

WILD VIRUS IMPACTS ON FITNESS OF *PANICUM VIRGATUM* (SWITCHGRASS)
AND THE GENETIC ARCHITECTURE OF DISEASE EXPRESSION

By

Michael P. Ryskamp

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Biology – Doctor of Philosophy
Ecology, Evolutionary Biology and Behavior – Dual Major

2023

PUBLIC ABSTRACT

The undeniable abundance of wild viruses in natural plant communities prompts many intriguing ecological and evolutionary questions including the following: *What are the fitness effects of wild plant viruses on their wild hosts? If a wild virus exerts negative effects, what are the specific nature of those effects and the time-frame over which they may be consequential? Can a wild plant tolerate a prolonged infection by an antagonistic virus? In the broader picture, to what degree do viral infection dynamics vary within plant populations and among ecotypes? And, what can we glean from genome-wide association studies about the nature of genetic-based defenses against viruses in wild plants? Are there any suggestions that these defenses differ from those in crop plants?* Here, I investigated interactions between the wild virus species switchgrass mosaic virus (SwMV) and its native perennial host (switchgrass) and as a model system to begin addressing these questions. In multi-year studies, I found that SwMV can maintain prolonged symptomatic infections within individual switchgrass plants, despite reducing the function of diseased tillers. In these cases, disease was localized to particular portions of the infected plants, while other parts remained disease-free; under these conditions, infected plants maintained gains in growth over multiple years. Additionally, when plants had moderate to high levels of a foliar fungal disease (anthracnose), I found that SwMV-inoculated plants had higher fitness than uninoculated plants. In genetic studies with collaborators, I found striking differences among regional switchgrass subpopulations in susceptibility to SwMV, and identified a large number of genes associated with the expression of SwMV disease. Taken together, these research findings highlight the dynamic nature of wild plant – virus interactions and demonstrate their complexity across scales. These findings further suggest that plant–virus interactions in wild plants may be regulated by a broader genetic network than has been identified in crops.

ABSTRACT

Despite increasing effort to sequence and characterize the ubiquitous wild viruses that inhabit wild plant communities, we have very little understanding of the impacts of wild plant–virus interactions. Given the demonstrable effects of crop-infecting viruses on both crops and wild plants, there has been much conjecture about the impacts of wild viruses on wild plants, as well as increasing realization that such interactions likely have significant ecological and evolutionary implications that are currently unrecognized. Investigation of wild plant – wild virus interactions has lagged in part because they can be more challenging to study than virus dynamics in crops. This challenge stems from several factors: wild plants are generally longer lived and more architecturally complex than crop species, and they typically exhibit greater phenotypic, genetic, and age-structure variation within populations. In this dissertation, I conduct a series of investigations to characterize interactions between a wild plant and a wild virus at organismal, genomic, and population scales. For a model system, I use switchgrass (*Panicum virgatum* L.)—a candidate species for biofuel development—and switchgrass mosaic virus (SwMV), a wild virus that circulates within perennial grasslands in Midwestern North America. The main research aims of this work were (i) to characterize disease dynamics and fitness effects of SwMV infection within individual switchgrass plants over time; (ii) to characterize within-population variation in SwMV disease expression; and (iii) to identify genes potentially associated with SwMV disease expression.

In the first research chapter, I present the results of several related field studies of naturally-infected established switchgrass plants, most notably a three-year longitudinal study of individual plants representing a range of SwMV disease extent (proportion of tillers with symptoms). These results demonstrate that while SwMV infection can cause significant disease

and loss of tiller performance, it may persist for years within individual plants that continue to grow and reproduce, particularly when infection is localized to only a portion of the plant.

The second research chapter describes an additional field study in which I experimentally inoculated young switchgrass plants to further explore multi-year fitness effects of SwMV. The results suggest that SwMV infection confers tolerance to hosts in the context of moderate to high foliar fungal disease (anthracnose) extents (proportion of plant with necrotic lesions).

The third research chapter reports work I conducted with collaborators to examine SwMV disease dynamics within a diversity panel of 500+ switchgrass accessions that originated from a broad swath of switchgrass' natural range in the USA. In a common garden in Michigan, this panel naturally accumulated SwMV infection and many accessions developed disease. There was striking variation in the extent and severity of disease among ecotypes and regional subgroups, underscoring the importance of developing better understanding of how virus pressures have shaped wild plant populations. Genome-wide association studies found associations between SwMV disease and a diverse set of genes widely distributed across the switchgrass genome.

Overall, these results suggest that there are complex and nuanced mechanisms by which pathogenic wild viruses persist within the long-lived and architecturally complex plants that dominate natural landscapes, including potential benefits in certain environmental contexts. Additionally, this work suggests that the subpopulations of a wild plant species can have divergent responses to wild virus infection, and that viral disease expression may be associated with a broader and more diverse genetic network than is currently indicated from studies of crop plant–virus systems.

Dedicated to my family—Rachelle, Kincaid, and Ainsley—who let me turn our vegetable garden into an experimental research site.

ACKNOWLEDGEMENTS

This dissertation would not be possible without the efforts, support, and hard work of many people. A team of undergrads helped with processing samples and conducting field work: Nicholas Console, Brendan Canavan, Bianca France, Brett Cornell, Brooke Desposato, Zach Wilson, and Jacob Gantz. Many site managers and technicians enabled my research, including Chris Difonzo, Brooke Wilke, Jason Bonnette, Lisa Vormwald, and additional staff at the Kellogg Biological Station (KBS) and the MSU Entomological Farm. The Graduate School; the College of Natural Sciences; the Ecology, Evolution, and Behavior Program; KBS; the Department of Plant Biology; the Environmental Science and Policy Program; and the DOE (grant DE-SC0021126) provided generous funding. Kay Gross welcomed me as a visiting graduate student to KBS. My KBS writing group provided helpful feedback and encouragement while writing. Carina Baskett gave generous support, feedback, and advice while co-coordinating IBIO 355L. My family provided countless hours of childcare and volunteer field work. Ellen Cole, Kota Nakasato, Katrina Culbertson, and François Maclot, for being amazing lab mates and for assistance with PCR, sequencing, and bioinformatics. Charles Geyer and Ruth Shaw provided instrumental guidance on aster modeling. Alice MacQueen, Lucie Tamisier, Acer VanWallendael, Tom Juenger, and the Lowry lab, for GWAS support and/or switchgrass expertise. My generous Committee Members, Marjorie Weber, Lars Brudvig, and David Lowry, provided instrumental feedback, guidance, and support. My adviser, Carolyn Malmstrom, for setting a high bar and helping reach it, and for being an exceptional model of “scientist as truth teller.” My wife, Rachelle, who probably knows too much about switchgrass and plant viruses than she cares to, provided unwavering support.

TABLE OF CONTENTS

INTRODUCTION	1
CHAPTER 1: ANTAGONISTIC WILD VIRUS PERSISTS AS LOCALIZED PATHOGEN WITHIN A NATIVE PERENNIAL GRASS HOST	8
APPENDIX.....	30
CHAPTER 2: EFFECTS OF A WILD PLANT VIRUS ON HOST FITNESS IN THE CONTEXT OF A FOLIAR FUNGAL PATHOGEN	46
APPENDIX.....	59
CHAPTER 3: EXPRESSION OF WILD VIRUS DISEASE IN SWITCHGRASS (<i>PANICUM VIRGATUM</i> L.) VARIES BY ECOTYPE, SUBPOPULATION AND IS REGULATED BY A DIVERSITY OF DISTRIBUTED GENES	66
APPENDIX.....	87
CONCLUSIONS.....	112
LITERATURE CITED	114

INTRODUCTION

A growing number of metagenomic surveys in wild plant communities have discovered novel wild viruses in wild plants in unmanaged landscapes (Muthukumar *et al.*, 2009; Roossinck *et al.*, 2010; Prendeville *et al.*, 2012; Bernardo *et al.*, 2018; Shates *et al.*, 2019; Hasiów-Jaroszewska *et al.*, 2021). Together with recent advances in identifying viral taxa hidden within plant transcriptome databases (Edgar *et al.*, 2022; Lee *et al.*, 2023), these surveys suggest that the current number of described plant viruses represents only a small fraction of global plant virus taxonomic diversity. This revelation is significant because viruses exert intimate, powerful effects on the physiology and development of their hosts; therefore, viruses may therefore be influencing ecological processes and evolutionary dynamics in plant communities in profound yet unrecognized ways.

Most plant viruses have positive-sense, single-stranded RNA (+ssRNA) genomes, which are analogous in structure and function to mRNA (Dreher & Miller, 2006). Upon transmission into host cells, +ssRNA viruses release viral RNA that competes with host mRNA for translation by ribosomes (Bushell & Sarnow, 2002; Schneider & Mohr, 2003; Thivierge *et al.*, 2005). Virus infection thus alters host gene expression directly, by introducing new genetic material that is translated into multi-functional proteins, and indirectly, by interfering with and influencing plant pathways that regulate development and growth (Callaway *et al.*, 2001; Morozov & Solovyev, 2003; Culver & Padmanabhan, 2007; Ivanov & Mäkinen, 2012).

Inferring potential impacts from crop virus investigations

Studies of crop virus interactions with wild plants offer a glimpse into the potential ecological implications of wild plant–virus interactions. Crop viruses have been shown to reduce wild plant fitness (Funayama *et al.*, 1997; Alexander *et al.*, 2017), influence native plant

population dynamics (Funayama *et al.*, 2001), and alter the composition of natural plant communities (Malmstrom *et al.*, 2005a,b). Viruses can also influence important plant–biotic interactions, including pollinator preferences (Murphy *et al.*, 2023), interactions with bacteria (Aguilar *et al.*, 2020), and preferences of leaf-chewing and sap-sucking herbivores (Belliere *et al.*, 2008; Mauck *et al.*, 2010; van Molken *et al.*, 2012).

Studies of crop and model plants suggest additionally that viruses may in some cases increase abiotic stress tolerance. In *Arabidopsis thaliana*, for example, virus infection influenced an abscisic acid pathway that increased drought tolerance (Westwood *et al.*, 2013). Likewise, in a greenhouse experiment limited to pot-grown plants, Xu *et al.* (2008) observed increased desiccation tolerance in several crop species infected with crop viruses.

Viruses therefore have potential to affect plants and plant communities directly, by impacting phenotypes and fitness, and indirectly, by influencing biotic interactions and abiotic stress tolerance. Careful, detailed ecological investigation of wild plant–virus interactions will undoubtedly uncover a wide-range of complex and multi-faceted interactions with impacts at organismal and community-wide scales.

A dearth of ecological studies with wild plant viruses

Prior to the efforts involving the switchgrass mosaic virus system discussed in this dissertation, two ecological studies that most directly investigated wild plant–virus interactions took place three decades and several continents apart. In an early study involving *Kennedia* yellow mosaic virus (KYMV) and dusky coral pea (*Kennedia rubicunda*) in eastern Australia, Gibbs (1980) compared the single-season growth and survival rates of infected and uninfected seedlings planted in the field. He found that uninfected plants had higher growth rates, but were preferentially browsed by herbivores, so that infected plants had higher rates of survival in the field.

In a more recent field investigation involving green milkweed (*Asclepias viridis*) and asclepias asymptomatic virus (AsAV) in Oklahoma, Alexander et al. (2020) compared infected and uninfected plants at several prairie sites. Because their primary aim was to characterize prevalence patterns over time, evaluations of fitness were limited to comparisons of mean trait values within and among sites. Nonetheless, the authors found some limited evidence for reduced height in AsAV-infected milkweed, and showed that AsAV persisted at sites for multiple years.

These investigations suggest that long-term field studies will be important to fully evaluate wild virus interactions with the long-lived plants that dominate terrestrial ecosystems—where the impacts of infection are contextualized within the broader abiotic and biotic environment. However, neither investigation was able to compare the quantified fitness of wild viruses on wild hosts. Doing so will be a fundamental component of elucidating wild plant–virus interactions.

Aster models for quantifying fitness effects of wild plant–virus interactions

Aster-life history modeling is a powerful statistical framework that can be used to evaluate the fitness effects of viruses on wild plants. It is especially useful for estimating the fitness of perennial plants, which is the multi-year outcome of interdependent life-stages associated with disparate metrics and data distributions, including binary states (e.g., alive or dead; a Bernoulli distribution), normally distributed measures (e.g., biomass), or count data (e.g., seeds or inflorescences (Poisson or negative binomial)). Aster models aid statistical modeling over life-stages by incorporating these probability distributions into a single joint distribution to estimate integrated fitness over time (Geyer *et al.*, 2007; Shaw *et al.*, 2008). Since they were first developed in 2007, aster models have been used to evaluate local adaptation in *Mimulus guttatus* (yellow monkeyflower) (Lowry & Willis, 2010); the effects of herbivory and domestication on

Helianthus (sunflower) (Dechaine *et al.*, 2009); and the fitness effects of a crop-associated virus on upland and lowland *Panicum virgatum* L. (switchgrass) genotypes (Alexander *et al.*, 2017).

A model system for investigating wild plant–virus interactions

The field of plant virology uses a diversity of model systems for investigation. Model host plants, however, are generally annual crop plants or short-lived species like *Arabidopsis thaliana*—fast-growing, tractable systems that can often be mechanically inoculated with crop viruses. A lack of established wild plant–virus systems has hampered progress (Shates *et al.*, 2019). However, the investigation of *P. virgatum* as a potential bioenergy feedstock has opened new possibilities and resources for examination of a long-lived native perennial species and its interactions with viruses.

Switchgrass is a phenotypically and genetically diverse perennial grass that exhibits local adaptation across its broad range in North America, from northern Mexico to southern Canada (Casler *et al.*, 2007; Morris *et al.*, 2011). Traditionally, variation within the species has been associated with ecotype (a morphological categorization), ploidy level, and latitude of origin (McMillan & Weiler, 1959; Casler *et al.*, 2007; Morris *et al.*, 2011; Lowry *et al.*, 2014). In a recent comprehensive evaluation, Lovell *et al.* (2021) conducted a detailed morphological evaluation of tiller and plant traits to categorize switchgrass genotypes into upland, lowland, and coastal ecotypes. The authors used full-length genome sequences to assess relatedness and distinguish three regionally adapted subpopulations: Midwest, Atlantic, and Gulf. These designations correspond to regional geography, and thus are correlated with latitude of origin. The ecotypes are distributed unevenly among subpopulations, such that there are five main ‘ecological subgroups’ (subpopulation-ecotype combinations). The Atlantic subpopulation consists primarily of upland and coastal ecotypes; the Gulf subpopulation, lowland and coastal ecotypes; and the Midwest subpopulation, predominantly the upland ecotype.

Switchgrass is susceptible to multiple crop viruses, including barley yellow dwarf virus (Schrotenboer *et al.*, 2011) and panicum mosaic virus (Stewart *et al.*, 2014), as well as to novel viruses only recently discovered. This dissertation focuses on one of them, switchgrass mosaic virus (SwMV), a positive-sense single-stranded RNA virus in the genus Marafivirus and in the Tymoviridae family. This wild virus was discovered in 2010 in leaves of symptomatic switchgrass plants growing in biofuel plantings in Michigan and Illinois (Malmstrom *et al.*, 2022; Agindotan *et al.*, 2012). It was subsequently named switchgrass mosaic virus for the pattern of chlorotic dots and dashes it forms in leaves of infected tillers. Additional investigations revealed that SwMV is endemic to the north-central USA (Malmstrom *et al.*, 2022) and transmitted by *Graminella aureovittata* (DeLong) (Agindotan *et al.*, 2013b)—a conservative North American leafhopper strongly associated with high-quality prairie communities (its coefficient of conservatism is 13.5, one of the highest rankings) (Wallner *et al.*, 2013).

After significant investment in switchgrass genomic research, there are now established resources and pipelines for conducting population-corrected genome-wide association studies (GWAS) with hundreds of re-sequenced switchgrass genotypes (Lovell *et al.*, 2021). In GWAS, trait measures conducted on plant genotypes are correlated to known SNP markers to identify loci and candidate genes that may be involved in regulation of traits of interest. In switchgrass, GWAS has been performed with a variety of traits, including flowering time (Grabowski *et al.*, 2017), panicle traits (Zhang *et al.*, 2022), and nitrogen use efficiency (Shrestha *et al.*, 2022). Recently, there have been increasing efforts to use GWAS to understand the genetic basis of switchgrass associations with microbes, including fungal (VanWallendael *et al.*, 2022) and

bacterial communities (Sutherland *et al.*, 2022), as well as fungal disease expression (Sutherland *et al.*, 2022; VanWallendael and Lowry *et al.* unpub.).

The growing number of GWAS involving switchgrass microbial interactions is significant because most GWAS involving plant pathogen disease has involved crop hosts, and characterizing the genetics of wild plant–pathogen interactions has been identified as a critical need that can be addressed with GWAS (Demirjian *et al.*, 2023). Among major pathogen groups, however, our understanding of the genetics of wild plant responses to viruses is most limited. GWAS has been implemented with dozens of crop plant–virus systems, but *Arabidopsis thaliana* is the only non-crop plant for which viral disease has been investigated with GWAS. To date, there have been no GWAS completed to evaluate genetic regulation of wild plant responses to wild viruses, which may differ considerably from the genetic regulation of viral disease in crops, and which has been identified as a priority research area for pursuing with GWAS investigation (Demirjian *et al.*, 2023).

Summary of dissertation research

This dissertation advances understanding of wild plant virus impacts by examining interactions between SwMV and its native switchgrass host from several different perspectives across scales. In it, I have quantified the impacts of SwMV on individual plants with field studies and aster models; evaluated patterns of SwMV disease within and among switchgrass subpopulations; and identified genes potentially associated with SwMV disease expression.

Chapter one describes a three-year study of natural infection in experimental stands that I conducted to evaluate (i) longitudinal patterns of disease in individual plants and (ii) fitness impacts of disease in tillers and multi-year whole-plant fitness with aster models. The results of this study indicate that while SwMV causes disease and reduces tiller performance, it may persist

for multiple years within an individual plant, sometimes as a localized infection that is maintained even as the plant makes continued growth gains.

Chapter two reports the findings of a second multi-year field study, this one a two-year study of experimentally-inoculated (virus and mock) switchgrass, in which I quantified the effects of SwMV on fungal disease severity and two-year fitness. The results suggest that SwMV infection reduces the extent of fungal disease (anthracnose) and is associated with increased tolerance at moderate to high levels of disease extent.

Chapter three describes collaborative work in which I examined SwMV viral disease dynamics within a tetraploid diversity panel with natural SwMV infection. We found that disease resistance and susceptibility varied widely within and among switchgrass subpopulations. Using GWAS approaches, we identified a distributed set of 67 top-candidate, small-effect SNPs, representing over 100 genes with diverse functions.

CHAPTER 1: ANTAGONISTIC WILD VIRUS PERSISTS AS LOCALIZED PATHOGEN WITHIN A NATIVE PERENNIAL GRASS HOST

Introduction

Deep-sequencing has revealed a treasure trove of novel virus infections in wild plants (Muthukumar *et al.*, 2009; Bernardo *et al.*, 2018; Susi *et al.*, 2019). These ‘wild viruses’ do not generally infect crops and their impact on plant fitness is unknown. It is thought that such infections are largely asymptomatic—not causing evident disease symptoms in the host—and thus unlikely to cause harm (Cooper & Jones, 2006; Roossinck, 2012). There is thus recurrent interest in the hypothesis that wild plant viruses are predominantly benign or beneficial in wild hosts (Roossinck, 2011, 2015a,b; Pagán *et al.*, 2016). However, only a few studies have attempted to test this idea empirically. This major gap in ecological understanding merits significant attention given the widespread occurrence of virus infections within natural plant communities.

Investigation of the fitness effects of wild virus infections requires care because natural plant communities are generally dominated by long-lived plants in which fitness effects accrue under the influence of fluctuating environmental conditions over many years. The impact of viral infection may be further modulated by the relative availability of host resources. For example, variation in host supply of nitrogen and phosphorus was shown to influence barley yellow dwarf virus dynamics (Lacroix *et al.*, 2017), and resource release from clipping treatments (e.g., soil moisture gains) was found to benefit grasses with virus infection more than those without (Malmstrom *et al.*, 2006). To assess virus fitness effects under natural conditions, we therefore conducted a suite of field studies with switchgrass mosaic virus (SwMV, Genus: *Marafivirus*, Family: *Tymoviridae*), a wild virus that infects native prairie grasses in North America. We then

used a quantitative approach to integrate several years of fitness component data into multi-year estimates of individual plant fitness.

SwMV has been found in several tallgrass prairie species in the USA, but the primary identified host is *Panicum virgatum* L. (switchgrass) (Agindotan *et al.*, 2013b; Malmstrom *et al.*, 2022), a long-lived tallgrass prairie-dominant with a broad geographic range (Zhang *et al.*, 2011). Due to its potential as a bioenergy feedstock, switchgrass has been extensively studied (Jin *et al.*, 2019; VanWallendael *et al.*, 2020; Lovell *et al.*, 2021), making it an attractive model wild grass species for this work. SwMV itself exhibits convincing hallmarks of a wild virus. It is endemic to the north central USA (Malmstrom *et al.*, 2022), has never been found in crop plants, and is transmitted by *Graminella aureovittata* (DeLong) (Agindotan *et al.*, 2013b), a North American leafhopper with high fidelity to high-quality prairies (Wallner *et al.*, 2013). Most intriguingly, this wild virus often causes symptomatic disease (foliar discoloration) in switchgrass (Agindotan *et al.*, 2012, 2013b), contrary to assumptions that wild virus infections typically are asymptomatic and benign in their wild hosts.

