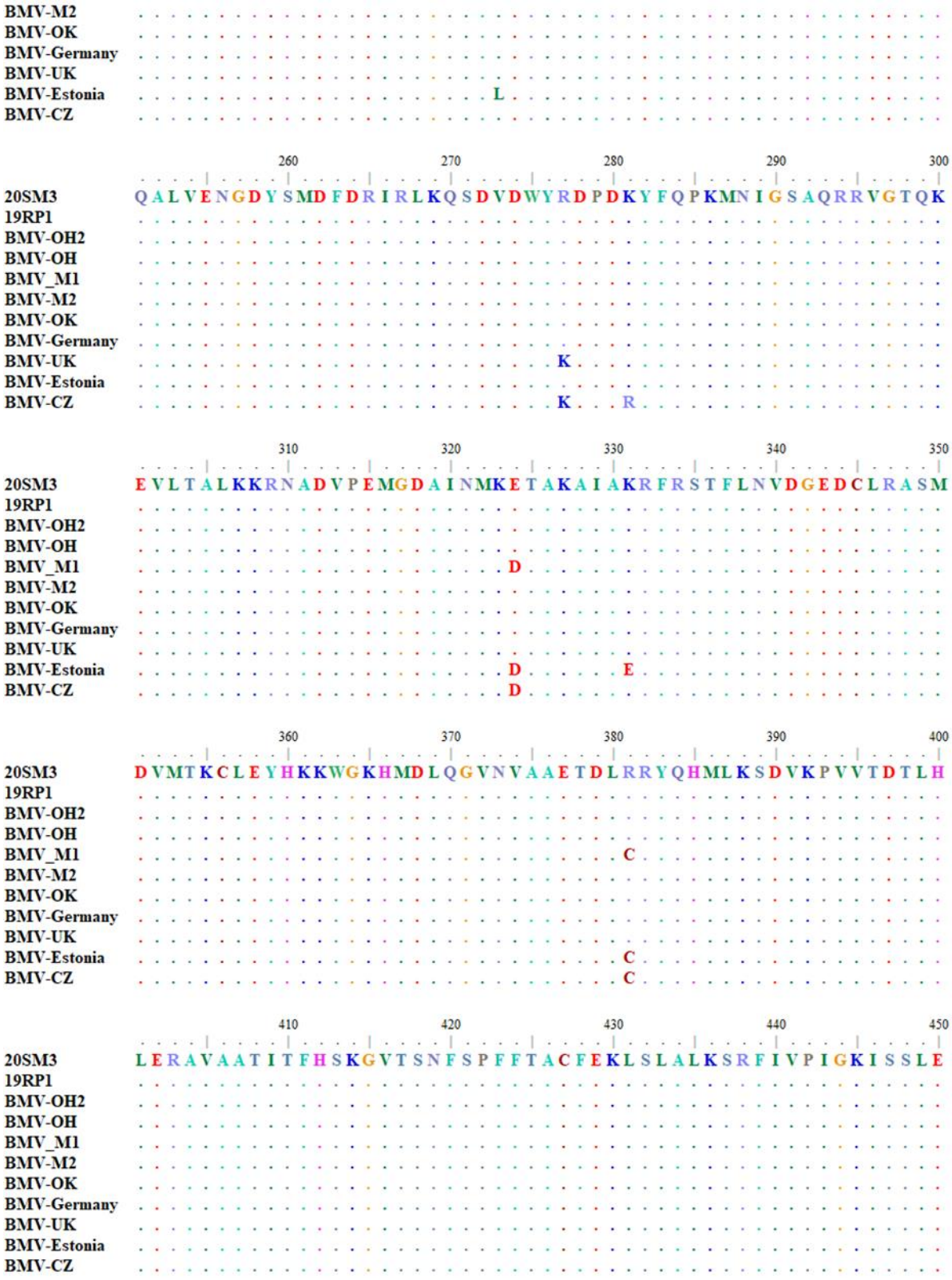

460 470 480 490 500
20SM3 LRLDSYADSFKFLTRLSNVEEFEQDSVPISRLRTFWTEEDLFDRLEHEVQ
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
510 520 530 540 550
20SM3 TAKTKRSKKKAKVPPAAEIPQEEFHDAPESSSPESVSDDVKPVTDVVPDA
19RP1
BMV-OH2 I A
BMV-OH T A
BMV-M1
BMV-M2
BMV-OK I
BMV-Germany
BMV-UK
BMV-Estonia E G
BMV-CZ M G
560 570 580 590 600
20SM3 EVSVEVPTDPRGISRHGAMKEFVRYCKRLHNNSESNLRHLWDISGGRGSE
19RP1
BMV-OH2
BMV-OH S
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ A
610 620 630 640 650
20SM3 IANKSIFETYHRIDDMVNVHLANGNWLYPKKYDYTVGYNEHGLGPKHADE
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ S
660 670 680 690 700
20SM3 TYIVDKTCACSNLRDIAEASAKVSVPTCDISMVDGVAGCGKTTAIKDAFR
19RP1
BMV-OH2
BMV-OH
BMV-M1