A key focal issue of this study was the complex modular architecture of individual switchgrass plants and the frequently nonuniform distribution of disease within them. Each plant comprises a population of annual stems (tillers, Fig. 1.1) that may compete or collaborate in resource acquisition. Tillers emerge from a branching network of connected perennial rhizomes (below-ground stems) that spread laterally, presenting a large potential resource for exploitation by SwMV, which replicates in plant vascular networks (phloem). In the field, it was clear not only that individual tillers could become substantially diseased, as evidenced by strong chlorotic symptoms, but also that not all tillers exhibited symptomatic disease. We wished to examine the extent to which this disease influenced the performance of individual tillers and the longer-term

consequences for a plant as a whole. To characterize fitness impacts of SwMV infection on switchgrass, we therefore evaluated how symptom expression (i.e., ‘disease’) was localized within plants and we tracked changes in disease extent within individual plants over time. We compared growth parameters and functional properties of tillers with and without disease and used aster life-history models (Geyer *et al.*, 2007) to estimate (1) single-season fitness impacts of infection on individual annual tillers and (2) three-year fitness impacts on whole plants. To extend understanding across resource environments, we further evaluated disease interactions with tiller performance in a long-term nitrogen-gradient study. Together, these results suggest that within-host localization of disease may be an important means by which moderately pathogenic viruses persist as multi-year infections in long-lived and architecturally complex hosts.

Materials and Methods

Location and biological materials

We conducted a suite of field studies in the diverse agro-ecological landscape of Michigan (north-central USA), a temperate seasonal climate region. Studies took place from 2017–2019 at three sites detailed below. SwMV is endemic throughout the region (Malmstrom *et al.*, 2021), and natural infection occurred within all study plots. The switchgrass plants studied were 3–9 years old and well-established from prior seed sowing in 2008 and 2012 for long-term experiments. The material seeded represented two upland octoploid accessions: ‘Cave-in-Rock,’ originating from a wild population in Illinois (USDA, 2011), and ‘Shawnee,’ derived from Cave-in-Rock after only one cycle of selection (Vogel *et al.*, 1996). Shawnee is phenotypically similar to Cave-in-Rock and performs similarly in field trials (Sanderson, 2008; Malmstrom, unpublished data).

Overview of approach

To confirm that this system was suitable for investigation of multi-year plant–virus interactions, we first determined that infection could persist across years by overwintering within plant tissues. We then initiated a suite of studies (i) to evaluate within-plant disease dynamics over time and (ii) to quantify the fitness effects of switchgrass mosaic virus (SwMV) at the tiller and whole-plant levels. Much of this work centered on a three-year longitudinal field study of naturally infected plants (‘longitudinal plants’) with varied disease extent (proportion of tillers with SwMV symptoms).

To investigate within-plant and whole-plant disease dynamics, we evaluated disease expression, growth, and reproduction over multiple years in the longitudinal study plants. To characterize how disease is distributed within the plant architectural network, we excavated a mature infected plant, mapped the belowground plant rhizome network, and determined the position of PCR-tested tillers within the mapped network.

We next used aster life-history modeling to compare the multi-stage fitness of diseased and disease-free tillers from the longitudinal plants. To explore the generality of these findings across soil types, we examined virus influence on tillers across a nitrogen-addition gradient. Finally, we used aster life-history modeling to estimate three-year whole-plant fitness as a function of plant disease extent.

Belowground persistence of viral infection

In temperate seasonal climates, the tillers of switchgrass plants senesce in fall, when temperatures drop below freezing. New tillers emerge in spring from belowground rhizomes. To evaluate rhizomes as potential winter reservoirs for viral infection, we excavated ~25 x 25cm belowground sections of switchgrass root and rhizome tissue (10 – 15cm deep) in late winter

2017 from established field plants at site NF02.E2 (Table S1.1), which had shown high prevalence of natural infection in the previous season. Plants were still six to ten weeks from spring green-up, and the ground was just beginning to thaw. We washed the material, collected tissue samples (ten root and four rhizome), and extracted total RNA with a Direct-Zol RNA Miniprep Plus Kit (Zymo, Irvine, CA, USA), following the manufacturer's protocol. 50mg fresh tissue samples were homogenized in 2-mL screw top vials with 1.0mm silica beads for two minutes in a BioSpec Mini-Beadbeater. Reverse-transcription (RT) PCR was then used to test for SwMV, as described in Malmstrom et al. (2022) (see also Methods S1.1). We re-planted the remaining rhizome material (Fig. 1.2a) in potting media in a highlight growth chamber (BioChambers TPC-19, Bio-Chambers, Inc, Winnipeg, Manitoba; 21°C, 50% rh, 16h light) and evaluated foliar disease expression when tillers emerged two weeks later.

Longitudinal field study of tiller and whole-plant performance

To quantify the effects of natural infection on established switchgrass plants in the field, we conducted a detailed three-year longitudinal study of ten mature 'Shawnee' switchgrass individuals growing in a naturalistic field experiment (Michigan State University campus, East Lansing, MI; site NF01.S3 in Table S1.1). The site was one of 15 planted in 2012 across mid-Michigan for a larger landscape-scale study and was selected for ease of access. Soil at this site was predominately Conover loam (0-4% slope, aquic hapludalfs), a prime farmland class. Average 30-year precipitation for the May through August growing season here is 36.5cm; growing season precipitation was 22% lower than average in 2017, 14% lower in 2018, but 14% greater in 2019 (Climate data from: <https://prism.oregonstate.edu>). Mean growing season temperatures were near the 30-year average of ~19°C. Plants at this site experienced natural herbivory, primarily from insects, and were exposed to natural microbial interactions; no

pesticides or pest control measures were taken. The site was mowed annually in winter when aboveground tissue was senesced; mowing, grazing, or prescribed burns are typical practices for prairie management in this region.

At study initiation in 2017, the plants were about three–five years old and 26–68 cm in diameter with 40–225 tillers each. Related testing found that SwMV was the dominant virus at this location and that ~50% of tillers were SwMV-infected (Malmstrom *et al.*, unpublished data). Individual plants were selected randomly from among those at least three years old (as estimated by structure) with discernible boundaries. At the time of selection, all plants that met these criteria exhibited some foliar symptoms of SwMV disease (Fig. 1.1b); none was disease-free. We therefore randomly selected individuals along a disease expression gradient (4%–68% diseased tillers).

Disease expression and tiller functional traits

From our field experience in this region, clear symptoms of SwMV are a strong and reliable indicator of infection, as verified with molecular diagnostics (Malmstrom *et al.*, 2022). For each plant in all three years, we quantified SwMV disease by visually rating foliar symptoms in each tiller with a 0–4 SwMV disease symptom scale (Fig. 1.3a). Any tiller with one or more leaves exhibiting symptoms at level 2 or greater was considered diseased. After observing that symptoms emerged in the same plant region in year two, we began documenting the spatial distribution of disease within each plant. For tiller trait evaluation, we marked 5 diseased tillers (with clear symptoms of SwMV), and 5 disease-free tillers in each plant (100 tillers total) in early June soon after green-up and monitored them over the summer. In early August before tillers were reproductive, we randomly selected two diseased and two disease-free tillers from the marked set on each plant for further analysis, and randomly selected an additional 20 disease-

free tillers from four neighboring disease-free plants. Disease scoring revealed that 1 of the 20 originally disease-free tillers from diseased plants had developed minor but clear symptoms; this tiller was reassigned to the diseased group. Of the 60 studied, there were thus 21 diseased and 19 disease-free tillers from plants with disease, and 20 disease-free tillers from disease-free plants.

To evaluate whether infection was associated with altered performance of individual tillers, we measured several traits associated with tiller growth: height, specific leaf area, biomass, and photosynthetic rates. We measured tiller heights (soil level to tip of tallest leaf pulled vertical), quantified leaf area of each tiller's top two fully-expanded leaves with a calibrated Licor 3100 leaf area meter (Licor, Lincoln, NE, USA), and then dried and weighed the leaves to determine specific leaf area (SLA). The remainder of each tiller was harvested and quickly air-dried. Total biomass for each tiller was calculated as the sum of its leaf and stem dry mass. To test tillers for SwMV infection, we extracted total RNA from 50mg samples of dried stem tissue, first with the Direct-zol™ RNA MiniPrep Plus kit (Zymo Research, Irvine, CA, USA) and then for confirmation with the Spectrum™ Plant Total RNA kit (MilliporeSigma, Burlington, MA). RT-PCR and Sanger sequencing was completed as described in Malmstrom et al. (2022) Ssee Methods S1.1 for additional details).

To quantify the effects of SwMV disease on photosynthesis rates, we measured leaf-level photosynthesis rates in the remaining unharvested tillers of the 100-tiller set marked in June. To enlarge the sample size, we measured rates in additional diseased and disease-free tillers of an 11th nearby plant, making a total of 53 tillers evaluated. For each tiller, we evaluated foliar disease and measured photosynthesis in the top-most fully expanded leaf with a Licor-6400 portable photosynthesis system (6 x 6mm leaf chamber, 6400-02B LED bulb; chamber conditions set to 2000 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity and constant 400 ppm CO_2). Measurements

(n=53) were taken over two sampling days with minimal cloud cover on between 10am and 4pm in late August to early September 2017.

To model variation in tiller traits, we used the *lme4* package in R to run mixed models with Plant ID as a random effect. We used the *emmeans* package to estimate mean trait values and to conduct planned contrasts of three tiller groups: diseased (PCR-positive with disease symptoms), asymptomatic (PCR-positive without disease symptoms), and uninfected (PCR-negative without disease). For photosynthesis measures, tillers were not PCR-tested, so diseased tillers were compared to disease-free tillers (uninfected and/or asymptomatic).

Fitness components of individual tillers and whole plants

For assessment of within-plant tiller fitness, we marked 10 diseased and 20 disease-free tillers (where numbers were sufficient) in the 10 plants at NF01 in June 2018. We selected the closest tillers to pre-established points distributed evenly throughout each plant, and we tracked the survival and reproduction of these ~300 tillers over the growing season. We re-assessed foliar disease symptoms in August and then measured panicle lengths in early fall before harvesting them for drying and weighing. These tillers were also used to quantify the proportion of initially disease-free tillers that developed symptoms later in the season.

To permit estimates of three-year whole-plant fitness, we evaluated plant survivorship and counted all tillers in early summer and fall (2017—2019). To assess reproductive output, we counted the number of panicles produced by individual tillers in fall and measured panicle length and mass on a sub-sample (~ 10%) from each plant. Floret production of individual tillers and whole plants was then estimated from panicle length and mass using relationships quantified in the nitrogen gradient study described below. To evaluate whole-plant patterns of disease expression and growth over time, we modeled change in tiller number, panicle production, and

disease (number of symptomatic tillers and plant symptom extent) over time by conducting a repeated measures analyses (*nlme* package in R) with a *Plant ID* term nested within *Time* as a random effect.

Soil nutrient influence on performance of diseased tillers

To assess the generality of our findings in the longitudinal study, we additionally examined tiller fitness across a gradient of N-addition at the Great Lakes Bioenergy Research Center Switchgrass Nitrogen Rate Experiment (SNRE in Table S1.1) at the W.K. Kellogg Biological Station in Hickory Corners, Michigan (sandy alfisols soils). The switchgrass ('Cave-in-Rock') was established in 2008, and broadcast fertilization of granular or liquid urea ammonium nitrate began in 2009 (Jach-Smith & Jackson, 2020). By the time of our experiment in 2018, the dense structure of these established stands permitted identification only of individual tillers, not individual plants (Fig. 1.1e). We examined tillers in treatments representing 0, 112, and 196 kg ha⁻¹ of annual N addition (three replicate plots per N level); for context, annual recommended N addition for maize in this area is ~120-200 kg ha⁻¹

(https://www.canr.msu.edu/news/corn_nitrogen_recommendations). Jach-Smith and Jackson (2020) reported that soil cores from this experiment from the low (0) and high (196) N treatments (in 2014) contained similar amounts of available phosphorus (orthophosphate) but that total inorganic N was about nine-fold higher in the high N treatment (8.3 vs 77.2 kg N ha⁻¹).

As we began measures in June, it was evident that SwMV disease at the site was relatively low in prevalence and spatially clumped. To compare diseased and disease-free tillers, we thus tagged ~three groups of ~six diseased tillers identified along pre-established central transects in each plot. We then randomly tagged a corresponding similarly-sized group of disease-free tillers 1m away from each diseased group. We tagged ~24 tillers in each plot and

213 tillers overall (105 diseased and 108 disease-free). We assessed tiller survivorship and disease symptoms in June, July, and August. In early fall, we measured final panicle lengths as seeds were ripening, and we collected, air-dried, and weighed samples. For each panicle, we counted florets under a dissecting scope, regardless of floret viability or condition. Relationships between panicle length/mass and floret counts were developed to predict panicle floret production for six panicles lost before harvest, and for evaluations of tiller and whole-plant fitness described above. Both panicle mass and length were consistent predictors of floret production of panicles ($R^2 > 0.9$). After aster modeling the fitness of tillers (described below), we used the aster estimated fitness and standard errors to conduct t-tests of diseased and disease-free tillers at the three treatment levels of N addition.

Within-plant disease distribution and rhizome network

Due to its architecture, an individual switchgrass plant can be conceptualized of as a population of tillers that emerge from a network of crisscrossing underground rhizomes. These individual tillers compete with each other (e.g., for sunlight), but may also work jointly (e.g., in supporting rhizome tissue with photosynthate). In the longitudinal study, we mapped the locations of diseased tillers within plants in 2018 and 2019. For destructive analysis of tiller and rhizome relationships, we excavated an entire infected plant from a different site in the same long-term experiment as the longitudinal study (NF09.C3 in Table S1.1), located near the nitrogen gradient study on similar Kalamazoo loam soils. In mid-June 2019, we tagged each tiller on the plant to be excavated, evaluated foliar disease symptoms, and harvested stem and leaf material for diagnostic testing. After excavating the plant from the soil, we washed away fine roots and soil to expose rhizomes and their connections (Fig. 1.4a). We mapped the complete rhizome network, including locations of basal nodes (segments of crown tissue from

which new tillers and rhizomes emerge) and proaxis buds (below-ground buds on rhizomes or basal nodes). We then used RT-PCR to test for infection tillers.

Estimating tiller and whole-plant fitness with aster models

Fitness is a result of a series of conditional time-dependent processes (e.g., annual survival, reproduction), often best described with differing statistical distributions. Aster models (Geyer *et al.*, 2007) capture these relationships with model structures that incorporate different life-history data distributions that depend on one another (i.e., ‘graphical nodes’ in aster models). The aster package (Geyer, 2021) in R (R Core Team, 2021) uses this nodal model structure to create a joint probability distribution for estimating fitness across data distributions and timespans (Shaw *et al.*, 2008). Our models of tiller fitness are based on i) the likelihood that spring-emerging tillers become reproductive (first graphical node), and ii) end-of-season floret counts (second graphical node). This model structure incorporates infection effects on tillers before they become reproductive, which would be excluded in analyses based solely on end-of-season measures.

To assess multi-year whole-plant fitness, we integrated seasonal tiller counts, panicle counts, and per-panicle floret production estimates (from sub-sampled panicle measures) and then evaluated these fitness estimates as a function of symptom extent (proportion of diseased tillers in each sampling year; initial disease extent and end-of-season disease extent for years two and three). To incorporate the sub-sampled panicles into the aster model estimates of plant fitness (~10% sub-sample of panicles from each plant), we included a sub-sample node in the graphical structure of the aster model as described in Stanton-Geddes *et al.* (2012). We also used previously unreported aster methods to evaluate over-dispersion in Poisson-modeled data; to

estimate the shape parameter for negative binomial distributions; and to estimate variance for multi-year subsampled fitness data. These methods are detailed in Ryskamp and Geyer (2023).

Results

Infection overwinters in rhizomes and roots

SwMV RNA was detected by RT-PCR in all (4/4) rhizome samples and in 30% (3/10) of the root samples that we excavated at winter's end (Fig. 1.2a-b). All tillers emerging from the rhizomes showed symptomatic SwMV disease (Fig. 1.2c-d).

Viral disease associated with altered tiller traits

Consistent with previous testing (Malmstrom *et al.*, 2022), the presence of SwMV foliar disease symptoms in tillers predicted positive RT-PCR tests for SwMV with 100% accuracy (Fig. 1.3c). Tillers without disease (i.e., symptoms) were either PCR-negative ('uninfected'; 54%, 21/39) or were asymptomatic and had mostly faint gel bands (56%, 10/18) (Fig. 1.3c), suggesting low viral titer. Disease-free tillers collected from disease-free plants were nearly twice as likely to test negative (70%, 14/20) than disease-free tillers collected from plants with disease in other tillers (37%, 7/19) (test of difference between two proportions, $Z = 2.076$, two-tailed $p = 0.0379$). Amplicon sequencing showed that nucleotide sequences of the viral coat-protein coding regions were >99% similar among plants (Fig. S1.1).

We compared traits of tillers by virus status as determined by PCR testing and symptom expression (diseased and PCR-positive; disease-free and PCR-positive; disease free and PCR-negative). In planned contrasts, diseased tillers accrued 41% less dry mass than uninfected tillers ($t = 5.15$, $df = 49.8$, $p < 0.0001$; Fig. 1.5a) and were 15% shorter ($t = 4.88$, $df = 51.2$, $p < 0.0001$; Fig. 1.5b). Diseased tillers produced 20% less leaf area ($t = 2.69$, $df = 56.9$, $p = 0.009$; Fig. 1.5c) and had significantly less leaf mass per unit area (30% increase in SLA; $t = 5.8$, $df = 50.2$, $p <$

0.001; Fig. 1.5d). The traits of asymptomatic tillers (i.e., disease-free PCR-positive tillers) were statistically indistinguishable from those of uninfected tillers ($t \leq 1.404$, $df \geq 48$, $p \geq 0.167$). In our separate comparison of leaf-level photosynthesis rates of diseased and disease-free tillers (not PCR-tested), diseased tillers had 15% lower rates of photosynthesis ($t = 2.48$, $df = 48$, $p = 0.017$; Fig. 1.5e).

Disease persists for years within individual plants that nonetheless maintain gains in growth and reproduction

SwMV disease persisted across years within all but one individual plant as each continued to make gains in growth and reproductive fitness components. Over the three years of the longitudinal study, mean total tiller number per plant increased 57% (repeated measures, $F_{5,45} = 17.8$, $p < 0.0001$; red line Fig. 1.6b) and mean panicle production increased 51% ($F_{2,18} = 22.4$, $p < 0.0001$; red line Fig. 1.6d).

Within each individual plant, the absolute number of diseased tillers remained relatively constant over the three years of the study (repeated measures, $F_{4,36} = 1.4$, $p = 0.248$; redline Fig. 1.6a), while their total tiller numbers of each plant and panicle production rose (Fig 1.6b, d). As a result, the mean proportion of symptomatic tillers (disease extent) per plant fell between 2017 and 2019, from 28% to 17%, on average ($F_{4,36} = 7.5$, $p = 0.0002$; red line in Fig. 1.6c). Intriguingly, disease was generally localized to the same sub-regions of each plant across years (Fig. S1.2).

Patterns of disease within a plant rhizome network

Analysis of the excavated rhizome network likewise revealed clustering of disease. At the plant's center was a dense clump of 10–20 relatively small basal nodes (Fig. 1.4b, red arrow), representing the oldest network portion established during the plant's first years; it is drawn as a single central region (Fig. 1.4c, red arrow) because it contained no living tillers and its density

made it difficult to delineate individual connections. Eleven sub-networks emanated from the center, each connected to it by a single, distinct rhizome (Fig. 1.4c, solid blue lines). Among these 11 sub-networks, we identified 54 nodes. Sixty-nine percent (37/54) of the nodes supported live tillers, all with a single tiller except for one node that supported two tillers. Of the tillers, 61% exhibited foliar disease symptoms, and 39% were disease-free. The remaining seventeen (31%) nodes did not support live tillers but were still living and had healthy rhizomes and/or buds attached. All 21 diseased tillers tested were PCR-positive for SwMV. The two other diseased tillers were too small to test (Fig. 1.4c, single asterisks). Nearly all disease-free tillers (93%, 14/15) also tested PCR-positive. Amplicon sequencing from 22 tillers, including four that were disease-free, found that the sequences were identical ('CA02' in Fig. S1.1).

Diseased and disease-free tillers were clustered by sub-network. Ten of the 11 sub-networks were composed entirely of either diseased or disease-free tillers. Within the single sub-section that contained both tiller types, diseased and disease-free tillers were segregated into two sub-groups branching from a single living node (Fig. 1.4c, '1'). Based on tiller and node locations, we estimate that this node originated ~2 years previously. The only PCR-negative tiller (Fig. 1.4c, '4') was located at the network periphery, where it branched from a node (Fig. 1.4c, '5') with an asymptomatic tiller that gave only a faint gel band (Fig. 1.4d, arrow). Across the 11 sub-networks, the number of tillers supported per sub-network fell as the proportion of diseased tillers in it rose (Results S1.1), consistent with our findings that disease reduced tiller performance.

Diseased tillers are less fit across resource (N) levels

In the nitrogen gradient study, SwMV disease was negatively associated with tiller performance—an association that was apparent across all three nitrogen-addition levels (0, 112,

and 196 kg N ha⁻¹ yr⁻¹). Disease-free tillers with asymptomatic or no infection were ~1.5x more likely to develop a panicle (78%, 84/108) than diseased tillers with symptomatic infection (57%, 60/105) ($Z = 3.217$, two-tailed $p = 0.0013$). Our integrated aster model estimates represent tiller fitness as the expected floret production of spring-emerging tillers. Diseased tillers were less fit than disease-free ones and produced fewer florets across all levels of N addition, with fitness differences most pronounced at higher N (Fig. 1.7). With no nitrogen addition, the fitness of diseased tillers was 41% lower than that of disease-free tillers ($t = 2.19$, $df = 70$, $p = 0.032$); at 112 kg N ha⁻¹, diseased tillers were 45% less fit ($t = 3.674$, $df = 69$, $p = 0.0005$), and at 196 kg N ha⁻¹, diseased tillers were 56% less fit ($t = 3.15$, $df = 68$, $p = 0.0024$).

In within-plant comparisons, diseased tillers typically have lower fitness

Consistent with the above patterns, the average estimated fitness of diseased tillers in our longitudinal study plants was about half that of disease-free tillers within the same plant (Fig. 1.8b), and the aster model that included *Symptom Extent* as a fixed effect explained significantly more variation than the null model ($\chi^2 = 20.998$, $df = 1$, $p < 0.0001$). There was among-plant variation in fitness differences between tiller types, and model comparisons supported inclusion of a *Plant*Disease* interaction term that allows for different fitness estimates at the individual plant level ($\chi^2 = 36.3$, $df = 9$, $p < 0.0001$), indicating that effects of SwMV disease may vary across plants.

Integrated three-year whole-plant fitness is greatest with moderate disease extent

We expected integrated three-year whole-plant fitness in the longitudinal study to decline as a simple linear function of the proportion of diseased tillers within each individual plant. The relationship between the proportion of diseased tillers (disease extent) per plant and three-year

plant-level fitness, however, appeared curvilinear, with peak fitness at moderate disease proportions (~20% of tillers) (Fig. 1.9).

Discussion

To our knowledge, this study is the first to quantify the multi-year fitness effects of a wild plant virus infection on a long-lived wild host in natural field conditions. Our findings provide a counter-example to the common assumption that wild viruses are asymptomatic and benign in wild plants, and our work provides new insight regarding possible mechanisms of long-term host–pathogen coexistence. We found that wild virus infection caused perceptible disease in perennial native switchgrass, which was associated with lower fitness of individual plant tillers. However, disease was distributed non-uniformly within plants and localized to specific host modules. Infections persisted in hosts that maintained annual gains in growth and reproduction, and the proportion of modules that were diseased generally fell over time as plants continued to produce new disease-free tillers. Multi-year host fitness was greatest in plants with moderate levels of disease, suggesting trade-offs in responses to infection or even possible relative benefits of infection. More broadly, this work demonstrates that antagonistic wild viruses can indeed persist in long-lived wild hosts, and it highlights the importance of within-host dynamics in shaping fitness outcomes and mediating long-term infection dynamics.

Testing assumptions about wild plant–virus interactions

Our work represents a novel effort that combines long-term observation of disease expression in plants with detailed seasonal measures to quantify multi-year fitness. This complementary approach offers one of the most detailed characterizations of wild plant–virus interactions to date—and specifically addresses several of the widespread assumptions regarding wild plant–virus interactions.

While there is increasing interest in wild plant viruses, there has been only limited study of the fitness effects of these widespread interactions. At present, most research on wild plant viruses focuses on sequencing the viromes of wild plant communities (Muthukumar *et al.*, 2009; Roossinck, 2012; Bernardo *et al.*, 2018; Hasiów-Jaroszewska *et al.*, 2021) or on characterizing wild virus populations in different landscape contexts (Rodríguez-Nevado *et al.*, 2017; Sallinen *et al.*, 2020; Susi & Laine, 2021; Maclot *et al.*, 2023). These virological surveys—along with analyses of global nucleic acid sequence databases (Edgar *et al.*, 2022; Lee *et al.*, 2023)—have conclusively demonstrated that wild virus infections are abundant and common in wild plants. Perhaps because many of these surveys were conducted by ‘blind’ selection of plant tissues (i.e., without regard for symptom expression), there has been recurrent speculation that wild plant viruses do not cause symptoms in wild hosts, and that wild viruses are therefore predominantly benign or beneficial in wild plant hosts (Fraile & García-Arenal, 2010; Pagán *et al.*, 2016). This conclusion appears to be predicated on the assumptions that only symptomatic infection causes harm, and that long-lived plants and pathogenic viruses could not co-exist. However, there is a dearth of investigation attempting to characterize wild plant–virus interactions and their long-term fitness impacts.

As a virus that is clearly associated with symptomatic disease, SwMV can be added to the list of symptom-causing wild viruses in wild plants, which includes viruses from South America (Jones & Fribourg, 1979), North America (Brunt & Stace-Smith, 1978), Western and Southern Africa (Thottappilly *et al.*, 1992; Vincent *et al.*, 2014), Australia (Randles *et al.*, 1976, 1981, 1982; Gibbs, 1980; Skotnicki *et al.*, 1992; Webster *et al.*, 2007; Wei *et al.*, 2012), and Asia (Yoo *et al.*, 2017). Our findings, alongside this list, demonstrate that wild plant viruses are not necessarily non-symptomatic, as often assumed, and they suggest that as investigation expands, a

wider range of symptoms associated with wild virus infection in wild hosts will likely be identified. Our study does, however, offer some support for the assumed relationship between symptom expression and negative host impacts. In the SwMV system, we observed that negative impacts were isolated to diseased tillers, and that the traits of asymptomatic and uninfected tillers were similar to each other. However, it is important to underscore that other plant viruses that do not cause obvious symptoms in leaves, may still have fitness and other impacts on hosts (Alexander *et al.*, 2017, 2020; Torrico *et al.*, 2018; Fukuhara *et al.*, 2020).

Lastly, our work also speaks to the assumption that wild plant viruses are mostly benign or beneficial in wild hosts. Our results demonstrate that diseased tillers were less fit than disease-free tillers, and our observations of disease at the whole-plant level suggest a possible means by which pathogenic viruses may coexist within long-lived hosts. Given the wide-ranging variety of expansive forms and architectures of the long-lived plants that inhabit natural ecosystems—many of which can live for hundreds of years—it seems possible that the pattern of localized disease may be a common means by which wild plants coexist with viral pathogens. Instead of being broadly benign, perhaps wild plant–virus interactions are nuanced and complex, much like plant interactions with bacterial and fungal symbionts are regarded to be.

Localized disease dynamics and coexistence

In our longitudinal study of ten plants over three years, we found strong evidence that nonuniform and localized disease allowed for persistent infection in plants that nonetheless continued to increase tiller and panicle production each year. Additionally, there is some suggestion that localized disease may be associated with relative fitness benefits for some plants. It is thus possible that nonuniform and localized disease may be structuring fitness effects and promoting long-term persistence of SwMV infection in switchgrass.

Other studies have reported nonuniform virus infections in perennials (Crosslin *et al.*, 1999; Ferri *et al.*, 2002; Jridi *et al.*, 2006; Kominek *et al.*, 2009; Dawson *et al.*, 2013). However, many of these studies have been conducted in agronomically important crops and are generally aimed at optimizing seasonal detection techniques for economically important viruses (but see Jridi *et al.* 2006 and Honjo *et al.* 2020). Therefore, they do not seek to characterize their observations within a context fitness or coexistence.

The interplay between hosts and viruses has often been cast as an evolutionary arms race of sorts. Instead of an arms race, wild plant–virus interactions may commonly be a *détente* in which diverse and nuanced controls lead to longer-term coexistence. Coexistence and/or moderated impact might be more likely to emerge when disease is directly or indirectly fire-walled within long-lived hosts that can continue to grow and reproduce—or when hosts are long-lived perennials that offer overwinter refuge, and not annual crops that only represent a single-season of host resources for rapid exploitation (e.g., Malmstrom *et al.*, 2017). In the case of the switchgrass system reported here, disease is localized and the virus overwinters in host rhizomes. Thus, despite its negative impact on diseased tillers, the virus is likely inhibited from accessing the complete host network—and may be experiencing some selective pressure for moderated harm.

Intra-plant heterogeneity and fitness impacts: potential benefits of infection

Given the reduced fitness of diseased tillers, we expected whole-plant fitness to decline linearly with disease extent (the proportion of tillers with symptomatic disease). Instead, we found support for a curvilinear relationship between disease extent and fitness, with peak fitness at moderate levels of disease (~20%). It is possible that fitness peaks associated with moderate disease extent represent an optimal immune response that balances growth costs associated with

heightened defenses with the costs of greater disease itself (Brown & Rant, 2013). Alternatively, intermediate levels of viral disease expression may confer some host benefit at the whole-plant level, perhaps by increasing intra-plant trait variation.

Increased trait variation has been linked to individual- and community-level benefits for productivity and resilience (Roscher *et al.*, 2012; Sakschewski *et al.*, 2016), resource acquisition (Givnish 1988, Osada *et al.* 2014, Ventre-Lespiauq *et al.*, 2017), defense against herbivores (Orians *et al.* 2002, Wetzal and Thaler 2016, Wetzal and Meek 2019) and pathogens (Jarosz and Burdon 1988, Burdon *et al.*, 1989, Orians and Jones 2001). For individual plants, trait variation may moderate the negative impacts of pathogens and herbivores, and may serve as a 'bet-hedge' against unpredictable fluctuation in pathogens and herbivores that often have specific trait requirements or preferences. For example, an herbivore may expend more energy foraging for leaves that meet specific criteria when a leaf canopy is variable; likewise, intra-plant variation may reduce plant-level damage by a pathogen that can colonize hosts only under specific conditions (Orians & Jones, 2001; Herrera, 2009, 2017; Wetzal & Thaler, 2016). In the SwMV system, the traits of diseased and disease-free tillers differed notably. While diseased tillers contribute less than disease-free tillers to floret production, they may expand the phenotypic range of tillers in a way that enhances whole-plant fitness. Moderate levels of disease may represent an optimal level of variation beyond which the negative effects on tillers begin to accrue.

Moreover, virus-induced effects on trait variation within or between individual plants has potential to influence several types of plant-biotic interactions, including herbivory and pressure from other pathogens. In our study, diseased leaves had greater specific leaf area (SLA), one of the four traits we measured, which has been shown to explain variation in herbivory in

switchgrass (Schuh *et al.*, 2019). It is likewise possible that SwMV-mediated trait variation reduces damage from leaf-chewing or sap-sucking invertebrates. The differences in physical traits we observed among diseased and disease-free tillers may reflect notable variation in tissue chemistry, which in some plants varies more than physical traits (Siefert *et al.*, 2015), but was not measured here. SwMV-mediated variation in tissue chemistry might reduce damage from herbivory, or might serve as a bet-hedging strategy to protect against pathogens that respond to specific tissue nutrient levels, such as those that cause rust (Danial & Parlevliet, 1995; Devadas *et al.*, 2014) and anthracnose (Nam *et al.*, 2006; Schmid *et al.*, 2018), which are common in switchgrass.

Future Directions

The SwMV-switchgrass system is a promising model for investigation of wild plant virus interactions. Future research could use experimental inoculations to conduct more targeted quantifications of the effects of SwMV on host traits and fitness and confirm causal relationships underlying associations identified here. Additional investigation is also needed to elucidate the physiological mechanisms and genetic regulation behind the pattern of localized infections in individual plants.

We conducted our work with a single upland octoploid accession (cv. ‘Cave-in-Rock’), but switchgrass is a diverse species characterized by regionally adapted populations throughout its range (Lovell *et al.*, 2021). Additional work is needed to evaluate population-level variation in switchgrass interactions with SwMV and the degree to which outcomes of infection vary among genotypes.

Overall, our work demonstrates that coexistence (i.e., long-term infection) between an antagonistic wild virus and a long-lived wild host is possible. We propose within-host

localization of disease as an important mechanism that enables coexistence between long-lived hosts and viruses. Given the numerous other reports of non-uniform virus infection in perennials, this phenomenon may be a common means of pathogen coexistence with wild hosts in unmanaged landscapes.

APPENDIX

Tables and figures

Figure 1.1. Virus interactions with switchgrass (*Panicum virgatum* L.) can be explored at different scales. (a) Tillers are leafy grass stems that can produce terminal reproductive structures (panicles, white arrow). When infected, tillers may exhibit (b) foliar disease symptoms. (c) Tillers emerge from proaxis nodes (stem base, red arrow) that occur along rhizomes (belowground stems, white arrows). Roots (blue arrow) emerge from nodes and rhizomes. (d) Individual plants are composed of populations of tillers (dozens to hundreds) that emerge in spring and senesce in fall. (e) In a dense, monospecific stand, the boundaries of individual plants are harder to discern, and competition among tillers and plants appears notable.



Figure 1.2. Switchgrass mosaic virus (SwMV) overwinters in rhizomes. (a) Switchgrass (*Panicum virgatum* L.) rhizomes (blue arrows), and to a lesser extent roots, harbor switchgrass mosaic virus (SwMV) infection over winter, as indicated by (b) positive RT-PCR tests of extracted material. (c–d) After planting PCR-positive rhizomes, the tillers that emerged exhibited SwMV disease symptoms.

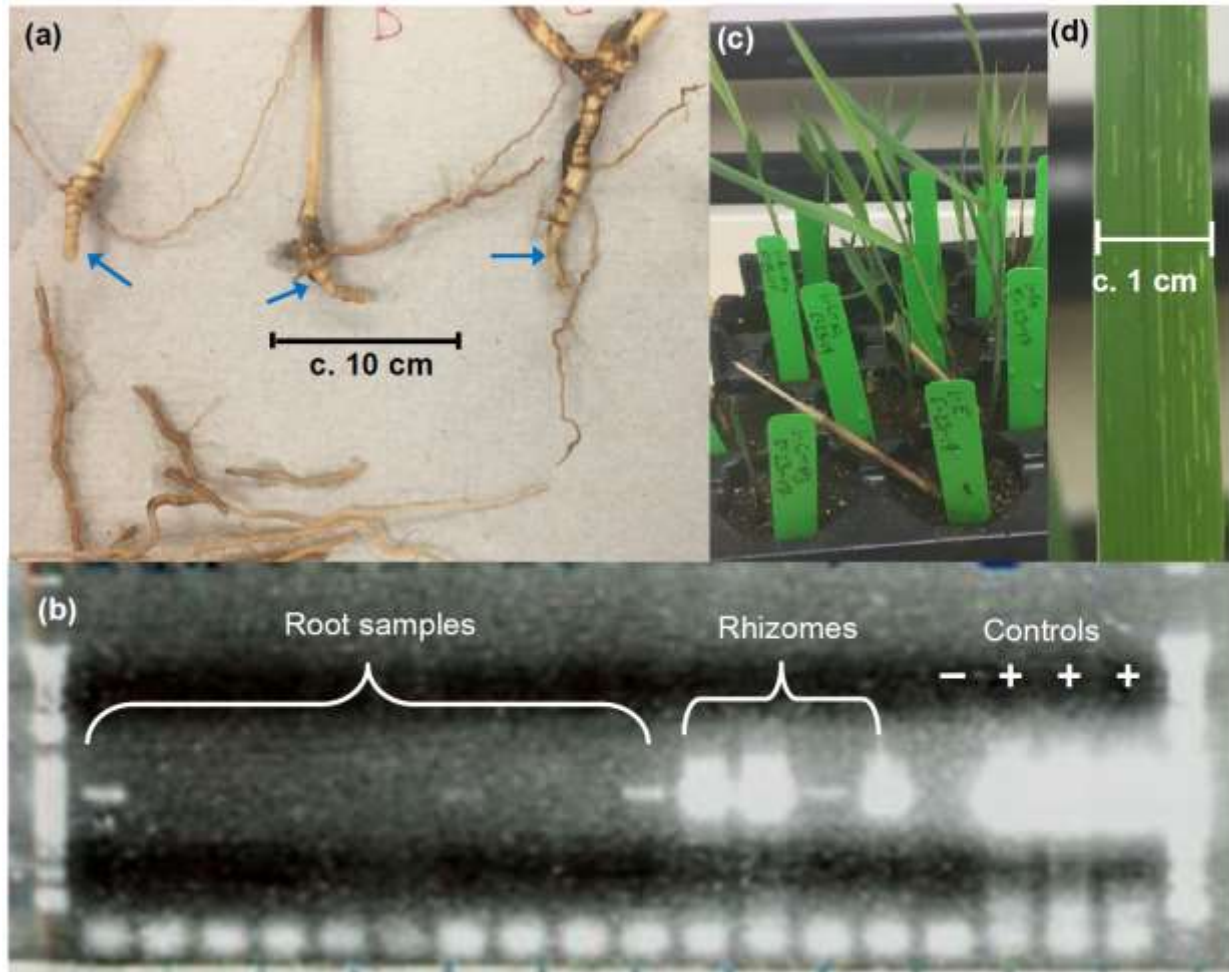


Figure 1.3. In switchgrass (*Panicum virgatum* L.), characteristic foliar disease indicates switchgrass mosaic virus infection. (a) Tillers were scored for maximum foliar symptom severity: 0 = no symptoms; 1 = possible symptoms - a few chlorotic dots/dashes in <10% of leaf area; 2 = characteristic pattern of chlorosis in 10–25% of leaf area; 3 = chlorotic pattern in 26–50% of leaf area; 4 = chlorotic pattern in >50% of leaf area. (b) In plants with disease, 100% of diseased tillers and 63.2% of disease-free tillers were PCR-positive (b, right panel). In disease-free plants, 30% of tillers were PCR-positive (b, left panel). (c) Bright SwMV PCR gel bands (~600 bp) from diseased tillers with 1kb plus ladder, confirmed with amplicon sequencing. Negative oat control shows non-target primer binding. (d) Example gel bands from disease-free tillers. Yellow brackets (in c & d) represent samples from two tillers that were used as additional positive controls for the non-diseased tillers (in d) as a means to compare relative band brightness across gels.

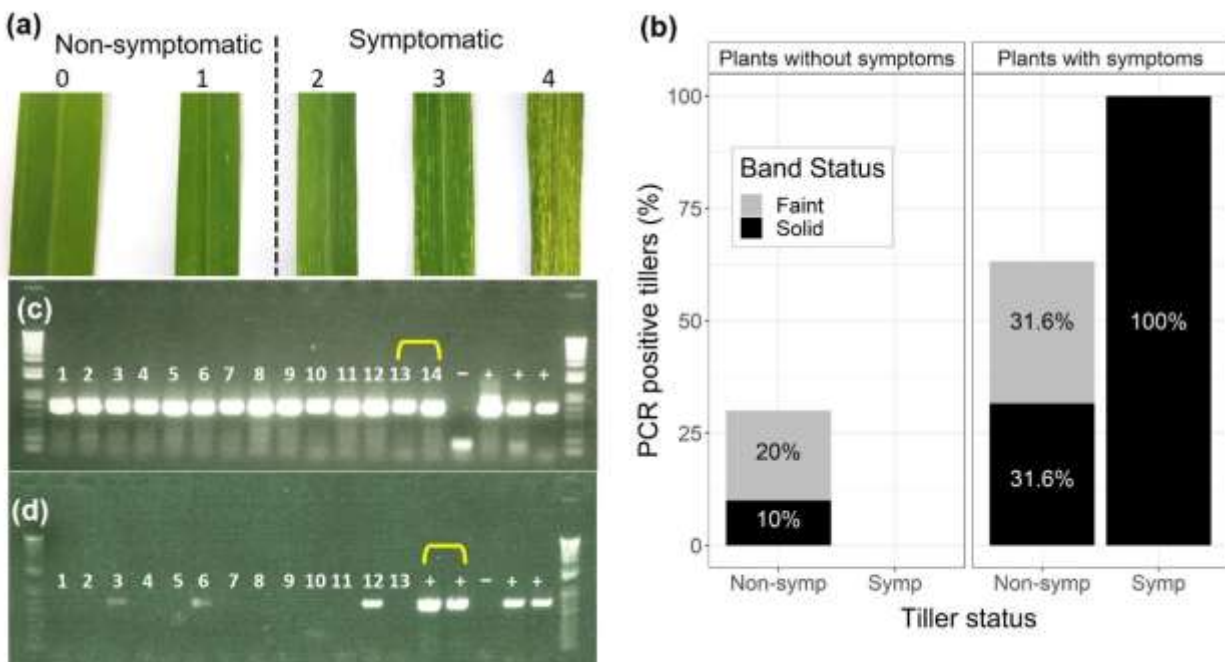


Figure 1.4. Disease is localized to rhizome sub-networks. (a) Aboveground and belowground structures used to evaluate within-plant virus distribution: rhizomes (black arrows), nodes ('1'), buds ('2'), and active tillers ('3'). Senesced tiller stumps from previous growth seasons were used to help determine the relative age of nodes within the network ('4'). (b) Excavated rhizome network after tillers were removed for RT-PCR testing. The network center had no active tillers (red arrows in b and c). (c–d) Map of the positions of tillers, nodes, and buds within the connected rhizome network. Most nodes had either one or no active tillers, but one node had two active tillers (double asterisk, c).