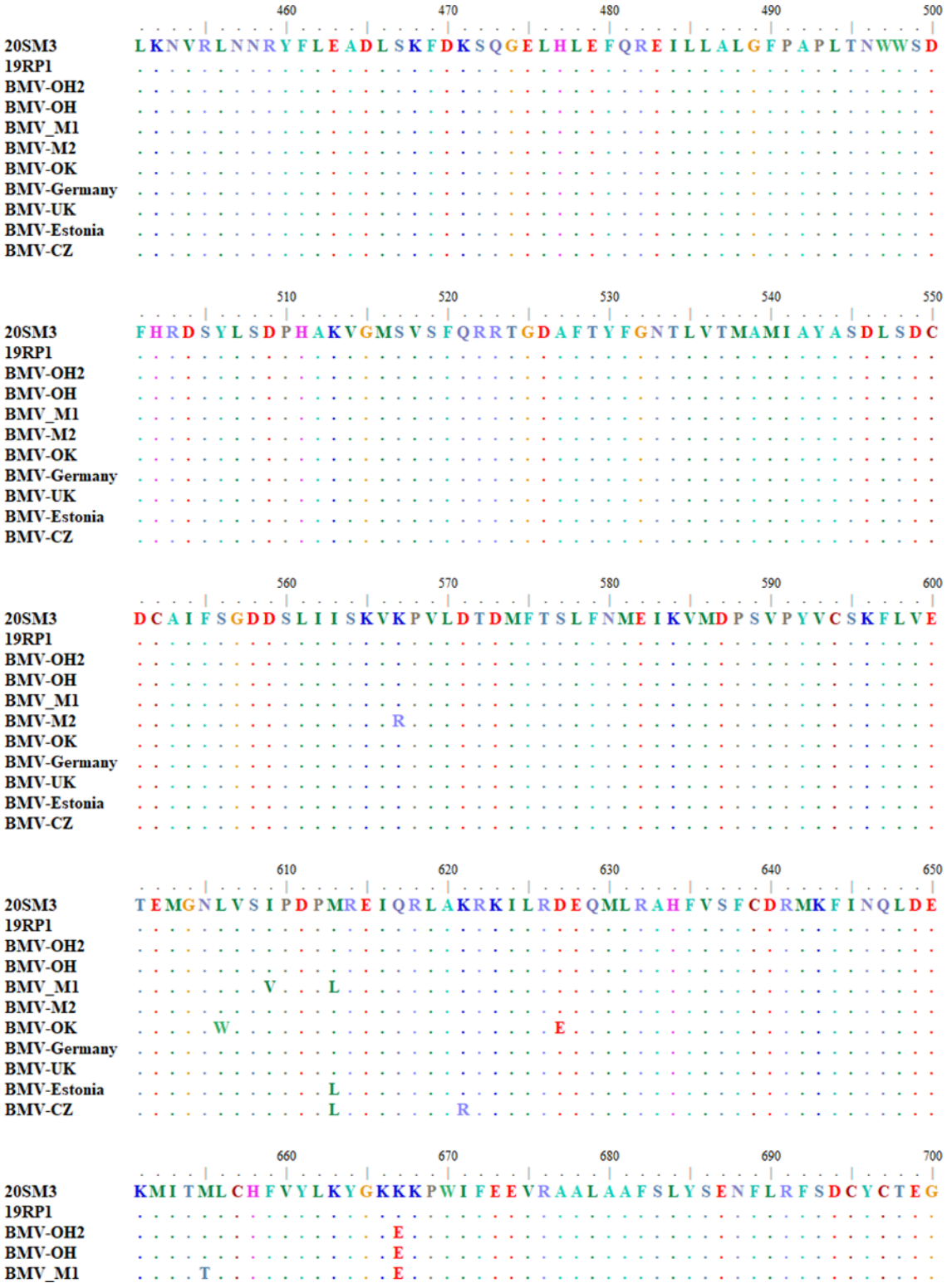

BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
710
720
730
740
750
20SM3 M G E D L I V T A N R K S A E D V R M A L F P D T F N S K V A L D V V R T A D S A I M H G V P S C H
19RP1
BMV-OH2
BMV-OH
BMV-M1 Y
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ Y
760
770
780
790
800
20SM3 R L L V D E A G L L H Y G Q L L V V A A L S K C S Q V L A F G D T E Q I S F K S R D A G F K L L H G
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
810
820
830
840
850
20SM3 N L K Y D R R D V V H K T Y R C P Q D V I A A V N L L K R K C G N R D T K Y Q S W T S E S K V S R S
19RP1
BMV-OH2
BMV-OH
BMV-M1 Q
BMV-M2 I
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia Q
BMV-CZ Q
860
870
880
890
900
20SM3 L T K R R I T S V L Q V T I D P N R T Y L T M T Q A D K A A L Q T R A K D F P V S K D W I D G H I K
19RP1
BMV-OH2
BMV-OH
BMV-M1 G
BMV-M2 A
BMV-OK
BMV-Germany A
BMV-UK A
BMV-Estonia
BMV-CZ G