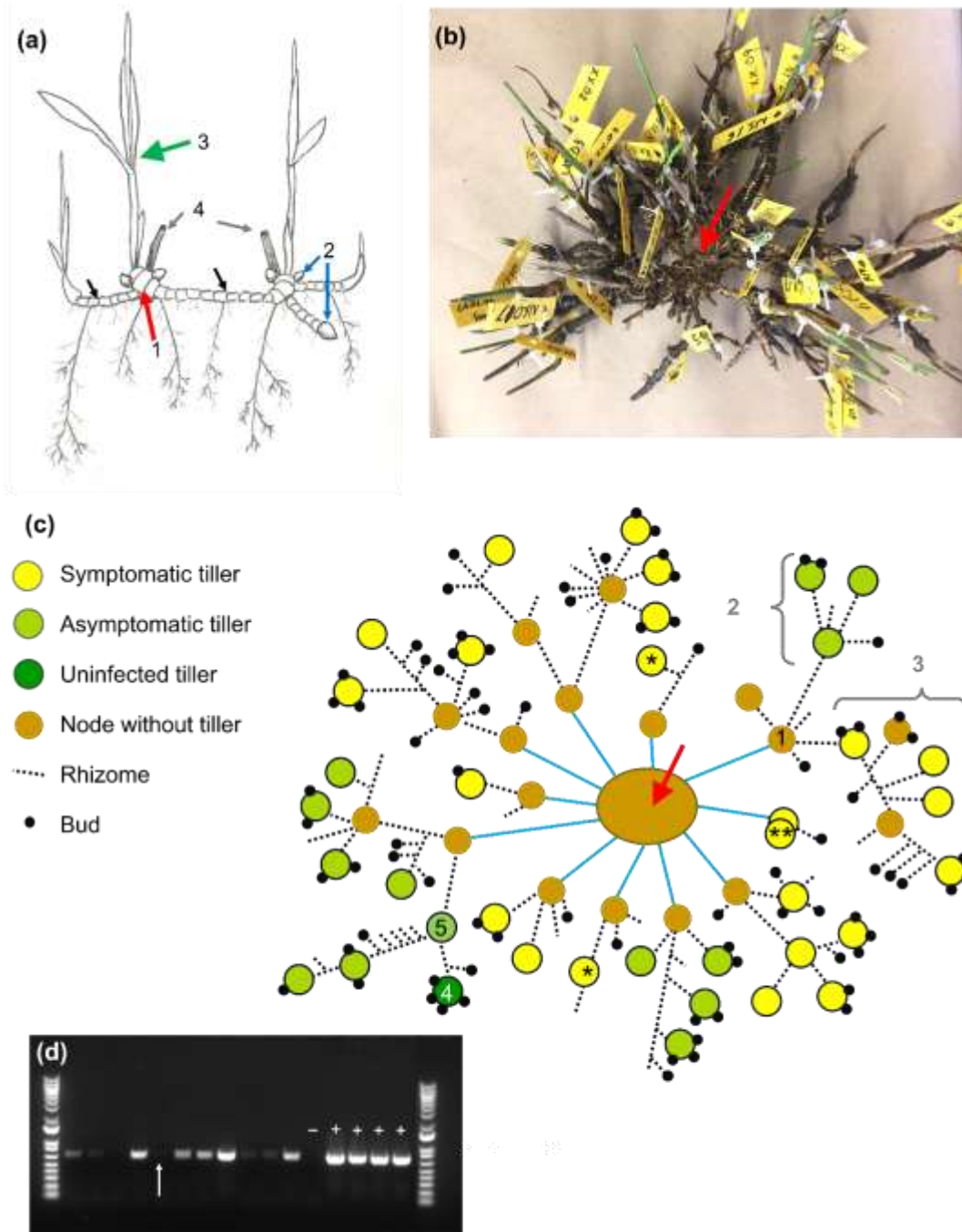


Figure 1.5. Traits of diseased tillers (yellow) differ from those of disease-free tillers, including asymptomatic (PCR-positive, light green) and uninfected tillers (PCR-negative, dark green). (a) Diseased tillers weighed less, (b) were shorter, (c) produced less leaf area, and (d) had higher specific leaf area (SLA). (e) Leaves of diseased tillers also photosynthesized at lower rates than those of disease-free tillers (asymptomatic and no infection indicated together by green bar). Tillers in a-d were PCR-tested. (f) Principal components analysis of tiller traits for PCR-tested tillers. Black lines represent ± 1 SE.

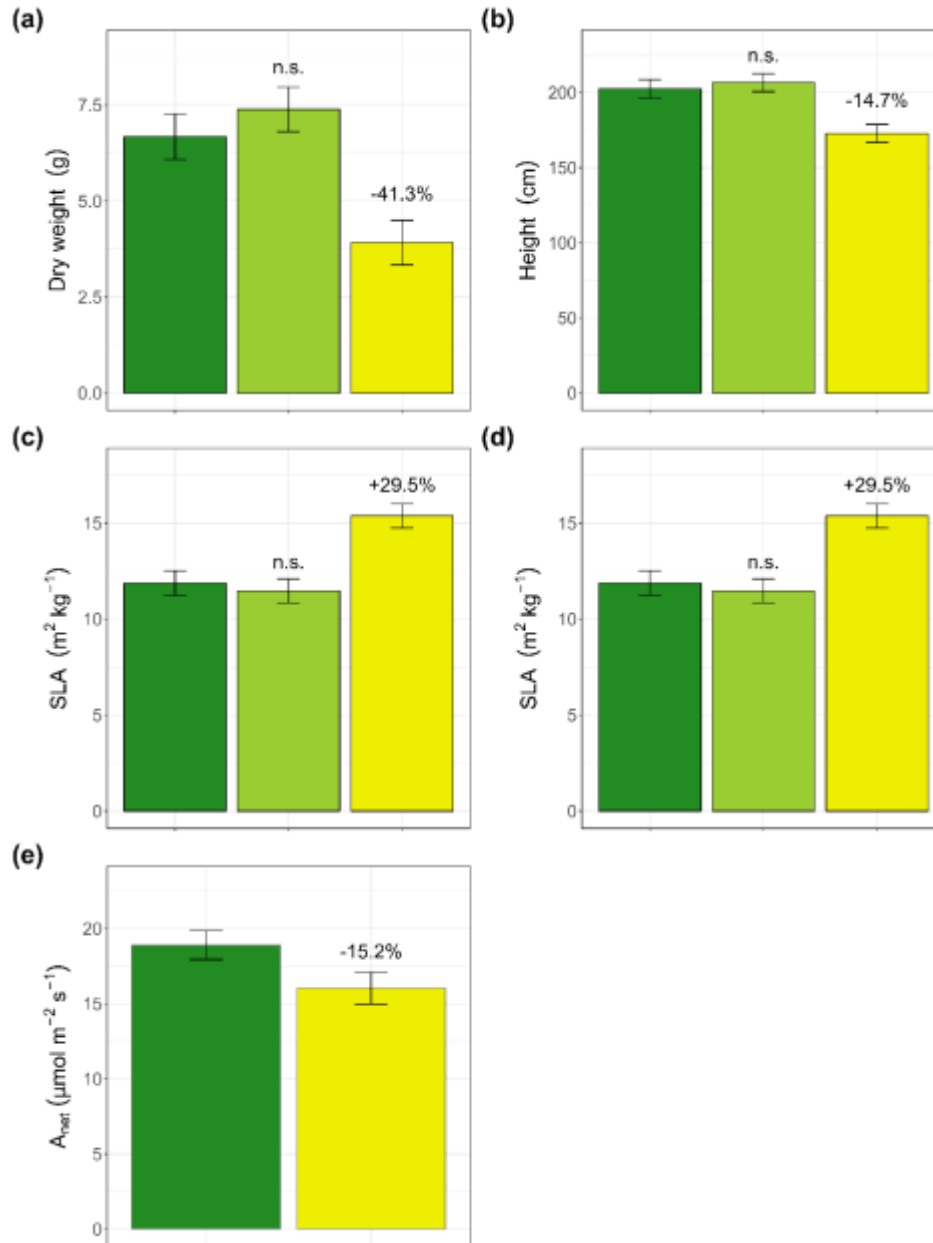


Figure 1.6. Over three growing seasons (2017–2019), disease persisted in plants that continued to grow and reproduce. (a-d) Change in diseased tillers, total tillers, symptom extent (proportion of diseased tillers), and panicle production over the three-year period. Thick red lines represent model estimated means for repeated measures tests.

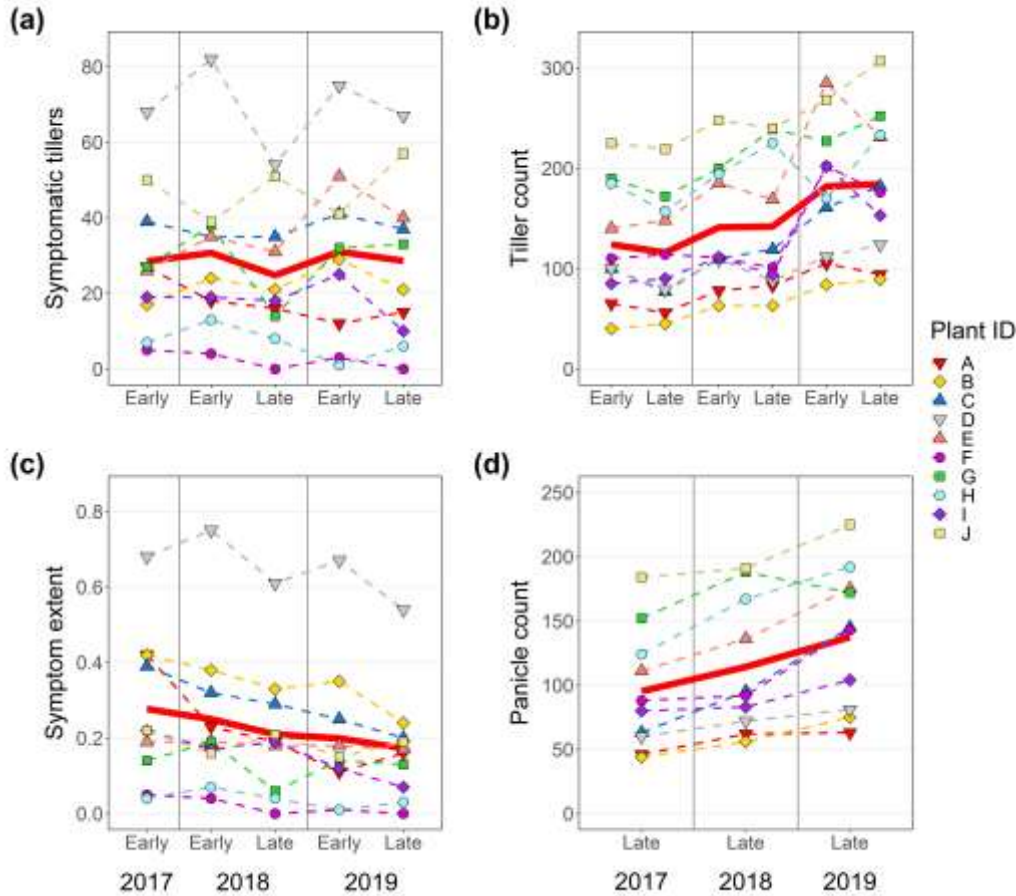


Figure 1.7. Diseased tillers were less fit at all resource (N) levels tested. Across all levels of N, symptomatic tillers were ~ 14% less likely to become reproductive than non-symptomatic tillers. The aster model incorporates this difference into the final fitness estimate of expected floret production. Vertical lines represent ± 1 SE. Estimates for means and standard errors for each group were generated with an aster model that incorporated an $N \times \text{Symptom Status}$ interaction term, which allows fitness effects of symptom expression status to vary at different levels of N.

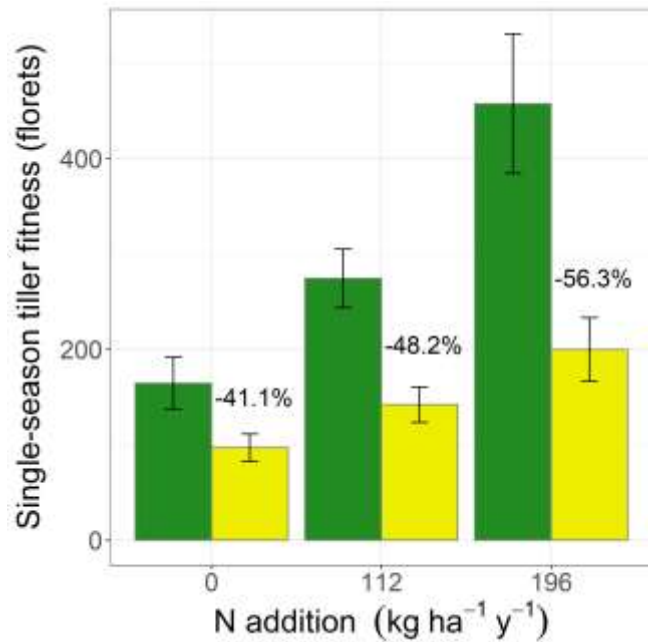


Figure 1.8. Within plants, symptomatic tillers are less fit than non-symptomatic tillers. (a) Among the ~ 200 non-symptomatic tillers that were identified in early summer, only 1.7% developed symptoms later in the growing season. (b) Aster model estimated fitness of symptomatic and non-symptomatic tillers for ten plants. No relationship was found between tiller fitness and plant-level disease extent (x-axis is ordered, left to right, from lowest to highest disease extent). Vertical lines represent ± 1 SE. The aster model used to estimate tiller fitness incorporates a *Plant*Symptom Status* interaction term that allows estimates to vary by Plant ID and symptom expression status of tillers.

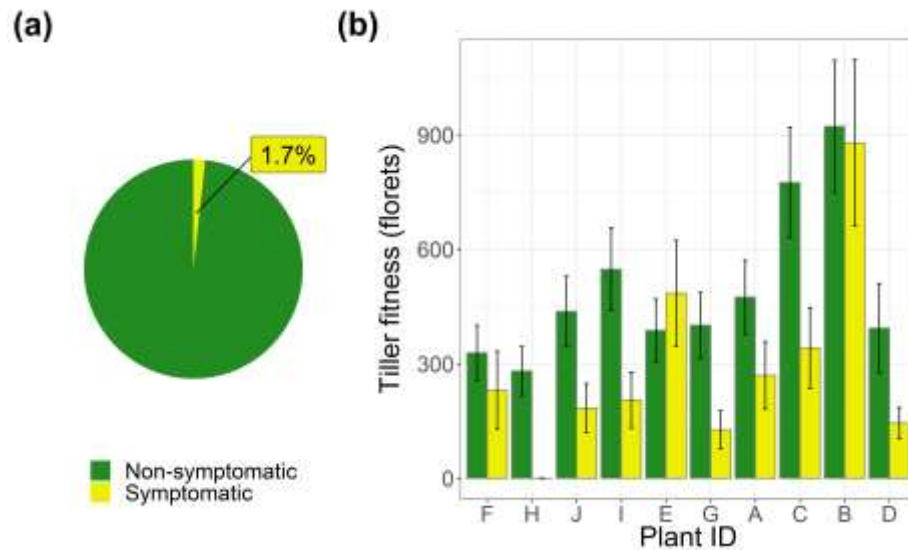


Figure 1.9. Three-year fitness (± 1 SE) as a function of mean symptom extent (proportion of tillers with symptoms). Fitness estimates are based on aster modeling that integrated three years of tiller counts, panicle counts, and floret production data. Our final model estimated three-year fitness as a function of annual symptom extent and includes a *Plant ID* term that allows fitness to vary at the individual plant level.

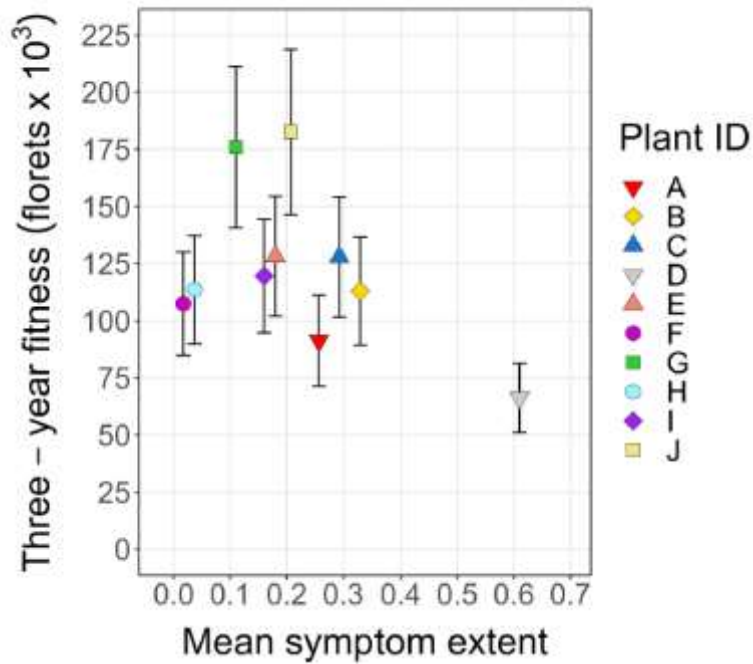


Table S1.1. Description of experimental field sites.

Site ID	Experiment(s)	Site details
NF02.E2	RT-PCR analysis of switchgrass roots and rhizomes	Located on private farm land in Clinton County, Michigan. This 400m ² plot was seeded in 2012 with <i>Panicum virgatum</i> (L.) (switchgrass, cv. "Shawnee"), <i>Andropogon gerardii</i> (big bluestem), <i>Sorghastrum nutans</i> (indiangrass), and <i>Elymus canadensis</i> (Canada wild rye). In previous field seasons, this plot had maintained relatively high levels of SwMV infection prevalence.
NF01.S3	RT-PCR analysis of 60 tillers and evaluation of tiller traits Within-plant comparison of symptomatic and non-symptomatic tillers Three-year longitudinal study of symptom expression and plant growth	Located in East Lansing, MI (plot coordinates: 42.69085, -84.49042). A replicate plot of mixed grasses (see NF02.E2 description). In 2017, we began a three-year longitudinal field study with 10 plants growing in this field plot (2017–2019). The 10 plants were also used to evaluate how detection of SwMV correlates with symptom expression in tillers (2017). In addition to the longitudinal survey, which focused on whole-plant dynamics, we also sub-sampled tillers to characterize virus dynamics at the tiller level. This plot was selected, in part, because in previous growing seasons, SwMV prevalence was near 50%, which meant we would have an ample supply of both infected and uninfected tillers for our investigations here.
NF09.C3	Map of SwMV within plant rhizome network	Located at the W.K. Kellogg Biological Station (KBS) in Hickory Corners, MI (plot coordinates: 42.41904, -85.37040). A 400m ² replicate seeded with switchgrass (cv. "Cave-in-Rock") in 2012.
SNRE	Tiller fitness at different levels of N addition	Located at KBS (plot coordinates: 42.39365, -85.37131). This site was seeded with switchgrass (cv. "Cave-in-Rock") in 2008. The site is divided into replicated sub-plots that receive different applications of ammonium nitrate in spring or early summer. Site maps and additional details are available in the Long-Term Experiments section of the KBS LTER website (https://lter.kbs.msu.edu/research)

Figure S1.1. Dendrogram based on 633 nt coat-protein coding region of switchgrass mosaic virus (SwMV) from 11 experimental plants and two GenBank sequences (an SwMV isolate from IL and a maize rayado fino (MRFV) isolate from Costa Rica). All 22 amplicons from the individual plant used for the rhizome network mapping (CA02) were identical. The four identical amplicons associated with plant A were associated with five RNA extractions from the same tiller, and were used as positive controls.

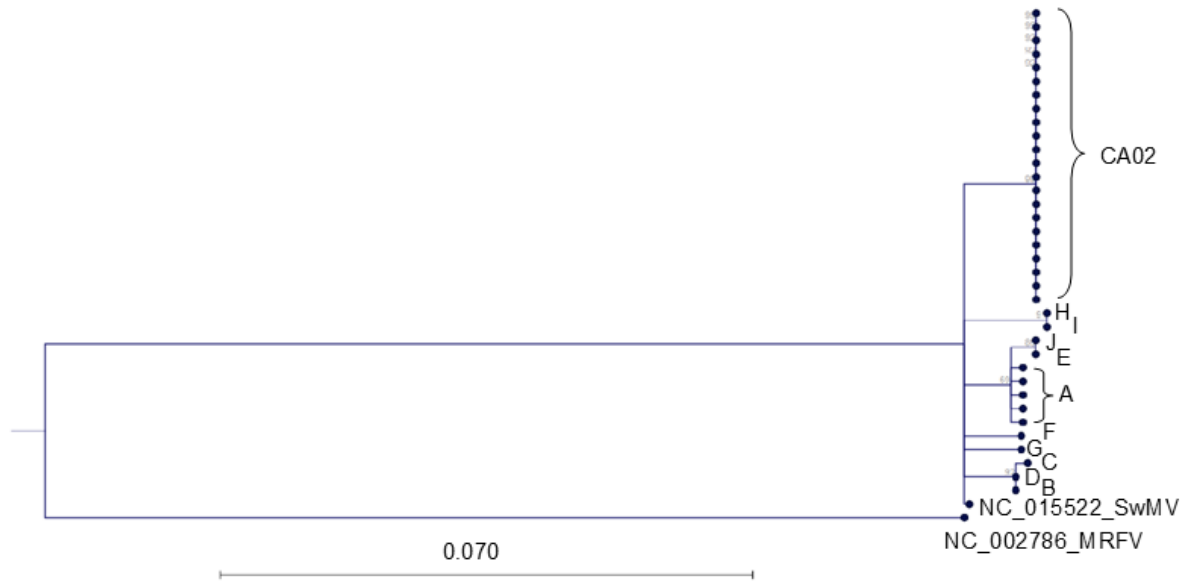
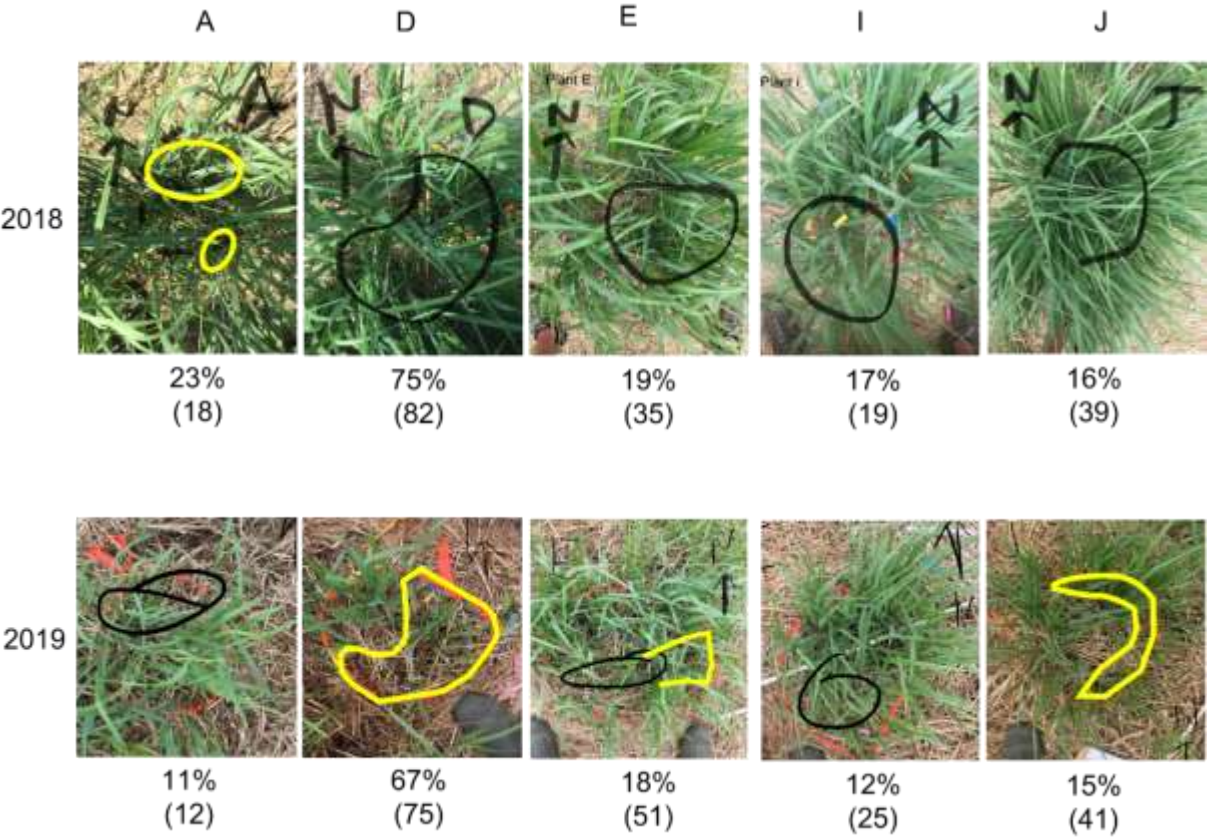


Figure S1.2. Localization of disease in longitudinal plants in is similar in 2018 and 2019. Black marks were added to the picture in the field to denote the locations of diseased tillers (not for A in 2018, which has black dashes highlighting the circumference of the plant), and yellow lines were added later after detailed inspection of the imagery.



Methods S1.1. Extraction, PCR, and Sanger sequencing.

Extraction and RT-PCR

In previous field seasons, we observed consistent and predictable PCR results after using the Zymo kit with stem tissue collected in July and August. However, in processing samples for virus detection in other plant species, we saw that Sigma's Spectrum™ Plant Total RNA kit (MilliporeSigma, Burlington, MA) may have a lower threshold of detection for virus in tissue collected in early summer (Malmstrom, unpub.). Therefore, after initially extractions with the Zymo kit, we ran a second round of extractions using the Sigma kit (25mg dried leaf tissue). In this second round of testing, we included all tillers that had previously been PCR-negative (after Zymo extraction) and a sub-sample of tillers that had been positive. We observed that the Sigma kit was able to pick up a consistent positive signal in tillers that had been positive after Zymo extraction, and there was a clear increase in virus signal in Sigma-extracted RNA from leaf tissue collected in early summer (consistent with kit comparisons with hop plant tissue collected in early summer). In our results, we report tillers as PCR-positive if they were associated with a band after extraction with either kit. Following RNA extraction, we conducted reverse transcription-PCR to detect the 633nt coat protein coding region of SwMV (see below).

The first reaction mixture for reverse transcription included 2µL total RNA, 1.0µL of 10mM dNTP, 0.4µL B88 reverse primer (Agindotan *et al.*, 2013a,b), and 8.6µL mixture of RNase free water and RNase Out (0.5 – 1µL) (all reaction mixtures are per sample volumes). This was incubated at 65°C for five minutes, then chilled for five minutes. A second reaction mixture consisted of 4.0µL 5x first strand buffer, 2.0µL 0.1M dithiothreitol, 1.4µL RNase free water, and 0.1µL Superscript II. Samples were incubated at 42°C for 50 minutes, 70°C for 15 minutes, and then stored at -20°C until conducting PCR.

The PCR reaction mixture consisted of 2.0 μ L 10x PCR buffer, 1.2 μ L 25mM MgCL₂, 0.4 μ L 10mM dNTP mix, 0.8 10uM B88, 0.8 μ L 10uM B89 forward primer (Agindotan *et al.*, 2013b), 0.2 μ L Amplitaq Gold, and 12.6 μ L RNase/DNase free water. We added 18 μ L of the reaction mixture and 2.0 μ L RT product solution (10% RT product diluted in RNase/ DNase free water) before running the following thermocycler program: 94°C for 10 minutes followed by 39 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 45 seconds. The final cycle was 10 minutes at 72°C. PCR product was stored at -20°C before gel electrophoresis.

Sanger sequencing and bioinformatics

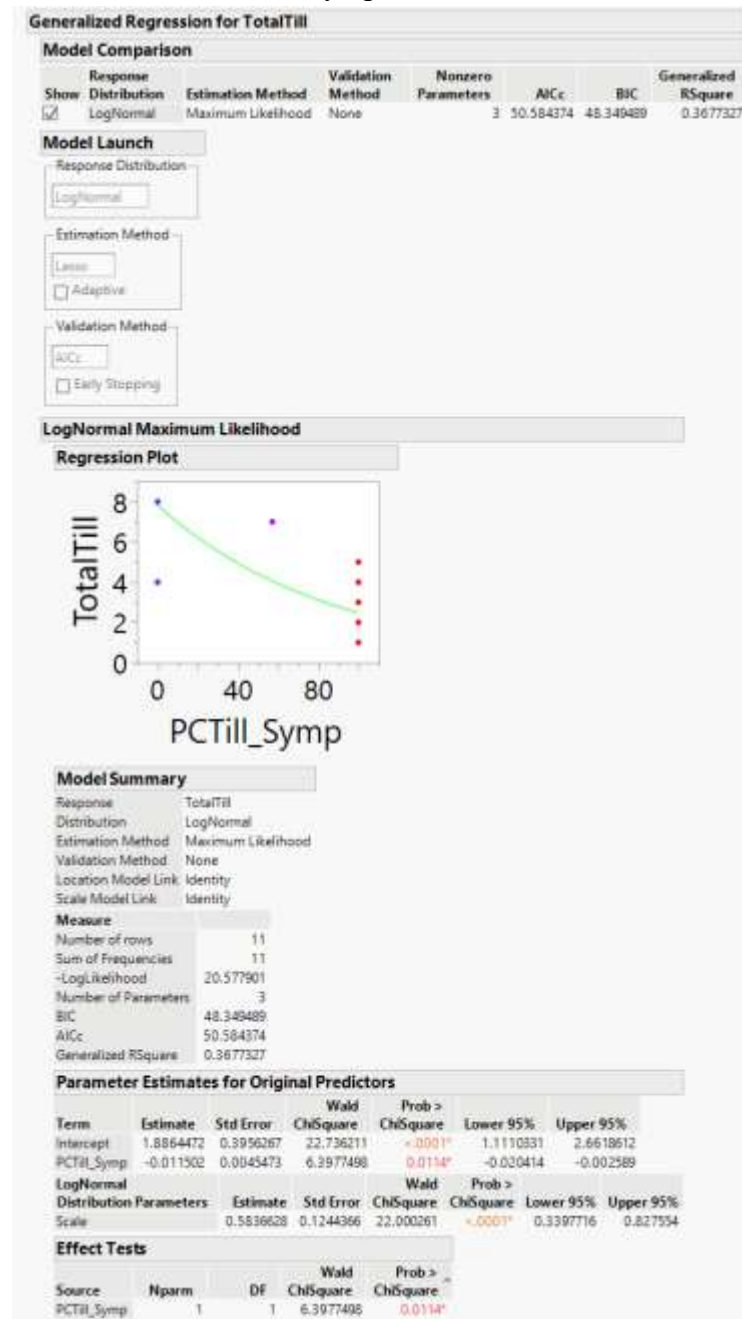
To confirm that PCR products were as expected, we Sanger-sequenced a subset of cDNA associated with tissue samples from each of the ten plants included in our longitudinal study and for 22 samples from the individual plant used for our network map. We prepared samples with a QIAquick® PCR Purification Kit (Qiagen, Venlo, Netherlands), added primers (both forward and reverse for each sample) and submitted them for sequencing at the Michigan State University Research Technology Support Facility.

We used CLC workbench to align forward and reverse strands and confirmed the sequence matched the reference genome. To explore variation in the sequences, we trimmed out the forward and reverse primers, and aligned the sequences to corresponding regions of the SwMV reference genome and maize rayado fino virus, which we used as the outgroup for the ML-based tree.

Methods S1.2. Estimating floret counts based on panicle measures.

We used panicle mass and length measures from panicles collected at the Nitrogen Rate Experiment site (SNRE) to estimate floret count data for two of our three aster models (within-plant tiller fitness and three-year plant fitness). The panicles from the SNRE study were sourced from tillers growing in N addition subplots of 0, 112, and 196 kg ha⁻¹ and therefore represent an extremely wide range of variation in switchgrass panicle sizes and floret production. To evaluate how well panicle length and mass predict floret counts, we plotted raw and transformed data to visually examine linear relationships, and we observed that panicle mass had a strong simple linear relationship with floret production, and that panicle length had a parabolic relationship. We then ran simple linear regressions (without intercepts) with raw and transformed data using panicle length and mass as single fixed effect terms to confirm that mass and length predicted panicle production. We selected models that had the highest r^2 values, and we selected coefficients from these models for slopes in predictive formulas for floret counts based on panicle mass and length. The r^2 value for the model predicting florets based on panicle mass was 0.936. Due to the parabolic relationship between panicle length and florets, raw panicle length explained 79% of variation in floret production, but we found that panicle length had a better linear fit to the square of floret count ($r^2 = 0.935$), and was therefore a more appropriate model for using in floret count estimation. The formulas for estimating floret production were (1) *florets = panicle mass (grams) * 533.5*, and (2) *florets = (panicle length (cm) * 0.5034)²*.

Results S1.1. A generalized regression evaluating the number of tillers of individual plant subsections as a function of symptom extent (% of subsection tillers with symptoms). Subsections with more symptomatic tillers had fewer tillers.



CHAPTER 2: EFFECTS OF A WILD PLANT VIRUS ON HOST FITNESS IN THE CONTEXT OF A FOLIAR FUNGAL PATHOGEN

Introduction

Increasing efforts to sequence and describe the viromes of wild plant communities are revealing a diversity of novel viruses that are widespread in unmanaged ecosystems (Bernardo *et al.*, 2018; Susi *et al.*, 2019; Maclot *et al.*, 2023). In light of the long timescale of wild plant–virus co-evolution (Mushegian *et al.*, 2016)—and the profound effects of known crop viruses on both crop and wild plants (Funayama *et al.*, 2001; Cooper & Jones, 2006; Malmstrom *et al.*, 2017)—wild plant viruses may exert notable but unrecognized influences on wild hosts and wild plant communities. An important next step in virus ecology is to quantify the influence of wild viruses on plants in the field setting, where diverse abiotic and biotic elements can interact to impact plants (Danial and Parlevliet 1995; Devadas *et al.*, 2014; Hoffman *et al.*, 2016).

Panicum virgatum L. (switchgrass) is emerging as a promising model system of a long-lived wild plant in which to examine microbial interactions. A C4 prairie grass, switchgrass is native to a broad geographic range in North America (Zhang *et al.*, 2011), where it is a co-dominant in tallgrass prairie communities. Due to switchgrass' potential as a biofuel feedstock, there has been concerted effort to characterize its broad genetic diversity (Lovell *et al.*, 2021), its morphological, physiological, and performance traits (Casler *et al.*, 2007; Sanderson, 2008; Shrestha *et al.*, 2022), and its interactions with pathogens and other microbes (Hoffman *et al.*, 2016; VanWallendael *et al.*, 2020, 2022).

Here we examine switchgrass interactions with switchgrass mosaic virus (SwMV), a wild virus endemic to the Great Lakes region of the United States (Agindotan *et al.*, 2013a; Malmstrom *et al.*, 2022). SwMV is horizontally transmitted by *Graminella aureovittata*, a North American leafhopper of high-quality tallgrass prairie communities (Wallner *et al.*, 2013) and the

only known vector of SwMV (Agindotan *et al.*, 2013b). Previously, we conducted a longitudinal (multi-year) field study with experimental switchgrass plants with natural SwMV infection (Chapter one, this dissertation), a ‘method 2’ study in plant virus ecology (Malmstrom & Alexander, 2016) that utilizes a controlled field experiment in which infection accumulates from natural sources. We found that SwMV infection can persist for multiple years within individual hosts, causing disease in some but not necessarily all tillers (shoots). Within plants, diseased tillers were less fit than non-diseased ones, but at the whole-plant level individuals with intermediate disease extents (proportion of diseased tillers) were most fit, suggesting that SwMV infection might confer fitness benefits in some circumstances, and/or that fitness effects of SwMV infection may depend on plant genotypes.

Here we probe fitness impacts further with a direct inoculation of experimental plants grown in the field. This study design (a ‘method 3’ study design in Malmstrom & Alexander 2016) involves experimental control at both the host and virus level and represents greater experimental power to quantify effects of virus on hosts. With a wild virus transmitted by a wild specialist leafhopper, such direct inoculations can be technically demanding but provide invaluable evidence of causality. For this work, we clonally propagated experimental switchgrass plants from multiple genets; we mock- and virus-inoculated clonal plants; and grew inoculated individuals in a common field plot exposed to natural field conditions for two years. We evaluated the extent of foliar anthracnose disease caused by a naturally-occurring fungal pathogen, which infected the experimental plants from the environment early in the study, and we evaluated how well anthracnose extent explained variation in biomass and two-year fitness for mock- and virus-inoculated treatment groups.

In the field, SwMV-inoculated plants had less extent of anthracnose disease than mock-inoculated plants, and demonstrated increased tolerance to the fungal pathogen with regard to biomass gain and fitness. We suggest that under natural field conditions, SwMV infection may increase the long-term fitness of hosts under pressure from some fungal pathogens or perhaps other environmental perturbations.

Methods

Plant material

To prepare plant material for mock and virus inoculation treatments, we clonally propagated 83 experimental plants from 23 individual mother plants grown from seeds of the switchgrass cultivar ‘Cave-in-Rock,’ a wildtype upland accession from southern Illinois that is commonly used in biofuel stands and prairie restorations (USDA, 2011). Seeds were germinated in early June 2019 in a research greenhouse at Michigan State University. Seedlings were transplanted into individual pots and greenhouse-grown for several weeks. In mid-August, these mother plants (i.e., ‘genets’) were each separated into 2–3 clones of relatively equal size and maintained in a temperature-controlled greenhouse with 16h daylight (natural illumination with supplemental lights). To ensure success, we used two methods to increase plant size. The most effective approach was hydroponics-based. In this approach, clones from eight genets (A–H, Table 2.1) were grown in aerated hydroponic containers containing half-strength Hoagland’s solution and divided as plant growth permitted. By September 2019, the eight original genets yielded a total 54 plants (4–11 clones per genet), which were transferred to potting media and grown for several more weeks to establish. The 29 remaining individuals (not grown in hydroponics, ‘CA-’ prefix, Table 2.1) were directly planted in potting media. These individuals (2-3 per genet) grew more slowly and were not divided further.

Virus inoculation

We assigned clones from each genet to both mock- and SwMV-inoculation treatment groups. For insurance against the possibility of low inoculation success, we assigned ~twice as many plants to the SwMV inoculation group. Plants from hydroponic-grown genets were more variable in size than the pot-grown ones because they had been divided more frequently. To ensure comparable size class distributions for the two inoculation treatments, we grouped the clones of each hydroponics-grown genet by size class and distributed members of each size class to the mock- and the SwMV-inoculation treatments. Each soil-grown genet included 2–3 similar-size clones; we haphazardly assigned one to the mock-inoculation group and the other one or two to the virus inoculation group. In total, 30 plants were assigned to the mock- and 53 were assigned to the SwMV-inoculation treatment (83 plants total).

At present, inoculation with SwMV can be done only with use of its known leafhopper vector, *Graminella aureovittata*. In our experience, this wild leafhopper can be challenging to maintain in an insectary for long periods. For this study, we therefore collected *Graminella* from wild populations with sweep nets prior to inoculation in October 2019. We swept established switchgrass stands near our field, separated leafhoppers by taxa, and caged them within insect tents in a greenhouse on greenhouse-grown switchgrass. We used *G. aureovittata* for the SwMV inoculation. For the mock-inoculation, we used a mixed colony of its congeners, *G. oquaka* and *G. mohri*. These two species are closely related to *G. aureovittata* and likewise feed on switchgrass, but they are not known to vector SwMV. We did not distinguish between *G. oquaka* and *G. mohri* because species identification requires destructive measures (Malmstrom *et al.*, 2022). We used the congeners for the mock treatment to ensure that any natural SwMV infection carried by the field-collected *G. aureovittata* was not transmitted to the mock-inoculated plants.

For virus acquisition, *G. aureovittata* was caged on SwMV-infected switchgrass to acquire virus and develop capacity to transmit it (Agindotan *et al.*, 2013b), while *G. oquaka/mohri* remained caged on SwMV-free switchgrass. For the inoculation, experimental plants were grouped together and caged with leafhoppers for the appropriate treatment for four weeks. The extended period was chosen to increase inoculation success; marafiviruses like SwMV propagate within the leafhopper vector and several weeks may be needed before transmission is possible (Agindotan *et al.*, 2012, 2013b). Afterwards, we removed the leafhoppers and applied permethrin insecticide with a short residual period (Astro, FMC Corporation) to ensure no live leafhoppers remained on plants.

SwMV overwinters in switchgrass rhizomes, and diseased tillers generally emerge with symptoms in spring (Chapter one, this dissertation). Therefore, to establish SwMV-inoculated plants with persistent infections in above- and below-ground tissue, we induced cold-season senescence in plants after the inoculation period. At this time, slight foliar symptoms of SwMV disease had already begun to emerge in new leaves of several SwMV-inoculated plants. To induce senescence, we first cold-primed plants overnight in growth chambers set at 11°C and then exposed them to cycles of 10 hours of light at 12°C and 14 hours of darkness at 5°C. Over five weeks, temperatures were gradually lowered to -3°C. After this simulated short winter, plants were transplanted into new pots and moved into a heated greenhouse with 16h supplemental lighting to renew growth. Of the 83 plants inoculated, 57 (63%) re-emerged (Table 2.1) and were treated with Mainspring, a systemic pesticide, to provide short-term (~2 month residual) control of common greenhouse pests (e.g., thrips). Emergence rates did not differ between inoculation treatments: 69.8% (37/53) of SwMV-inoculated plants and 67.7% (20/30) of

mock-inoculated plants re-emerged ($z=0.297$, $p = 0.767$; significant difference between two proportions test). Plants were maintained in the greenhouse for five months until field planting.

Field site establishment

In early June 2020, all 57 inoculated plants that emerged (37 virus-inoculated, 20 mock inoculated) were planted into a field site in Grand Rapids, MI, within the natural range of switchgrass (Lovell *et al.*, 2021). This location was chosen to facilitate childcare while conducting fieldwork during the global SARS-CoV-2 pandemic. Prior to establishment, the site had been lightly used as a vegetable garden for several years, with semi-annual applications of shredded leaves/bark and a few limited applications of fertilizer. Soils at the site consist of 15–20cm topsoil (5% organic matter; 14% silt, 14% clay, and 72% sand) with sandy sub-soils (~90% sand). Prior to planting, we established a honeycomb plot layout of staggered rows (0.67m apart). Plants were randomly assigned to positions within rows (0.5m apart), and buffer plants were planted around the plot perimeter. In the first growing season, we watered the plot during several hot weeks with little or no precipitation.