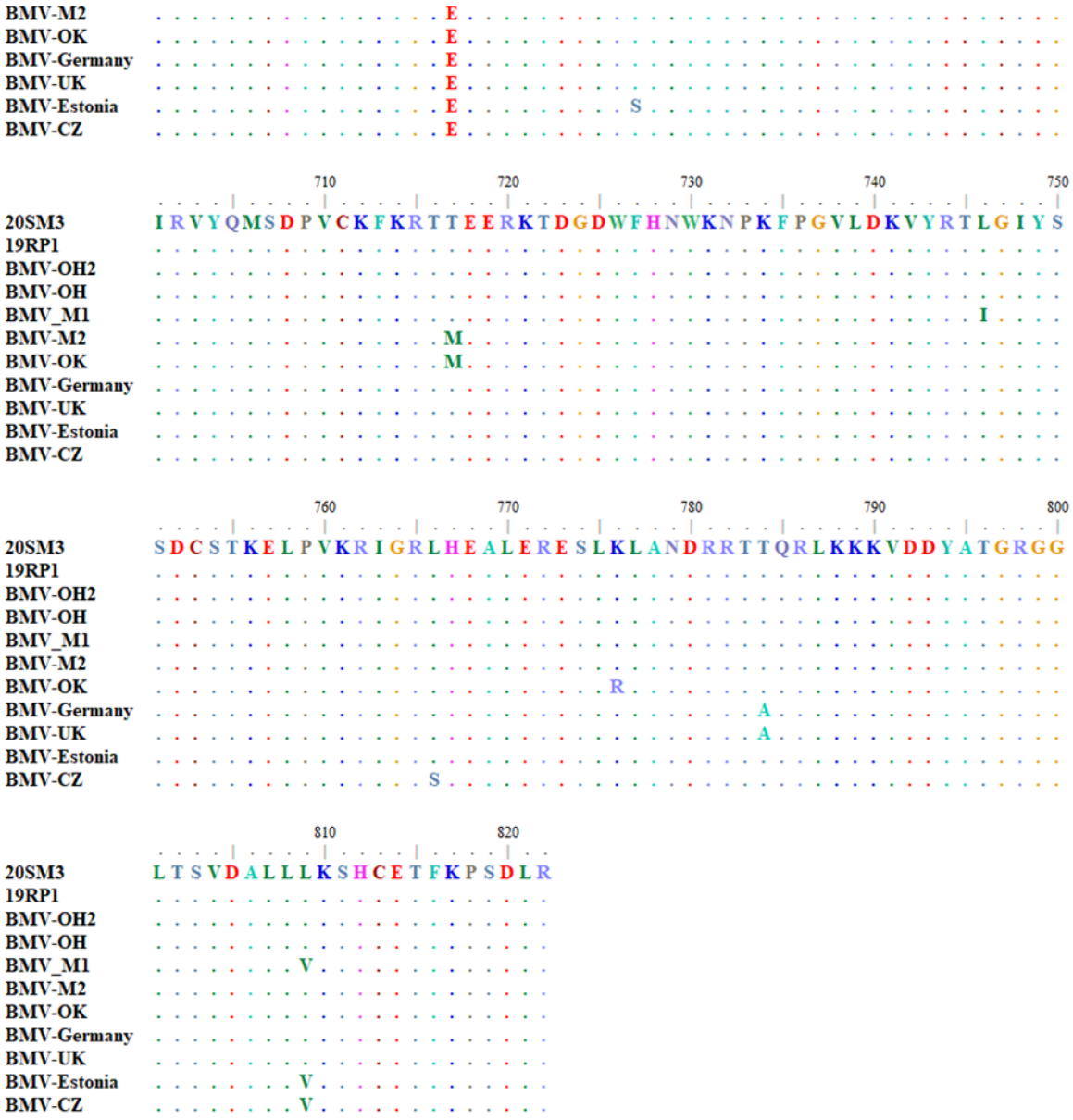

910
920
930
940
950
20SM3 T V H E A Q G I S V D N V T L V R L K S T K C D L F K H E E Y C L V A L T R H K K S F E Y C F N G E
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
BMV-OK A
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
960
20SM3 L A G D L I F N C V K
19RP1
BMV-OH2
BMV-OH
BMV-M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ

Supplementary Figure B.2. Amino acid sequence alignment diagram for RNA2b of brome mosaic virus. The alignment was obtained by using muscle alignment in MEGA X. Identical amino acids in all isolates are indicated by dots. The substitution of amino acid in some isolates appears as a letter