Symptom development and infection persistence

Among the 37 emergent SwMV-inoculated plants, ten plants (27%) developed symptomatic SwMV disease and expressed disease symptoms in every tiller for the entire two-year period; we refer to these individuals as ‘V-SYMP’. Twenty-two of the 37 (59.5%) never developed symptoms and were asymptomatic throughout the field experiment (‘V-ASYMP’ plants). The remaining 5 (13.5%) initially showed faint signs of SwMV disease, but never developed clear symptoms and their later growth was not diseased; these were also categorized as V-ASYMP. None of the 20 emergent mock-inoculated plants (MOCK) showed signs of SwMV disease at any time. We incorporated these three categories—MOCK, V-ASYMP, V-

SYMP—as a fixed factor in statistical models (*‘Inoculation Group’*) of biomass, anthracnose, and two-year fitness (described below).

Evaluating persistence of infection

All 57 field plants survived for two growing seasons—the duration of the study. At the end of the second growing season, we used RT-PCR to test for virus infection in the youngest fully expanded leaves of tillers that were randomly selected from each plant. We extracted total RNA from dried tissue samples with a Spectrum™ Plant Total RNA kit (MilliporeSigma, Burlington, MA), and conducted RT-PCR with SwMV specific primers, as described in Chapter one, Methods S1 (this dissertation). This testing found that all ten V-SYMP individuals were strongly PCR-positive (Fig. 2.1a), suggesting that infection persisted throughout the two-year study. Three of 19 V-ASYMP plants (15.8%) also tested positive, but gave only faint bands (Fig. 2.1b). In addition, four of 20 MOCK plants (20%) tested positive with barely visible bands, likely due to infections acquired through naturally extant vectors in the second year.

Quantifying anthracnose extent

By mid-summer of year one (2020), most experimental plants exhibited large (<5 mm) necrotic lesions characteristic of naturally-occurring anthracnose. This is a fungal disease caused by *Colletotrichum* spp (Crouch *et al.*, 2009), which is known as an important hemibiotrophic pathogen in maize and other crops (Oliveira Silva *et al.*, 2022). We therefore characterized anthracnose extent as a potentially important covariate and element for models of plant traits and fitness. We assessed anthracnose extent in late summer of both years for every plant as a percent (to the nearest 5%) of leaves with more than one necrotic lesion. We modeled variation in anthracnose extent by running linear models in R with *Inoculation Group* (MOCK, V-ASYMP, and V-SYMP) as a fixed effect and used likelihood ratio tests to compare it with a null model.

We used the *emmeans* package to estimate group means and evaluate pairwise differences among the three *Inoculation Groups*.

Evaluating biomass and fitness as a function of anthracnose extent and inoculation group

To quantify aboveground biomass, we clipped tillers near the soil surface at the end of each growing season (late September), and dried the aboveground biomass from each plant in labeled paper bags. To parameterize aster models of two-year fitness, we counted all tillers and panicles for each plant at the end of both growing seasons.

Modeling biomass

We modeled biomass as a function of *Inoculation Group* and *Anthracnose* using linear models in R. We constructed two single-term fixed effect models (*Inoculation Group*, *Anthracnose*) and two models with both fixed effects, one of which included an *Inoculation Group*Anthracnose* interaction term. We conducted model comparisons among models and we selected the best performing model based on likelihood ratio tests and AICc values (*AICc*tab in ‘*bbmle*’ package). We used the *summary* function in R to determine if the slopes of each *Inoculation Group* were significantly different than zero (*Inoculation Group*Anthracnose* model only).

Multi-year fitness using the aster framework

The *aster* package in R is a powerful tool for life-history analysis that enables estimation of fitness across the disparate and interdependent distributions associated with perennial plant life stages (Geyer *et al.*, 2007). The statistical structure of an aster model is composed of graphical nodes that correspond to life stages (e.g., early survival; number of tillers). We constructed an aster model to evaluate the two-year fitness of the 57 plants that established in the field site. This model is based on (i) tiller counts in years one and two (negative binomial); and

(ii) panicle counts in years one and two (Bernoulli; modeled as a proportion of tillers that become reproductive and conditioned on tiller counts in each year). The final estimate of plant fitness, which is integrated over two years, incorporates any differences among plants/treatments in tiller production and in the likelihood that tillers become reproductive (i.e., develop panicles).

We evaluated two-year fitness as a function of two fixed effects: *Inoculation Group* and anthracnose extent ('*Anthracnose*'). We first fit a series of models with all combinations of the two fixed effects: 1) *Inoculation Group*; 2) *Anthracnose*; 3) *Inoculation Group* + *Anthracnose*; and 4) *Inoculation Group***Anthracnose*. We evaluated fit of Poisson and negative binomial distributions for panicle counts, and we estimated the shape parameter for the negative binomial distribution as described in Ryskamp and Geyer (2023). We used likelihood ratio tests to compare models (including a null model) and to select the model that best explained variation in two-year fitness. After finalizing the aster model, we used aster-estimated fitness values for individual plants to explore potential fitness differences among *Inoculation Groups* at different levels of anthracnose extent (Tukey pairwise comparisons), and we used polynomial regressions to explore the relationship between anthracnose and fitness for each *Inoculation Group*.

Results

Anthracnose extent

Among the three inoculation groups, there were significant differences among the three *Inoculation Groups* ($F_{2, 54} = 5.48$, $p = 0.007$) (Fig. 2.2). Compared to MOCK plants, V-SYMP plants had 61% less anthracnose ($t = 3.09$, $p = 0.0086$; Tukey's HSD) and V-ASYMP plants had 36% less anthracnose ($t = 2.42$, $p = 0.049$) (Fig. 2.2). Differences between V-SYMP and V-ASYMP plants were not significant ($t = 1.31$, $p = 0.395$).

Dry mass

In models of dry mass, all models that included *Anthracnose* and *Inoculation Group* were significant relative to the null model, but the model with the *Anthracnose*Inoculation Group* interaction term (i.e., ‘interaction model’) best explained variation in two-year dry mass (Table 2.2), indicating that the relationship between anthracnose extent and biomass varied by *Inoculation Group* ($F_{5, 51} = 5.72$, $p = 0.039$).

For MOCK plants, the estimated two-year biomass (interaction model) declined steeply from the lowest anthracnose levels to the highest, from 510 grams (± 59.0 , 1 SE) to 75.4g (± 59.8) ($t = 4.31$, $p < 0.0001$, t-test for non-zero slope) (Fig.2.3). The biomass of V-ASYMP plants also fell with increasing anthracnose extent, ($t = 2.34$, $p = 0.023$), but to a lesser degree, from an estimated 441.2g (± 45.0) at low anthracnose levels to 368.6g (± 65.6) at the highest levels. For V-SYMP plants, the estimated two-year biomass rose slightly from low to high anthracnose extents, from 306.4g (± 58.1) to 322.8g (± 59.6), but variation in anthracnose extent did not significantly explain variation in biomass for V-SYMP plants ($t = 1.74$, $p = 0.0873$).

Two-year fitness

All models with *Inoculation Group* and/or *Anthracnose* included as fixed effect terms were significant in likelihood ratio tests against the null model ($\chi^2 = 569$, $df = 2$, $p < 0.00001$). However, the aster model that explained the most variation included *Anthracnose*Inoculation Group* interactions, indicating that the relationship between two-year fitness and anthracnose extent depends on inoculation group ($\chi^2 \geq 251$, $df \geq 4$, $p < 0.00001$). Overall, the fitness of SwMV-inoculated plants (both V-SYMP and ASYMP plants) was similar across anthracnose levels ($t = 0.57$, $df = 54$, $p = 0.835$, Tukey’s HSD), and in linear regressions, anthracnose extent did not explain variation in fitness for either V-ASYMP ($F_{1,25} = 0.244$, $p = 0.6258$) or V-SYMP

plants ($F_{1,8} = 0.63$, $p = 0.452$). Fitness for MOCK plants however, declined with increased anthracnose extent ($F_{1,18} = 7.23$, $p = 0.015$, simple linear regression), and was best modeled as cubic a regression (Fig. 2.4, dark green line) that declined steeply from ~275 panicles at low anthracnose levels (<5%) to an average of ~60 panicles at $\geq 20\%$ anthracnose ($F_{2,16} = 43.1$, $p < 0.0001$).

After plotting the aster-predicted fitness for all plants, we observed that the cubic regression line fit for the MOCK-plant fitness fell below the average fitness of SwMV-inoculated plants (~100 panicles) near a mean anthracnose extent of 20%. We therefore conducted Tukey pairwise comparisons of the three *Inoculation Groups* to compare fitness estimates above ($\geq 20\%$) and below this threshold (red line, Fig. 2.4). At low levels of anthracnose (<20%), mean fitness of MOCK plants (195 panicles) was more than two-times that of SwMV-inoculated plants (both V-ASYMP and V-SYMP) ($t \geq 3.77$, $df = 18$, $p \leq 0.0014$). The majority of MOCK plants (75%), however, had higher levels of anthracnose (20–100%). Within this range, the mean two-year fitness for MOCK plants was far smaller, at 64 panicles (per plant). V-ASYMP and V-SYMP plants, in contrast, had ~65% greater fitness than MOCK plants at this range of anthracnose extent ($t \geq 4.98$, $df = 33$, $p \leq 0.00002$).

Discussion

This study represents a unique effort to quantify the multi-year impacts of a wild plant virus in experimentally inoculated wild plant hosts. SwMV-inoculated plants had less extensive anthracnose disease overall and greater two-year fitness than mock-inoculated plants at moderate–severe anthracnose levels ($\geq 20\%$). Our results suggest that SwMV confers benefits for plants in the context of foliar fungal pathogens like anthracnose. Our results also highlight that context-dependent effects of wild plant – virus interactions may serve as a means by which moderately antagonistic viruses coexist and coevolve with perennial host plants.

Virus influence on plant interactions with foliar fungal pathogens

Relative a ‘method 2’ study in virus ecology, which allows natural infection to accumulate within an experimental field stand, a ‘method 3’ experiment enables powerful quantification of the effects of inoculated virus in experimental plants relative a mock-inoculated control group (Malmstrom & Alexander, 2016). In our field plot, we planted clonally propagated mock- and SwMV-inoculated plants, which allowed us to control for some of the stochasticity that is introduced due to genotype-specific variation. Overall, our results suggest that virus-inoculated plants not only have less extensive anthracnose than MOCK plants, but that they also have greater tolerance to anthracnose disease and higher relative fitness when the extent of anthracnose disease reaches levels 20% and higher. These results suggest that SwMV may confer increased tolerance to anthracnose disease, and/or that SwMV may have contextually beneficial impacts in the context of moderate to severe foliar fungal pathogen outbreaks.

Interestingly, V-ASYMP plants performed similarly to V-SYMP plants in many of our results. Despite having intermediate levels of anthracnose extent, the relationship between biomass and anthracnose extent for V-ASYMP plants was significantly less negative than that of MOCK plants. Additionally, at moderate to severe levels of anthracnose, the two-year fitness of V-ASYMP and V-SYMP plants were similar, and significantly higher than MOCK plants. The performance of V-ASYMP was interesting because they had the lowest proportion of PCR-positivity (10%) among treatment groups at the end of the experiment, which suggests that any infection they acquired (as evidenced by initial symptoms) tended not to persist. Perhaps the prolonged exposure to infectious vectors during the inoculation stage resulted in an immune priming that conferred some protections against anthracnose (Desmedt *et al.*, 2021).

Conclusions and future directions

In this study, we conducted a controlled inoculation with clonally propagated plants. Surprisingly, SwMV-inoculated plants overall experienced much less anthracnose disease from the environment than did mock-inoculated plants. Among individuals that contracted moderate to high levels of anthracnose ($\geq 20\%$), those in the SwMV-inoculation treatment were more fit than those that were mock-inoculated, suggesting the possibility that SwMV may confer benefits to switchgrass in the context of environmental anthracnose pressure. Future studies are needed to explore interactions between SwMV infection and fungal pathogens in switchgrass more broadly and to examine their potential mechanistic bases.

APPENDIX

Tables and figures

Table 2.1. Genets and associated clones in mock- and SwMV-inoculation treatments. All mock-inoculated plants remained free of signs of SwMV disease throughout the experiment. Of the 37 plants inoculated with SwMV, 10 showed distinct SwMV symptoms that remained evident throughout the study. Five others initially showed faint but distinct signs of disease that faded early in the first growing season. The remaining 22 plants in the inoculation treatment never showed any SwMV disease symptoms. ‘Surv’ in the table indicates the number of plants that survived until the establishment of the field site. ‘PCR+’ refers to results of PCR testing at the end of the two-year experiment (in 2021).

Symptoms in SwMV-inoculated clones	Genet ID	Initial clones (n)	SwMV-inoc clones				Mock-inoc clones			Notes about symptom expression in SwMV-inoc clones
			n	Surv (n)	Num. symptomatic (symp desc)	PCR + (n)	n	Surv (n)	PCR+ (n)	
Genets with persistent disease	G	7	5	5	5 (persistent)	5	2	2	0	G02 had faint symptoms at first H03 (SwMV trt) w/ mild symptoms at first
	H	5	3	1	1 (persistent)	1	2	1	1	
	D	4	2	2	2 (persistent)	2	2	2	0	
	C	5	3	1	1 (persistent)	1	2	0	0	
	CA06	2	1	1	1 (persistent)	1	1	0	0	CA06A had mild symptoms at first
	Sub:	23	14	10	10	10	9	5	1	
Genets with transient signs of disease	A	11	8	4	2 (faint/distinct & transient)	1	3	2	1	A02 = distinct; A10 = faint; neither was PCR+ in 2021 F06 & F08 = faint CA01A had distinct symptoms that faded
	F	10	6	4	2 (faint & transient)	0	4	3	0	
	CA01	3	2	2	1 (distinct & transient)	0	1	1	0	
	Sub:	24	16	10	5	1	8	6	1	

(Table 2.1, cont'd)

Symptoms in SwMV-inoculated clones	Genet ID	Initial clones (n)	SwMV-inoc clones				Mock-inoc clones			Notes about symptom expression in SwMV-inoc clones
			n	Surv (n)	Num. symptomatic (symp desc)	PCR + (n)	n	Surv (n)	PCR+ (n)	
Genets with no signs of disease	<i>B</i>	7	5	3	0	0	2	1	0	
	<i>E</i>	6	4	1	0	0	2	2	0	
	CA03	3	2	2	0	0	1	1	0	
	CA05	2	1	1	0	0	1	1	0	
	CA10	2	1	1	0	0	1	1	0	
	CA14	2	1	1	0	0	1	1	0	
	CA15	2	1	1	0	1	1	1	1	
	CA11	2	1	0	0	0	1	1	1	
	CA02	2	1	1	0	0	1	0	0	
	CA04	2	1	1	0	0	1	0	0	
	CA13	2	1	1	0	0	1	0	0	
	CA07	1	1	1	0	1	0	0	0	
	CA08	1	1	1	0	1	0	0	0	
	CA09	1	1	1	0	0	0	0	0	
	CA12	1	1	1	0	0	0	0	0	
	Sub:	36	23	17	0	3	13	9	2	
	Total:	83	53	37	15	14	30	20	4	

Table 2.2. Variation in two-year biomass is best explained by *Anthracnose*Inoculation Group* interactions. The model with the interaction term (model ‘3’) had the lowest AICc value, and was significant in likelihood ratio tests in against the null model and test models (models 1 and 2). In the ‘ Δ AICc’ column, the AICc values for each model are listed relative model 3’s base AICc value of 735. The shaded columns list the results of likelihood ratio tests. The ‘LRT’ column specifies the Model IDs of the likelihood ratio test comparison (relative the model represented in each row).

Model ID	Model terms	Δ AICc	LRT	χ^2	df	p-value
0	Null	+13.3	NA	NA	NA	NA
1	Inoc. Group	+10.7	vs. 0	7.20	2	0.0273
2	Inoc. Group + Anthracnose	+2.1	vs. 1	10.9	1	0.00095
3	Inoc. Group x Anthracnose	0 (735)	vs. 2	7.25	2	0.0267

Figure 2.1. In RT-PCR gels, tissue samples from persistently infected plants were associated with strong, bright gel bands (asterisks). (a) Most tissue samples from non-diseased tillers did not produce bands in gels (white brackets). (b) Gel bands for V-ASYMP and MOCK tissue samples were less bright, especially for MOCKs, which were almost non-detectable for two of the four bands (white circles). The brightness of these gels was adjusted slightly for visibility (+~20%). Negative and positive controls are indicated with a minus and a plus sign, respectively.

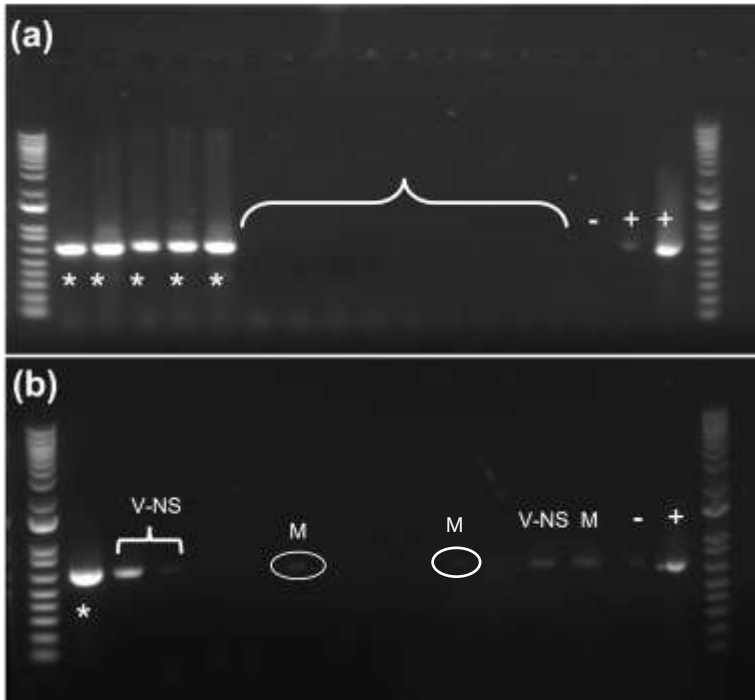


Figure 2.2. Compared to MOCK plants (dark green), V-ASYMP plants (light green) had 36% less anthracnose extent, and V-SYMP plants (yellow) 61% less. The best performing model did not account for Genet. Bars represent model estimated means (*eemeans* package, R) from a model with *Inoculation Group* as a single fixed effect. Vertical lines represent ± 1 SE.

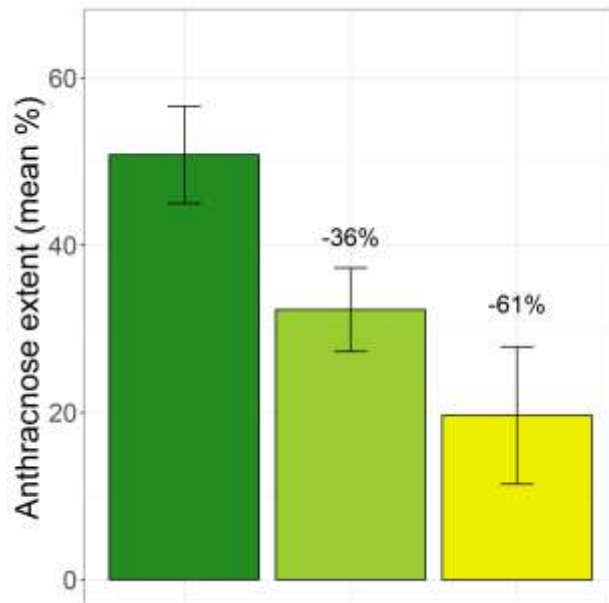


Figure 2.3. The relationship between anthracnose extent and biomass depends on *Inoculation Group* (MOCK, V-ASYMP, or V-SYMP). From low to high anthracnose extent, the decline in biomass was most negative for MOCK plants (green circles). The decline in biomass for V-ASYMP (light green diamonds) plants was less negative, and for V-SYMP plants (yellow squares), anthracnose extent did not explain variation in total biomass. The shapes in the graph represent raw data for plants, and the model estimates are represented as regression lines $\pm$ 1 SE (colored areas around lines) that correspond to each *Inoculation Group*.