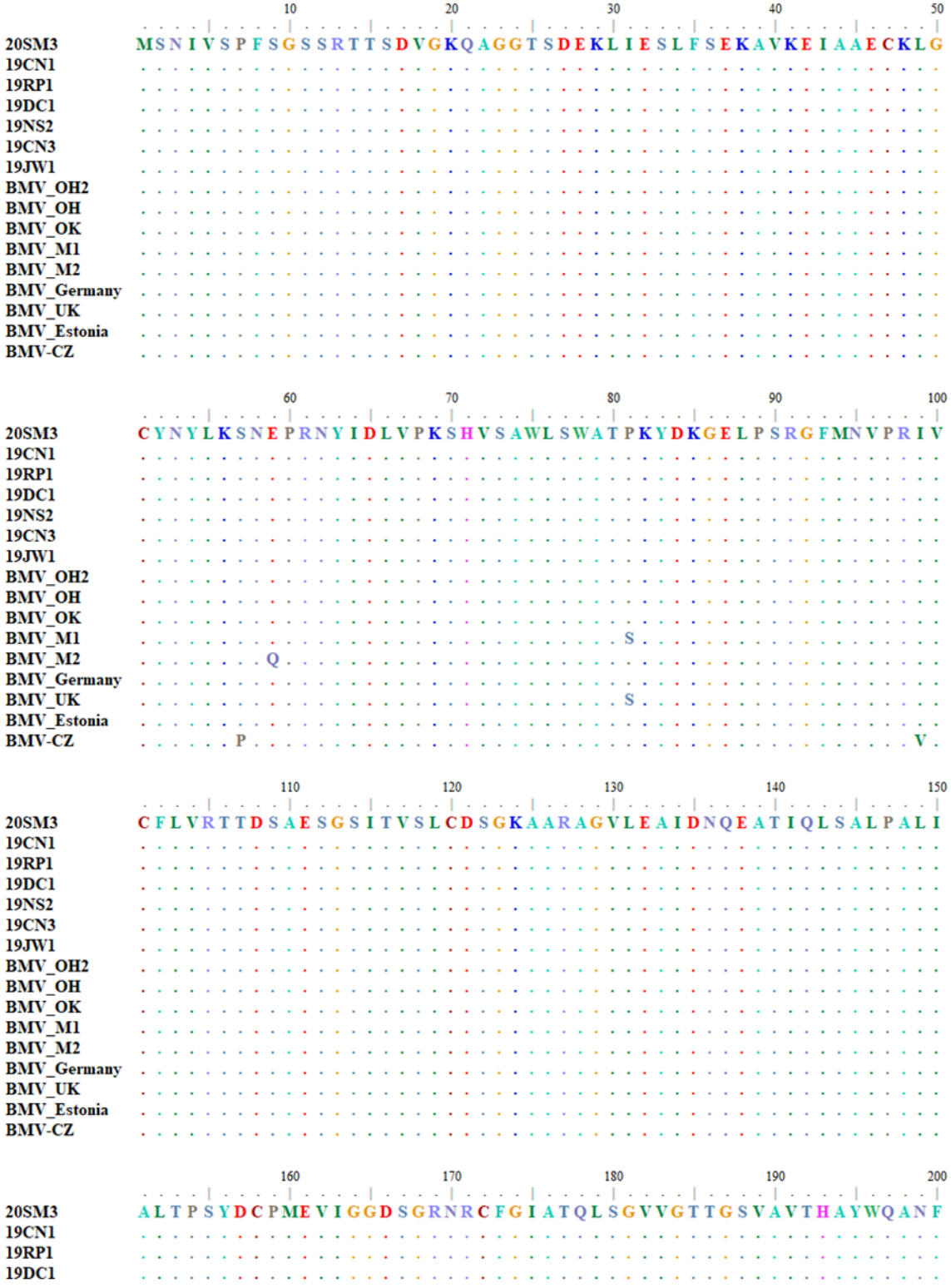


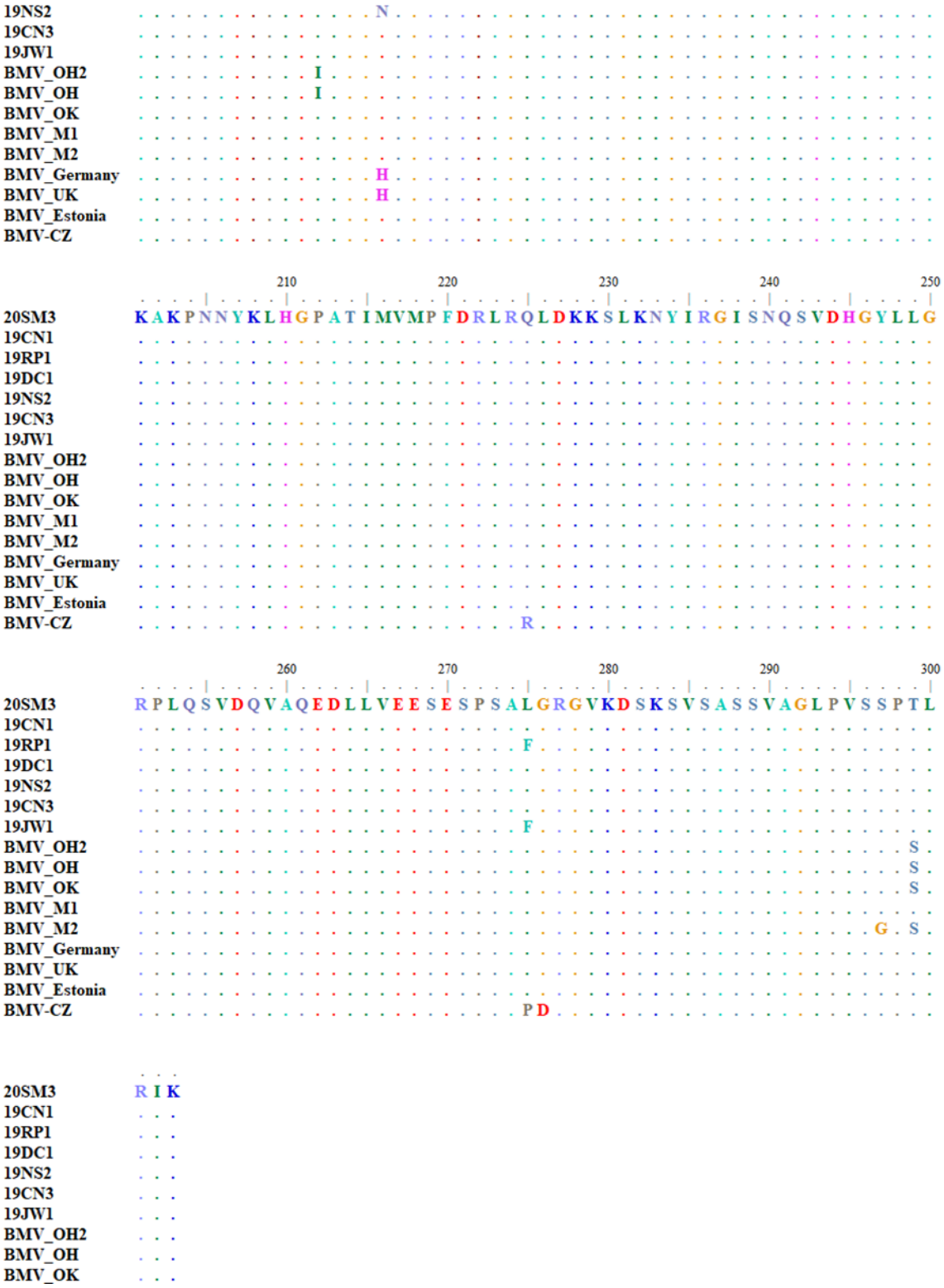

BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
260
270
280
290
300
20SM3
Q A L V E N G D Y S M D F D R I R L K Q S D V D W Y R D P D K Y F Q P K M N I G S A Q R R V G T Q K
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
310
320
330
340
350
20SM3
E V L T A L K K R N A D V P E M G D A I N M K E T A K A I A K R F R S T F L N V D G E D C L R A S M
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
360
370
380
390
400
20SM3
D V M T K C L E Y H K K W G K H M D L Q G V N V A A E T D L R R Y Q H M L K S D V K P V V T D T L H
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
410
420
430
440
450
20SM3
L E R A V A A T I T F H S K G V T S N F S P F F T A C F E K L S L A L K S R F I V P I G K I S S L E
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ


460 470 480 490 500
20SM3 LKNVRLNNRYFLEADLSKFDKSQGELHLEFQREILLALGFPAPLTNWWSD
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
510 520 530 540 550
20SM3 FHRDSYLSDPHAKVGMSVSFQRRTGDAFTYFGNTLVTMAMIAYASDLSDC
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
560 570 580 590 600
20SM3 DCAIFSGDDSLIISKVKPVLDTDMFTSLFNMEIKVMDPSVPYVCSKFLVE
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2 R
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
610 620 630 640 650
20SM3 TEMGNLVSIPDPMREIQRLAKRKILRDEQMLRAHFVSFCDRMKFINQLDE
19RP1
BMV-OH2
BMV-OH
BMV_M1 V L
BMV-M2
BMV-OK W E
BMV-Germany
BMV-UK
BMV-Estonia L
BMV-CZ L R
660 670 680 690 700
20SM3 KMITMLCHFVYLKYGKKKPWIFEEVRAALAAFSLYSENFLRFSDCYCTEG
19RP1
BMV-OH2 E
BMV-OH E
BMV_M1 T E