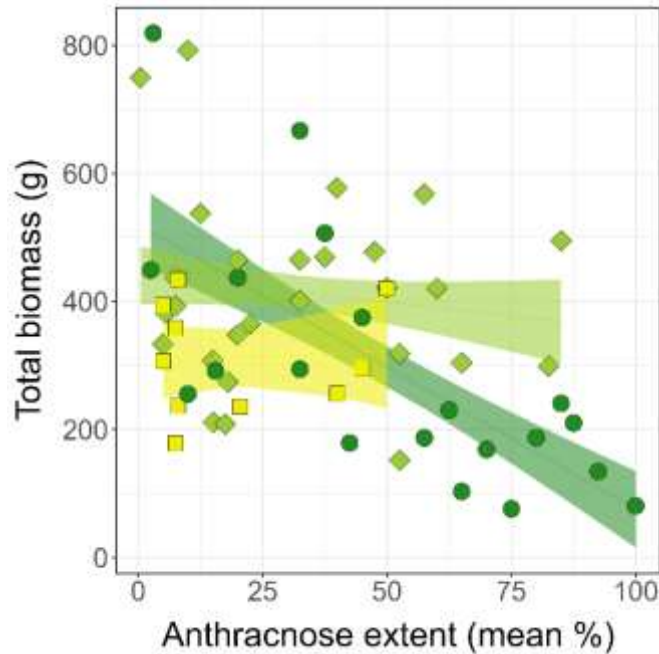
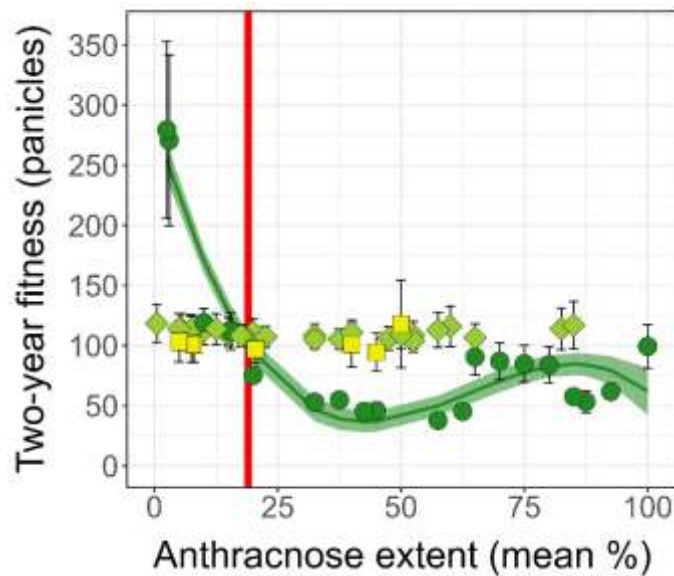


Figure 2.4. Two-year fitness is a function of *Anthracnose*Inoculation Group* interactions. The fitness of MOCK plants (dark green circles) falls from high levels at low anthracnose extents, and dips below the mean fitness of SwMV-inoculated plants (both V-ASYMP and V-SYMP) at 20% anthracnose extent (red line). The dark green line represents a cubic regression fit to the aster-estimated fitness of MOCK plants (± 1 SE). The two-year fitness estimates of V-ASYMP (light green diamonds) and V-SYMP (yellow squares) plants were similar, and did not significantly vary with anthracnose extent. The black lines above/below shapes represent ± 1 SE.



CHAPTER 3: EXPRESSION OF WILD VIRUS DISEASE IN SWITCHGRASS (*PANICUM VIRGATUM* L.) VARIES BY ECOTYPE, SUBPOPULATION AND IS REGULATED BY A DIVERSITY OF DISTRIBUTED GENES

Introduction

Land plants invoke multi-layered defense strategies that include structural defenses (Hanley *et al.*, 2007), basal defenses triggered by general pathogen characteristics (Zhang & Zhou, 2010), and defense pathways triggered by translated resistance (R) gene products that recognize specific pathogen strains or races (Yang *et al.*, 2013). In crop protection, introgression of single R genes and even resistance gene pyramids often do not provide durable resistance, which has spurred notable study of R gene families and their rapid evolution (Yang *et al.*, 2013; Stam & McDonald, 2018). Investigation to-date has generally found greater complexity of defenses in wild plants than in domesticated ones (Fernandez *et al.*, 2021; He *et al.*, 2022), but the full suite of wild plant defenses has yet to be mapped. Its exploration will likely yield key understanding about how plants navigate complex environmental challenges, as well as new strategies for crop protection.

Viruses comprise the majority of emerging plant disease threats and merit significant attention (Anderson *et al.*, 2004). In nature, wild viruses are abundant in plant communities (Muthukumar *et al.*, 2009; Bernardo *et al.*, 2018; Maclot *et al.*, 2023), having infected terrestrial plants since early in their evolution (Mushegian *et al.*, 2016). Some percentage of these wild virus taxa are benign to beneficial (Gibbs, 1980; Márquez *et al.*, 2007); others are antagonistic but may increase host fitness in particular circumstances (Ryskamp, Chapters one and two, this dissertation). Here we seek to expand understanding of the genetic architecture of viral defenses in wild plants through multi-year study of a disease-causing wild virus in a genetically diverse native grass.

Our focal host plant species is *Panicum virgatum* L. (switchgrass), a perennial C4 prairie grass that ranges throughout North America from northern Mexico to southern Canada (Casler *et al.*, 2007; Morris *et al.*, 2011). This native grass offers unparalleled opportunities for examining wild plant–virus interactions because its genome has been extensively sequenced to support its development as a potential bioenergy feedstock (Casler, 2012; Lowry *et al.*, 2015, 2019; Lovell *et al.*, 2021).

Switchgrass exhibits much phenotypic and genotypic variation throughout its geographic range, but variation in viral disease susceptibility has yet to receive broad attention. Historically, phenotypic variation in switchgrass was characterized by two contrasting morphologically based ecotypes: upland and lowland, with the former more common in the USA north and the latter in the warmer south where adaptations to periodic flooding allow it to inhabit riparian zones (Porter, 1966; Cortese *et al.*, 2010; Lowry *et al.*, 2014). More recently, Lovell *et al.* (2021) identified a third ‘coastal’ ecotype defined by upland coloration and stature and lowland leaf and tiller traits. In their study of more than 600 genotypes, the authors further distinguished three SNP-based subpopulations that correspond to US regional geography: Midwest, Atlantic, and Gulf. A first aim of our study was to investigate differences among these ecologically interesting ecotypes and subpopulations in susceptibility to viral disease.

Most of what is known so far about switchgrass disease resistance has come from studies of rust (*Puccinia* spp.) (Uppalapati *et al.*, 2013; Serba *et al.*, 2015; Kenaley *et al.*, 2018; VanWallendael *et al.*, 2020) and other fungal diseases, including anthracnose (*Colletotrichum* spp.) (Crouch *et al.*, 2009) and smut (Kenaley *et al.*, 2019). These studies have identified significant variation in fungal resistance but found that lowland ecotypes, more strongly represented in warmer regions, typically tend to be more fungal-resistant than upland ones

(Uppalapati *et al.*, 2013; Hoffman *et al.*, 2016). In genetic analyses of a four-way mapping population grown at multiple sites, VanWallendael *et al.* (2020) found evidence that the nature of effective resistance might differ by environment; they identified two large-effect quantitative trait loci associated with rust resistance at northern sites in the United States, and several small-effect resistance loci in southern sites.

Viral disease in switchgrass can be as noteworthy as fungal disease in the field but has been less recognized or examined. As a result, little is known either about broad patterns of viral susceptibility within and among switchgrass ecotypes and subpopulations or about the genes that regulate switchgrass viral defenses. It is clear, however, that at least several viruses can reach high prevalence in switchgrass stands (Stewart *et al.*, 2015; Malmstrom *et al.*, 2022) and reduce switchgrass performance (Alexander *et al.*, 2017; Malmstrom *et al.*, 2017). Moreover, there is evidence that selection of switchgrass for increased productivity may increase its susceptibility to viral infection (Schrotenboer *et al.*, 2011), highlighting the importance of understanding the genetic mechanisms responsible. One small clue from previous field studies is that the southern lowland ecotype ‘Kanlow’ from Oklahoma has appeared less susceptible (Stewart *et al.*, 2015) and/or more tolerant to viral infection (Alexander *et al.*, 2017) than several more northerly upland accessions (e.g., ‘Shawnee’). Extrapolating from this evidence and from the studies of fungal disease, we predicted that lowland ecotypes in general might be the least susceptible to viral disease, perhaps because so many of them originate from warmer regions where disease pressure might be greatest.

To examine broad patterns of virus susceptibility and associated genes in switchgrass, we worked with the same tetraploid switchgrass diversity panel described in Lovell *et al.* (2021), a common garden planting of >500 genotypes from across the US that is used for genome-wide

association studies (GWAS). We quantified the susceptibility of each accession to viral disease at a site with notable virus pressure, and then used GWAS to identify SNPs and candidate genes associated with two viral disease phenotypes (severity and extent). We focused on naturally occurring disease caused by switchgrass mosaic virus (SwMV), a wild virus endemic to the Great Lakes region of the United States (Agindotan *et al.*, 2012; Malmstrom *et al.*, 2022, Malmstrom *et al.* unpub.). This wild virus is transmitted by *Graminella aureovitta* (Agindotan *et al.*, 2013b), a North American leafhopper of high-quality grasslands (Wallner *et al.*, 2013). SwMV prevalence can reach high levels in switchgrass, in some circumstances maintaining prevalence of >50% for multiple years (Malmstrom *et al.*, 2022, Malmstrom unpub.).

We found significant variation in switchgrass susceptibility to SwMV disease, with the Atlantic subpopulation markedly least susceptible, and upland ecotypes from the Midwest and lowland ecotypes from the Gulf population most susceptible. The 12 GWAS we conducted identified 1,100 significant single nucleotide polymorphisms (SNPs) associated with disease expression from which we identified 67 top candidates representing the most significant SNPs within linkage disequilibrium blocks in multiple GWAS; these were associated with 143 switchgrass genes. Only a minority of these genes (4.4%) belong to recognizable R gene families, and 21% (30/143) lack functional annotations altogether due to lack of homology in the rice and *Arabidopsis* model systems. Gene ontology (GO) enrichment identified pathways related to immune responses and signaling, as well as to metabolism, development, and abiotic stress responses. These findings demonstrate the breadth of viral interactions with wild plants and suggest opportunities for switchgrass selection for disease resistance.

Methods

Location and overall approach

We rated SwMV disease severity and extent in the switchgrass diversity panel planting at Michigan State University's W.K. Kellogg Biological Field Station in Hickory Corners, Michigan (42.4189, -85.3714). The panel was established in early summer 2018 and contains plants spaced 1.56m apart in a honey-comb design (plot establishment is described in greater detail in Lowry et al. (2019) and Lovell et al. (2021). The panel includes 595 tetraploid individuals that represent 512 re-sequenced genotypes used for GWAS analysis. AP13, from which the switchgrass reference genome is derived, is represented by 56 clonal replicates distributed throughout the panel planting as a check plant. A limited number of other genotypes are represented by several clonal replicates (27 in total).

The diverse genotypes in the panel represent regionally adapted accessions collected from localities throughout the native range of switchgrass in North America. They comprise three SNP-based subpopulations (Atlantic, Gulf, and Midwest) and the three morphology-based ecotypes (coastal, lowland, and upland) (Lovell *et al.*, 2021). Each subpopulation contains one or two dominant ecotypes, giving five well-represented subpopulation-ecotype combinations (subgroups): Atlantic-coastal, Atlantic-upland, Gulf-coastal, Gulf-lowland, and Midwest-upland (Table 3.1). For more information about this diversity panel, see Lovell et al. (2021).

Scoring SwMV disease phenotypes

At this site, SwMV infection accumulated naturally over time in the diversity panel from environmental virus pressure. It was the single most dominant viral disease at this site. We evaluated SwMV disease expression in individual plants for three years (2020–2022) in both early summer (May 26–June 8) and mid-summer (July 3–16), beginning two years after panel

establishment. We rated disease severity as the maximum foliar symptom expression found on any individual leaf and rated it from zero (no symptoms) to four (severe symptoms, >50% of leaf) (Fig. 3.1). Scores of 2 or above represent reliable indications of SwMV disease, as validated with molecular diagnostics (Chapter one, this dissertation). Because SwMV disease can be distributed non-uniformly throughout a plant, sometimes localized to only portions, we also quantified disease extent—which we defined as the proportion of tillers with severity score greater than two—and scored this metric on an ordinal scale from zero to 10. Taken together, the disease severity and extent measures give a holistic view of disease expression in each individual.

Heritability

We estimated narrow-sense heritability (h^2) with the *heritability* package in R. We first calculated a genetic relatedness matrix based on 1M randomly selected SNPs, and we used the *marker_h2* function to estimate h^2 .

Disease dynamics over time

We evaluated simple panel-wide disease prevalence as the percentage of individuals ($n = 595$) with disease severity score ≥ 2 . For evaluation of disease severity and extent, however, we corrected for genotype replication to ensure that AP13 and other replicated genotypes were not over-represented in statistical models. To do this, we calculated the mean severity and extent for each genotype (at each time point), and then conducted analysis on these means.

To evaluate disease susceptibility by ecotype, subpopulation, or subgroup across the multi-year study, we conducted repeated measures analysis by fitting individual models with a fixed-effect interaction term with each group affiliation and time (e.g., *Ecotype x Time*), using the *nlme* package in R. We included *Genotype ID* as a random effect to represent that measures

were repeated on genotypes, and we used a first-order autoregressive variance-covariance structure. To test the significance of each model, we conducted likelihood ratio tests (*anova* function) against null models that included *Time* as the sole fixed effect. We then evaluated how well ecotype, subpopulation, and ecological subgroup explained variation in disease severity and extent by comparing AICc values and by conducting log-likelihood based comparisons among the three full models for each disease phenotype. We used the *eemeans* R package to generate model estimates and conduct Tukey-corrected pairwise comparisons of the groups represented in models at each time point.

GWAS analyses and candidate genes

To identify genes potentially associated with virus susceptibility, we conducted genome-wide association analyses for disease severity and extent at each time point (2 seasons x three years). For genotypes with more than one clonal replicate, we used median phenotype scores at the genotype level. All GWAS were conducted with the *switchgrassGWAS* package in R (R Core Team, 2021; Lovell *et al.*, 2021), which performs fast statistical analysis with large SNP arrays encoded as matrices. This package implements the current best practices in human genetics for principal component analysis of population genetic data (Privé *et al.*, 2020). Genome assembly and SNP calling is described extensively by Lovell *et al.* (2021). Only SNPs with <20% missing data and minor allele frequencies >0.05 in the 512 genotypes were used in univariate GWAS, resulting in 12.9M SNPs retained for the analysis. The *pvddiv_standard_gwas* function in *switchgrassGWAS* was used to run all GWAS. As part of this function, we used singular value decomposition (SVD) on all 12.9M SNPs in all genotypes to create 15 genetic principal components (PCs) for population structure correction. Once a sufficient number of PCs are included to capture true population structure effects, GWAS results are not sensitive to the

number of PCs used (Price *et al.*, 2006). To choose the number of PCs that best controlled for population structure, we ran linear regressions for each phenotype including 0 to 15 PCs as covariates, then selected the smallest set of PCs that reduced λ_{GC} , the genomic inflation factor, below 1.05. All 12 of our GWAS had $\lambda_{GC} < 1.035$.

To identify significant SNPs from each GWAS, we used the Bonferroni-corrected $-\log_{10}$ P-value ($\alpha = 0.05$) threshold of 8.41. To curate a conservative list of candidate SNPs, we used the *BigSnpR* R package to identify the most significant SNPs within linkage-disequilibrium blocks ($r^2 = 0.20$). To further focus our analysis, we then filtered this SNP-pruned set to include only SNPs identified in two or more of our 12 GWAS ('top candidate SNPs').

To identify candidate genes associated with these SNPs, we used the *pdiv_table_topsnps* function in the *switchgrassGWAS* package to identify genes within 20kbp window (± 10 kbp) association of each top candidate SNP. This is a conservative distance consistent with $<50\%$ linkage disequilibrium decay (Grabowski *et al.*, 2017; Lovell *et al.*, 2021). The gene annotation file for switchgrass was accessed from the US Department of Energy's Joint Genome Institute (<https://genome.jgi.doe.gov/portal/>).

The switchgrass genome includes over 95,000 genes, and the functions of most of these genes are unknown. To evaluate potential gene function associated with SwMV disease phenotypes, we thus compared identified candidate genes to homologous genes in two better-annotated genomes: *Arabidopsis thaliana* (TAIR 10) and rice (*Oryza sativa*, v7) (accessed from Phytozome, <https://phytozome-next.jgi.doe.gov>). We then used the *Arabidopsis* gene homologs to conduct gene ontology enrichment with ShinyGO v.0.75 (<http://bioinformatics.sdstate.edu/go75/>) to explore biological processes associated with candidate genes for disease severity and disease extent (false discovery rate of 0.05).

Hierarchically clustered heat maps

To visualize the distribution of the set of top viral disease-related SNPs within and among ecological subgroups, we generated heat maps based on the alleles for each plant genotype at the associated loci and grouped them with hierarchical clustering based on Euclidean distance and complete linkage, using Morpheus (<https://software.broadinstitute.org/morpheus>). We generated separate heat maps for disease severity and extent using the top candidate SNPs that were associated with GWAS analyses for the separate phenotypes. The heat maps were derived from matrices of switchgrass plant genotypes (columns) and the combination of alleles selected at each SNP position. The cell values of each matrix indicate whether genotypes are homozygous for the reference allele ('0'); heterozygous ('1'); or homozygous for the alternate allele ('2') at each SNP position. To reduce figure size, we excluded redundant genotypes within ecological subgroups that were genetically identical at the selected disease severity/extent loci. For both heat maps, each genotype effectively represents a unique combination of alleles within ecological subgroups (i.e., subpopulation-ecotype affiliation).

Results

Panel-wide disease progression

Symptoms of SwMV infection were evident in the GWAS diversity panel beginning in 2019, the year after panel establishment. Disease prevalence rose from 6% in 2020 to saturate at about 21% in late 2021 (Fig. 3.2a). Mean disease severity and extent values likewise initially rose quickly and then peaked in 2021 before dipping slightly in 2022 (Fig. 3.2b-c).

Variation in disease phenotype among ecotypes, subpopulations, and ecological subgroups

In separate models for repeated measures analysis, morphologically-based *Ecotype* (upland, lowland, coastal) and SNP-based *Subpopulation* (Atlantic, Gulf, Midwest) each

independently explained significant variation in disease severity and extent (relative to null models, Table 3.2). *Ecotype* was a somewhat less effective predictor than *Subpopulation*.

Nonetheless, it was clear that in 2021 and 2022, contrary to our prediction, the lowland ecotype experienced the most (not the least) viral disease, while the coastal ecotype experienced the least, and upland was intermediate (Fig. 3c-d). Lowland individuals expressed ~2–7x more severe ($t \geq 2.43$, $p \leq 0.041$, Tukey's HSD) and ~8–14x more extensive disease than coastal plants ($t \geq 3.64$, $p \leq 0.00087$). And, compared to upland plants, lowlands had ~1.5–3x more severe disease ($t \geq 4.56$, $p \leq 0.00002$) and 1.3 –2.5x more extensive disease ($t \geq 2.45$, $p \leq 0.0387$).

Among the SNP-based subpopulations, the Atlantic subpopulation had markedly and unexpectedly low levels of disease, while the Midwest subpopulation expressed the most and the Gulf subpopulation only slightly less (Fig. 3a-b). The Midwest and Gulf subpopulations experienced disease that was 2.5 – 7.5x more severe ($t \geq 3.67$, $p \leq 0.0008$, Tukey's HSD) and 14 – 39x more extensive than the Atlantic subpopulation ($t \geq 2.62$, $p \leq 0.025$) in 2021 and 2022.

The best predictor of disease severity and extent, however, was the analysis based on ecological *Subgroup*, which outperformed the *Subpopulation* and *Ecotype* models of disease severity ($\chi^2 \geq 163.5$, $df = 12$, $p < 0.0001$) and extent ($\chi^2 \geq 68.2$, $df = 12$, $p < 0.0001$) (Table 3.2 and 3.3). The additional variation explained by *Subgroup* was most pronounced for the Gulf subpopulation, which was represented as having intermediate levels of disease (in the *Subpopulation* model), but in which there was clear within-subpopulation divergence by ecotype. From mid-2020 through mid-2022, Gulf-lowland genotypes experienced disease that was at least 2.3x more severe ($t \geq 5.56$, $p < 0.00001$) and 4.5x more extensive ($t \geq 3.46$, $p \leq 0.0053$) than Gulf-coastals. In contrast, Gulf-coastals experienced notably little disease, similar to the

Atlantic-upland and Atlantic-coastal subgroups (Severity: $t \leq 2.06$, $p \geq 0.24$) and (extent: $t \leq 1.23$, $p \geq 0.74$) over the three-year period (Fig. 3.3e-f).

A parallel result was found for the upland ecotype, which was estimated to have intermediate levels of disease in the *Ecotype* models of severity and extent. After accounting for subpopulation affiliation (in the *Subgroup* models), however, it was evident that while the Midwest-upland genotypes experience much disease (comparable to the Gulf-lowlands), the Atlantic-uplands had very little. Compared to the Atlantic-upland subgroup, Midwest-upland plants had at least 3x more severe disease ($t \geq 7.14$, $p < 0.0001$) and 8x more extensive disease ($t \geq 4.37$, $p \leq 0.00015$) between early 2021 and mid-2022.

After accounting for both subpopulation and ecotype in the *Subgroup* models, our results reveal that viral disease expression over time in the diversity panel depended upon both subpopulation and ecotype affiliation, as best evidenced by the Subgroup delineations. Once viral disease had accumulated in the planting, Midwest-upland and Gulf-lowland accessions exhibited at least 2x more severe and 4x more extensive disease than the Atlantic- and coastal subgroups (Atlantic-upland, Atlantic-coastal, Gulf-coastal).

Heritability

Narrow-sense heritability (h^2) for both disease phenotypes generally increased with disease prevalence in the diversity panel plot (Table 3.4). The highest h^2 value for a severity GWAS was 0.49 (mid-2021 and mid-2021), and the highest h^2 value for extent was 0.33 (mid-2022).