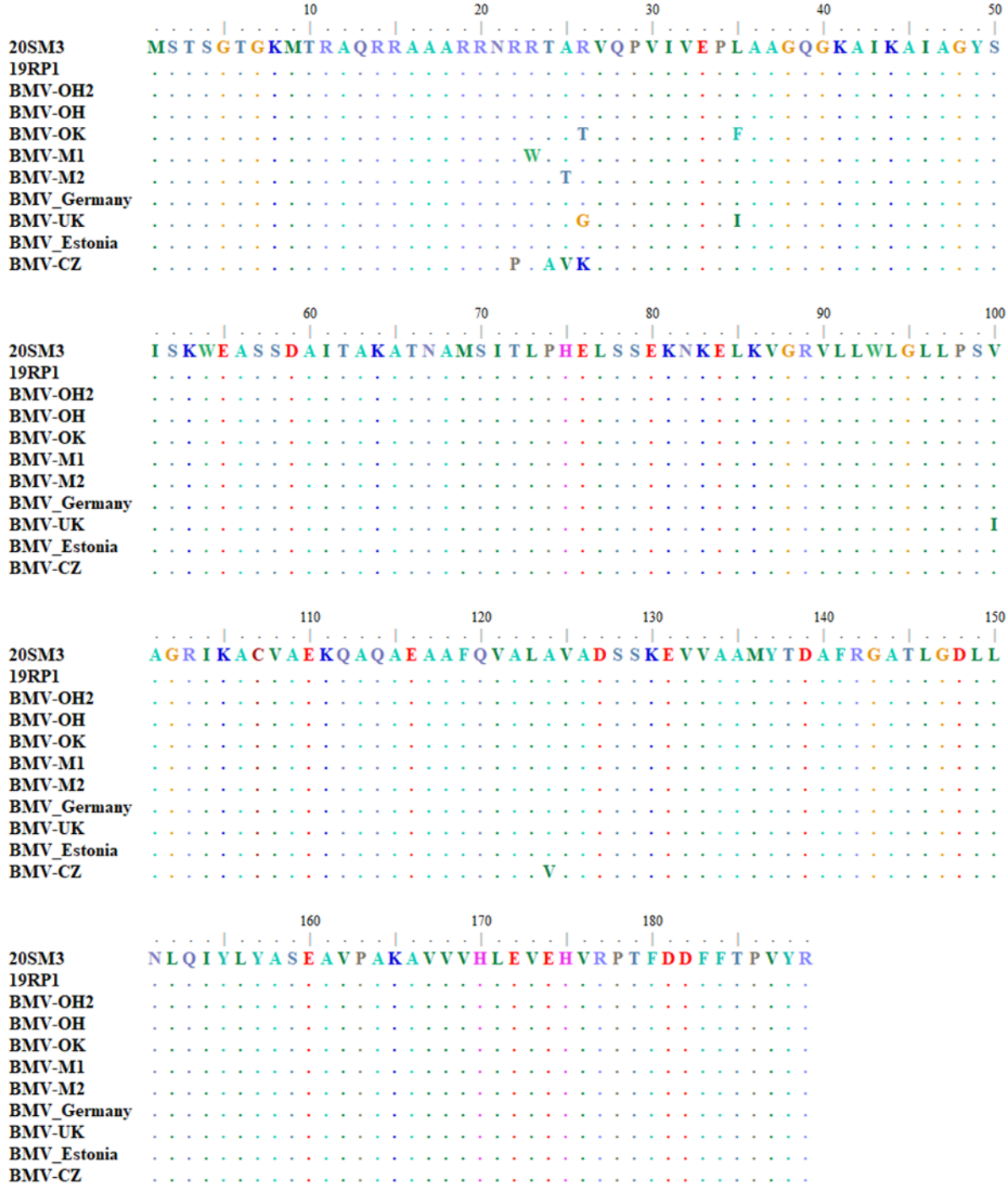

BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
710
720
730
740
750
20SM3
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
760
770
780
790
800
20SM3
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ
810
820
20SM3
19RP1
BMV-OH2
BMV-OH
BMV_M1
BMV-M2
BMV-OK
BMV-Germany
BMV-UK
BMV-Estonia
BMV-CZ

Supplementary Figure B.3. Amino acid sequence alignment diagram for RNA3a (Movement protein) of brome mosaic virus. The alignment was obtained by using muscle alignment in Mega X. Identical amino acids in all isolates are indicated by dots. The substitution of amino acid in some isolates appears as a letter




19NS2
19CN3
19JW1
BMV_OH2
BMV_OH
BMV_OK
BMV_M1
BMV_M2
BMV_Germany
BMV_UK
BMV_Estonia
BMV-CZ
210
220
230
240
250
20SM3 K A K P N N Y K L H G P A T I M V M P F D R L R Q L D K K S L K N Y I R G I S N Q S V D H G Y L L G
260
270
280
290
300
20SM3 R P L Q S V D Q V A Q E D L L V E E S E S P S A L G R G V K D S K S V S A S S V A G L P V S S P T L
20SM3 R I K

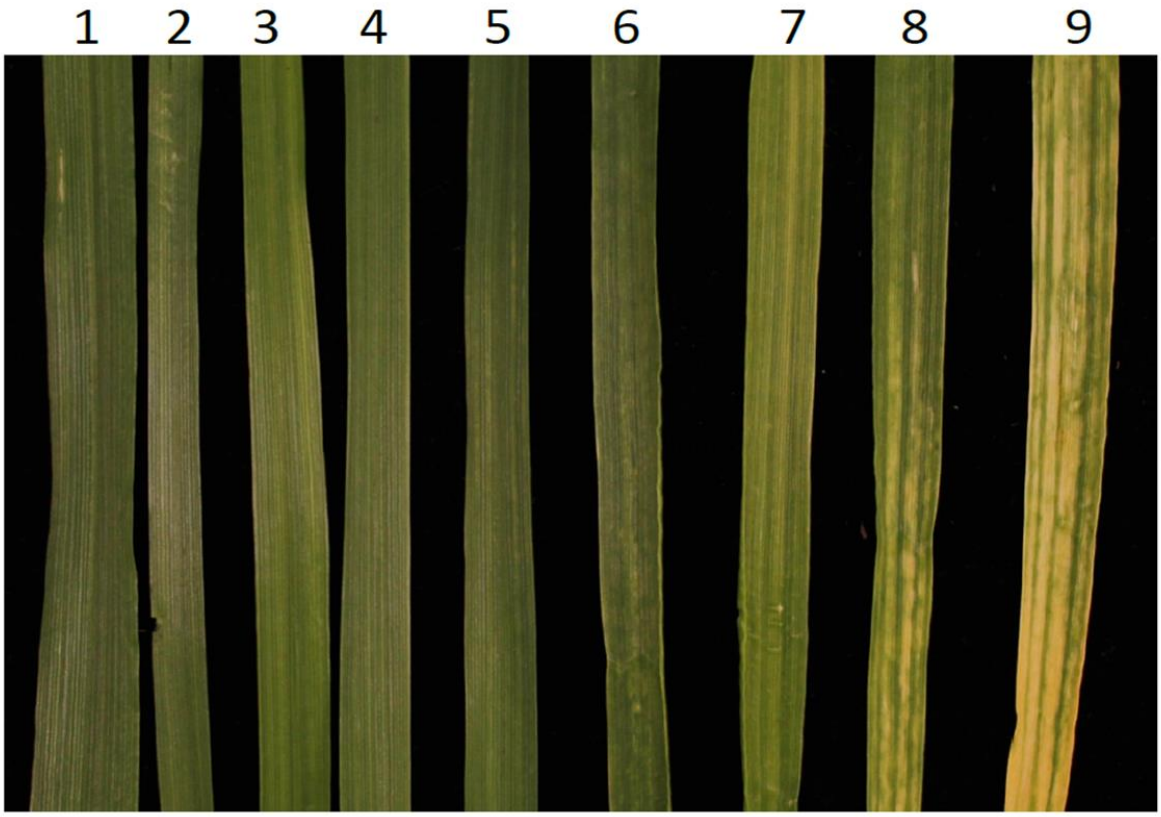
Supplementary Figure B.4. Amino acid sequence alignment diagram for RNA3b (coat protein) of brome mosaic virus. The alignment was obtained by using muscle alignment in Mega X. Identical amino acids in all isolates are indicated by dots. The substitution of amino acid in some isolates appears as a letter



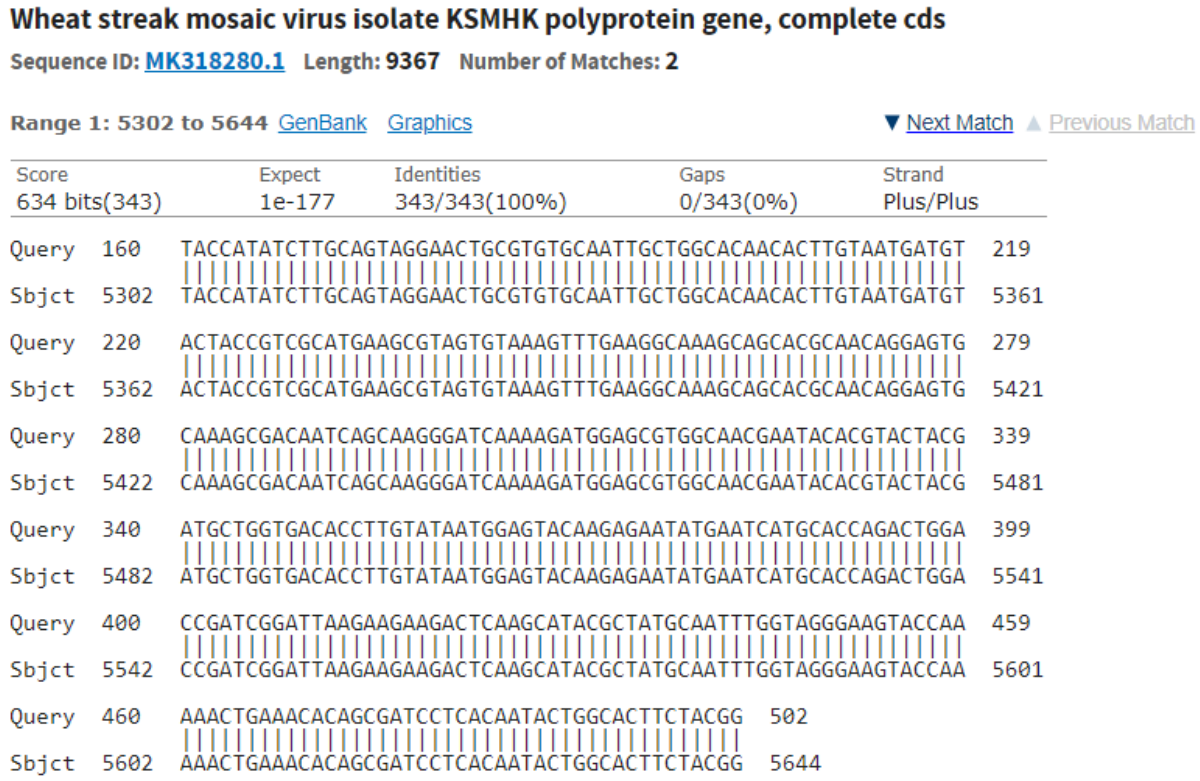
Appendix C - Supplementary Figures and Tables from Chapter 4

Appendix C. Figures and Tables

Supplementary Figure C.1. Wheat streak mosaic virus rating scale adapted from (Rupp, 2015). From left to right, resistant to susceptible rating scale 1-9 representing an effective continuum of the severity of phenotypic virus rating



Supplementary Figure C.2. Blast result of plasmid insert sequence for the determination of WSMV inserted amplicon (topoisomerase-activated vector, pCR® 2.1-TOPO) sequenced with universal sequencing primer M13 forward (-20) and M13 reverse. The blast result confirmed that the inserted amplicon matched with WSMV KSMHK isolate with 100% identities.



Wheat streak mosaic virus isolate KSMHK polyprotein gene, complete cds

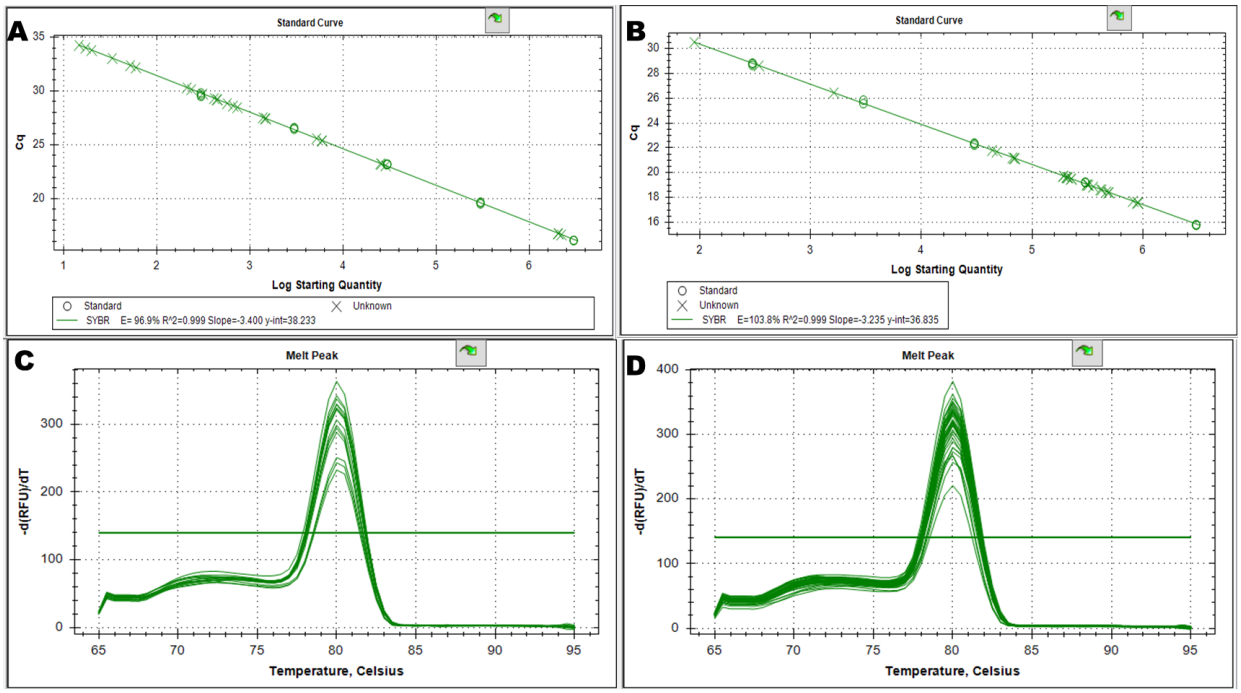
Sequence ID: MK318280.1 **Length:** 9367 **Number of Matches:** 2

Range 1: 5302 to 5644 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
634 bits(343)	1e-177	343/343(100%)	0/343(0%)	Plus/Plus

```
Query  160   TACCATATCTTGCAGTAGGAACTGCGTGTGCAATTGCTGGCACAACACTTGTAATGATGT  219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5302  TACCATATCTTGCAGTAGGAACTGCGTGTGCAATTGCTGGCACAACACTTGTAATGATGT  5361

Query  220   ACTACCGTCGCATGAAGCGTAGTGTAAAGTTTGAAGGCAAAGCAGCACGCAACAGGAGTG  279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5362  ACTACCGTCGCATGAAGCGTAGTGTAAAGTTTGAAGGCAAAGCAGCACGCAACAGGAGTG  5421

Query  280   CAAAGCGACAATCAGCAAGGGATCAAAAGATGGAGCGTGGCAACGAATACACGTACTACG  339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5422  CAAAGCGACAATCAGCAAGGGATCAAAAGATGGAGCGTGGCAACGAATACACGTACTACG  5481

Query  340   ATGCTGGTGACACCTTGTATAATGGAGTACAAGAGAATATGAATCATGCACCAGACTGGA  399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5482  ATGCTGGTGACACCTTGTATAATGGAGTACAAGAGAATATGAATCATGCACCAGACTGGA  5541

Query  400   CCGATCGGATTAAGAAGAAGACTCAAGCATACGCTATGCAATTTGGTAGGGAAGTACCAA  459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5542  CCGATCGGATTAAGAAGAAGACTCAAGCATACGCTATGCAATTTGGTAGGGAAGTACCAA  5601

Query  460   AAACTGAAACACAGCGATCCTCACAATACTGGCACTTCTACGG  502
              |||||||||||||||||||||||||||||||||||||||||||
Sbjct  5602  AAACTGAAACACAGCGATCCTCACAATACTGGCACTTCTACGG  5644
```

Supplementary Figure C.3. Standard curves and melting curves obtained by an SYBR Green RT-qPCR using a serial dilution. A serial dilution of plasmid DNA carrying the previously cloned WSMV fragment was used as a template in a 5 point, 10-fold dilution series. The five serial dilutions from 3×10^6 to 3×10^2 of plasmid were used. A and B, a standard curve of WSMV plasmid and estimated WSMV copy numbers of the samples of two assays analyzed and estimated from BioRad CFX96 Real-Time System. C and D, are the melting curves of amplicons representing A and B assay.



Supplementary Figure C.4. A standard curve of WSMV plasmid and estimated WSMV copy numbers of the samples from four separate assays obtained by an SYBR Green RT-qPCR using a serial dilution of plasmid DNA carrying the previously cloned WSMV fragment. The plasmid was used as a template in a 5 point, 10-fold dilution series. The five serial dilutions from 3×10^6 to 3×10^2 particles of plasmid were used