Disease phenotypes are associated with numerous widely distributed genes

We ran 12 GWAS analyses which identified in total 1,100 significant SNPs ($-\log_{10}(p) > 8.411$) that were widely distributed across all 18 chromosomes of the tetraploid switchgrass

genome (Fig. 3.4). From this set, LD-based SNP-pruning identified 649 unique SNPs, represented across all chromosomes (Fig. 3.5). For further analysis, we focused on the ‘67 top candidate’ SNPs identified as the most significant SNPs within LD blocks in multiple GWAS. These were distributed across all chromosomes, except 3N and 4K, with chromosomes 1N, 2N, and 5N containing the most (31%, 21/67). Of the top candidates, 10.4% (7/67) were associated uniquely with disease severity; 49.3% (33/67) with disease extent; and the remaining 40% (27/67) with both.

Hierarchically clustered heat maps

In hierarchically clustered heat maps (representing plant genotypes at 67 top candidate SNP positions), genotypes segregated into two main groups for both severity and extent. One branch was almost exclusively associated with the Midwest subpopulation, and the other branch contained all other ecological subgroups. Compared to the distinct branch separating out the Midwest-uplands, the differences between Gulf-lowlands and Gulf-coastals was more gradual, and the Gulf genotypes gradated from lowland to coastal approaching the relatively homogenous group of genotypes representing the Atlantic subpopulation (Fig. 3.6).

Candidate genes and homologs in *Arabidopsis thaliana* and *Oryza sativa* (rice)

Of the top candidate SNPs, 21% (14/67) were located in intergenic regions within the AP13 reference genome and were not within +/-10kbp association of known switchgrass genes. For the remaining 53 SNPs, we identified a total of 143 genes in +/-10kbp association of switchgrass genes, 79% (113) of which had homology with rice genes and/or *Arabidopsis* genes. There were 96 homologous *Arabidopsis* genes identified and 110 for rice. Only five switchgrass genes were homologous to genes with coiled-coil, nucleotide-binding, and leucine-rich repeat domains, which are hallmarks of R genes. Others represented potential defense perception genes

(receptor-like kinases), defense signaling pathways (e.g., protein kinases, SA/Et signaling pathways, ABC transporters) and genes that encode proteins that target viral proteins for degradation (ubiquitin, pentatricopeptide). A complete list of switchgrass gene candidates and homologous genes is available in Table S3.1).

Gene ontology (GO) enrichment with the 96 homologous *Arabidopsis* genes revealed 30 enriched pathways associated with disease severity and 21 for disease extent (Fig. 3.7). The pathways with most support for disease severity and extent differed. For severity, top pathways featured regulation of immune system responses, including a pathway associated with programmed cell death (Fig. 3.7a). For disease extent, the pathways with the most support were associated with signaling and cell communication (Fig. 3.7b).

Discussion

Throughout their evolutionary histories, wild plant species likely encounter a broad range of plant-infecting viruses, some beneficial or benign, some pathogenic. These viral interactions have strong potential to shape host genomes and population dynamics. Here, we report the first large-scale genetic analysis of wild virus interactions with a long-lived wild plant, a priority goal in pathosystem research (Demirjian *et al.*, 2023). Through multi-year longitudinal analysis of disease phenotypes and GWAS of two disease metrics over time, we identified clear differences in viral disease susceptibility among ecological subgroups of switchgrass and a suite of associated candidate genes based on conservative selection of SNPs. We found 30 candidate switchgrass genes without homology in *Arabidopsis* and rice that may represent novel elements regulating plant virus interactions in natural ecosystems. Our results reveal that susceptibility to SwMV disease appears to be associated with a broad network of many small-effect genes implicated in diverse cellular, metabolic, and immune response pathways. Our results further

indicate that wild virus disease expression in switchgrass is equally or more heritable relative to expression of crop virus diseases in common crops considered to date.

Extending GWAS to wild plant – wild virus interactions

Crop viruses known to cause economic damage compose only a small proportion of the plant-infecting viruses discovered to date (Edgar *et al.*, 2022). In crop protection, genome-wide association studies have broadened understanding of the genetic underpinnings of crop defenses against crop viruses and other pathogens in hosts such as soybean (Chang *et al.*, 2016; Che *et al.*, 2020), maize (Sitonik *et al.*, 2019; Rossi *et al.*, 2020), and rice (Cubry *et al.*, 2020). A consensus finding is the polygenic nature of defense responses, indicating much greater complexity in plant disease genetics than indicated by qualitative disease responses mediated by R genes (Demirjian *et al.*, 2023), which were an earlier focus in crop improvement.

To fully understand mechanisms of disease resistance and emergence, however, we have much yet to learn from the diversity of wild plant–virus interactions in nature and their influence on plant genomes. This first study of a wild virus in a long-lived wild host strengthens the conclusion that the genetic basis for viral interactions with hosts is multi-layered and complex. The only previous viral GWASs with wild hosts, to our knowledge, have examined crop virus responses in *Arabidopsis thaliana* (Pagny *et al.*, 2012; Rubio *et al.*, 2019; Butković *et al.*, 2021). However, the short-lived life history strategy of *Arabidopsis* has more in common with crop plants than with most other wild plants, which are long-lived and thus likely navigate more diverse microbial interactions over their lifespan.

To capture the rich complexity of wild plant defense responses, we examined switchgrass dynamics in the field, not indoors, with plants grown in real soil profiles, not pots or trays; measured disease responses over a three-year time period appropriate for a perennial plant; and

evaluated ~12.9M polymorphic markers, more than any previous plant virus GWAS (Demirjian *et al.*, 2023). During the three-year study, plants were exposed to a full range of natural environmental conditions—heat, cold, herbivory, and microbial interactions. This included natural exposure to wild leafhoppers that transmitted infection to the plants; no mechanical inoculation was used (or was possible). No comparable genetic study of wild plant viral disease in an unconstrained outdoor environment has been conducted. Our findings thus permit initial comparison of disease responses and their genetic architecture in long-lived wild plants with those of short-lived model plants and of crops.

Switchgrass susceptibility to viral disease varies within and among ecotypes and subpopulations

Field-based phenotypic measures revealed that susceptibility to viral disease varies significantly among and within switchgrass subpopulations. Strikingly, the Atlantic subpopulation (dominated by coastal ecotypes) and coastal ecotypes within the Gulf subpopulation experienced very little viral disease, whereas the Midwestern subpopulation (dominated by upland ecotypes) and lowland ecotypes in the Gulf subpopulation experienced ~5x more. The sharp difference in disease susceptibility between the coastal and lowland ecotypes within the Gulf subpopulation is particularly notable because the Gulf subpopulation is considered the most genetically homogenous of the three subpopulations (Lovell *et al.*, 2021).

We can only speculate about the ecological factors that have driven evolution of viral disease resistance in the coastal ecotypes along the Gulf and throughout the Atlantic subpopulation or favored its relaxation in the Gulf-lowland and Midwestern-upland subgroups (Fig. S3.1). It is notable that the upper Midwest region inhabited by the more virus-susceptible Midwest switchgrass subpopulation strongly overlaps the known geographic distribution of SwMV. It is not intuitive that local host genotypes might be more susceptible to a local virus, but

similar observations have been noted before. For example, in a study of barley yellow dwarf virus effects on California native grasses, a population of native grass located at the collection point of the viral inoculum was among the most susceptible to it (Malmstrom *et al.*, 2005).

The pattern of viral disease susceptibility in switchgrass seen in our study is similar in some ways to patterns seen with fungal diseases. In a comparable study of rust in the same diversity panel, Midwest genotypes acquired much more extensive infections than Atlantic or Gulf-coastal plants (VanWallendael, pers. communication). In a different study, northern upland cultivars (mostly Midwest) likewise experienced greater anthracnose fungal disease than did southern cultivars predominantly originating from the Atlantic coast (Hoffman *et al.*, 2016). More broadly, the foliar phytobiomes of Midwestern subpopulations are distinct from those of Atlantic and Gulf genotypes (all ecotypes within them considered together) (Bowsher *et al.*, 2020; VanWallendael *et al.*, 2022). Taken together, these findings suggest that either (i) there may be some shared genetic basis for defense responses against several different microbial groups, including viruses, or (ii) there exists some congruity in broad levels of microbial pressure in different regions and/or habitats that patterns separate defense responses similarly. Whatever the cause, the emerging body of evidence indicates that Atlantic and coastal genotypes of switchgrass generally have greater disease resistance than Midwest and Gulf-lowland genotypes.

GWAS: general characteristics of SNP profile in switchgrass

In comparison to findings for crop plant–virus GWAS to date, one of the most striking features of our GWAS of wild viral disease in switchgrass is the larger number and wider chromosomal distribution of significant SNPs identified. GWAS of viral disease in maize (Rossi *et al.*, 2020), rice (Cubry *et al.*, 2020), and soybean (Che *et al.*, 2020) have typically identified just a dozen or so significant SNPs within relatively few genomic regions. In *Arabidopsis*, the

number of significant SNPs found in viral disease analyses has ranged from one (Pagny *et al.*, 2012) to ~25 SNPs concentrated at eight genomic regions (Rubio *et al.*, 2019). In contrast, individual GWAS in the switchgrass study reported here frequently implicated 100s of SNPs that were significant above a conservative $-\log_{10}(p)$ value. After SNP pruning, we still identified >600 significant SNPs that were widely distributed across all chromosomes. This number of significant SNPs is notably more than found in switchgrass GWAS for other phenotypic traits. For example, analyses of flowering time (Grabowski *et al.*, 2017), nitrogen use (Shrestha *et al.*, 2022), climate adaptation (Lovell *et al.*, 2021), and fungal phytobiome (VanWallendael *et al.*, 2022) identified 1–25 significant SNPs (the latter two studies were included on the same diversity panel and workflow pipeline as ours). GWAS of switchgrass leafspot disease (*Bipolaris*) found more SNPs, but still an order of magnitude fewer than the number we identified.

It is possible that the large number of SNPs we identified is a result of genomic inflation due to population structure. However, this seems unlikely because we used protocols to correct for population structure, and genomic inflation factors associated with each GWAS ranged from 0.836 – 1.032 (average = 0.971). An inflation factor of 1.0 is associated with no inflation, and 1.10 is generally an accepted maximum threshold (Williams *et al.*, 2021). Moreover, our GWAS was conducted with the same genotypes used in other switchgrass GWAS studies that identified much smaller sets of SNPs, suggesting that the larger number of SNPs found here is trait-related and not an artefact of population structure.

Estimates of heritability vary widely across GWAS involving viral diseases in crop and model plants. Most studies calculate broad-sense heritability (H^2)—the proportion of phenotypic variation due to genetic factors (including dominance and epistasis)—which requires replication

among genotypes or known pedigree. We calculated narrow-sense heritability (h^2)—the proportion of phenotypic variation explained by additive genetic variation (i.e., SNPs)—which excludes variation due to dominance and epistasis. Narrow-sense heritability estimates from *Arabidopsis* viral disease GWAS ranges from 0.03 to 0.51 (Montes *et al.*, 2020; Butković *et al.*, 2021). The latter value was calculated from a controlled greenhouse experiment with fewer genotypes, which may reduce overall phenotypic variation (and thus inflate h^2). Still, our h^2 estimates of 0.49 and 0.32 for disease severity and extent (after infection fully accumulated in the panel) are comparable to the high end of those found in *Arabidopsis*–virus GWAS, and are ~10x higher than the lower estimates. Estimates of H^2 from other plant–virus GWAS range from 0.03–0.79. Given that h^2 is frequently 50–90% less than H^2 (Wang *et al.*, 2013), our h^2 estimates are comparable or higher than values found for other crop- and *Arabidopsis*-virus GWAS. Because virus transmission in our study was conducted through wild leafhopper activity, the values for switchgrass represent heritability of both vector and viral susceptibility. Overall, our results indicate that SwMV disease in switchgrass is a highly heritable trait.

Candidate genes regulating SwMV disease in switchgrass

With switchgrass, we identified 143 candidate genes in +/-10kbp association with a conservative list of most significant pruned SNPs represented in multiple GWAS (i.e., ‘top candidate SNPs’); this stringent filtering reduced our initial focal list of SNPs by an order of magnitude. Of these, thirty are novel genes with no annotation or known homology in rice or *Arabidopsis*.

Comparing the number of candidate genes found in this study with those found in studies of crop and model plants is complex because identification of candidates is heavily influenced by study-specific criteria, which vary. Furthermore, in publications some investigators provide

information about only a selected set of specific genes, and do not publish the full set of candidates. Nonetheless, the number of associated genes we found is still significantly higher than for most other crop-virus GWAS.

In comparing different host types, another interesting axis to consider is the percentage of candidate genes identified as R genes. Of the 113 candidate switchgrass genes annotated through homology, only 4.4% bore hallmark features of R genes, including coiled-coil, nucleotide-binding, and leucine-rich repeat domains (CC-NB-LRR). One SNP on the 2N chromosome, which was among the most frequently identified (four GWAS), was associated with a switchgrass gene (Pavir.2KG420100) homologous to a coiled-coil protein in rice (Os07g28470.1). Additional SNPs on the 5K, 7K, and 8K chromosome were also associated with multiple GWAS and annotated as *A. thaliana* and/or rice CC-NB-LRR genes (Table S3.1). Perhaps the most comparable list of candidate genes associated with viral disease comes from a study of rice yellow mottle virus in African rice (Cubry *et al.*, 2020), which identified ~200 rice genes by screening for candidates within 100kbp association of significant SNPs; these were generally localized around a single genomic peak. Among the list of described rice genes Cubry *et al.* identified, the proportion of genes bearing R gene hallmarks was 5%, similar to the 4.4% found in switchgrass. Kinases showed similar representation in rice and switchgrass, with 2.6% and 2% receptor-like kinases and 1.3% and 2% protein kinases, respectively.

Intriguingly, our analysis with switchgrass identified a homologous rice gene that was also identified as a candidate gene by Cubry *et al.*, a plasma-membrane type calcium-transporting ATPase associated with endoplasmic reticulum (LOC_Os03g42020). Given its appearance in these two distinct viral studies, this gene merits careful investigation as a potential general determinant of viral susceptibility. The ascribed nature of its function makes this

association plausible, as the endoplasmic reticulum has been identified as an important organelle that influences virus–host interactions (Ravindran *et al.*, 2016).

Another distinguishing feature of our switchgrass candidate gene profile, compared with those for rice and *Arabidopsis*, is the greater prominence of genes involved in RNA/DNA processing and modification, including transcription and translation factors; DNA/RNA-binding/recognition genes; and genes associated with ribosomes, polymerases, histones, and methyltransferase. Comparisons with other published plant–virus GWAS gene profiles indicate much less representation of RNA/DNA-processing genes (Pagny *et al.*, 2012; Rubio *et al.*, 2019; Cubry *et al.*, 2020; Butković *et al.*, 2021).

In switchgrass, we also identified several candidate genes potentially involved in defense-related signaling, which was a major biological pathway in GO enrichment for disease extent. Among these are several possible phytohormone signaling genes, including SUPPRESSOR OF AUXIN RESISTANCE1 (AT33410.2), ethylene responsive element binding factor 4 (AT3G15210), and two genes associated the shikimate pathway (AT06350.1, Os12g34874.2), which regulates SA production, an important phytohormone in plant defense (Ding & Ding, 2020).

We likewise identified several candidate genes associated with ubiquitin, an important plant defense protein that targets virus gene products for degradation (Wu *et al.*, 2019). A candidate gene on the 6N chromosome (significant in four GWAS) was homologous with a ubiquitin-like protease gene in rice (Os04g44640.1), and several on the 9N and 9K chromosomes were homologous to ubiquitin genes in *A. thaliana* (AT42160.1, AT55860.1). In available gene profiles from GWAS of virus disease in rice (Cubry *et al.*, 2020) and *Arabidopsis* (Rubio *et al.*, 2019), ubiquitin-associated genes were absent and or listed only once (respectively). In long-

lived wild plants, diverse ubiquitin encoding genes may play a more prominent role in defense against wild viruses.

Our results suggest that the suite of genes regulating wild virus disease in long-lived wild plant hosts may be larger and more diverse than those of crop plants and short-lived wild plants, but more GWAS with wild plants and wild viruses are needed to better understand genetic patterns of crop vs wild plant defense against viruses. Further investigation is also required to conduct standardized comparisons of GWAS results associated with crop and wild host responses to viruses, but initial comparisons indicate a prominent role for DNA/RNA processing genes and suggests that these translated gene products are important points of interaction between wild hosts and wild viruses.

Conclusions and future directions

Our results demonstrate that SwMV disease expression varies within and among switchgrass subpopulations, and they raise interesting ecological questions about the selective pressures that gave rise to the geographic and ecotypic patterning in switchgrass. Future work is needed to investigate the degree to which SwMV disease resistance and susceptibility is a function of selective pressures exerted on switchgrass by SwMV. GWAS revealed a notably large number of significant SNPs associated with diverse homologous genes, suggesting that the genetic architecture of SwMV disease expression in switchgrass is demonstrably larger and more diverse than that of crop virus disease in crops and *Arabidopsis*. The dearth of wild plant–virus GWAS limits our ability to understand the full depth of strategies and adaptations employed by plants and viruses as they interact. Additional investigations are needed to characterize the genetics of wild plant viral disease more broadly, and will likely illuminate additional intriguing patterns regarding the geographic distribution of viral disease susceptibility and resistance.

APPENDIX

Tables and figures

Table 3.1. Among the three subpopulations (rows) and the three ecotypes (columns), the 512 tetraploid genotypes used for GWAS represent five main ecological subgroups (bolded): Atlantic-coastal, Atlantic-upland, Gulf-coastal, Gulf-Lowland, and Midwest-upland. For one Gulf plant, ecotype was unknown. The five main sub-populations (bolded) were used as the focal groups to explore how well ecological subgroup (ecotype-subpopulation combination) explains variation in disease expression over time.

	<i>Coastal</i>	<i>Lowland</i>	<i>Upland</i>	<i>Unknown</i>	
<i>Atlantic</i>	78	3	119	0	200
<i>Gulf</i>	75	126	1	1	203
<i>Midwest</i>	4	0	105	0	109
	157	129	225	1	512

Table 3.2. Model comparisons for repeated measures analysis of disease severity and extent from 2020-2022. Shaded columns represent the results of individual comparisons between each test model (which included an interaction term between *Time* and each model term) and the corresponding null model (i.e., with *Time* as the lone fixed effect). Chi-squared values, degrees of freedom, and p-values were generated from likelihood ratio tests using the *anova* function in R. For both disease phenotypes (severity and extent), significant variation was explained by morphological ecotype (upland, lowland, coastal); SNP-based subpopulation (Atlantic, Gulf, Midwest); and ecological subgroup (subpopulation-ecotype combination for the five well-populated subgroups bolded in Table 3.1). For both dependent variables, ecological subgroup was the strongest predictor, as further described in Table 3.3 (highlighting direct comparisons of three test models of both disease phenotypes). Test statistics were generated using the *anova* function in R for each individual model. After confirming that there were significant differences within *Subgroup*, *Subpopulation*, and *Ecotype* for both disease phenotypes, we conducted Tukey-adjusted comparisons to identify significant pairwise differences (reported in results section). For the null models, ('*Time*' as the lone fixed effect), the base AICc values are listed parenthetically.

Disease phenotype	Model term	Test statistic	ΔAICc	χ^2	df	p-value
Severity	Null (<i>Time</i>)	$F_{(5,2525)} = 70.0$	0 (7963.2)	-	-	-
	<i>Subgroup</i>	$F_{(4,493)} = 43.8$	-272.3	321.0	24	<0.0001
	<i>Subpopulation</i>	$F_{(2,503)} = 52.2$	-204.7	229.0	12	<0.0001
	<i>Ecotype</i>	$F_{(2,503)} = 41.7$	-132.6	156.8	12	<0.0001
Extent	Null (<i>Time</i>)	$F_{(5,2520)} = 23.7$	0 (9226.4)	-	-	-
	<i>Subgroup</i>	$F_{(4,492)} = 22.2$	-159.9	208.6	24	<0.0001
	<i>Subpopulation</i>	$F_{(4,502)} = 28.4$	-118.3	142.5	12	<0.0001
	<i>Ecotype</i>	$F_{(4,502)} = 21.0$	-84.4	108.6	12	<0.0001

Table 3.3. Comparisons among test models. For both disease phenotypes, the ecological subgroup (*Subgroup*) models outperformed the *Subpopulation* and *Ecotype* models in comparisons of three test models of change in disease severity and extent over time. The values in the shaded columns were generated with the *AICctab* function in R (*bbmle* package). The base AICc value is listed parenthetically in the row that corresponds to the *Subgroup* model, which had the lowest AICc for both disease severity and extent. Column ‘df’ lists the degrees of freedom for each individual model, and ‘Weight’ represents the Akaike weights (out of 1.0) for the three models used to describe variation in disease severity and extent.

Disease phenotype	Independent variable	$\Delta AICc$	df	Weight
Severity	<i>Subgroup</i>	0 (7575)	33	1
	<i>Subpopulation</i>	139.1	21	<0.001
	<i>Ecotype</i>	196.9	21	<0.001
Extent	<i>Subgroup</i>	0 (88781)	33	1
	<i>Subpopulation</i>	43.8	21	<0.001
	<i>Ecotype</i>	66.4	21	<0.001

Table 3.4. Narrow-sense heritability (h^2) estimates generally increased with disease prevalence. ‘CI’ column represents 95% confidence intervals, and additive genetic variance and error variance are represented in columns ‘ σ^2_A ’ and ‘ σ^2_E ’, respectively.

Year	Season	Phenotype	Prevalence	h^2	CI (+/-)	σ^2_A	σ^2_E
2020	Early	Severity	0.06	0.15	0.20	0.05	0.30
	Mid			0.24	0.21	0.27	0.87
	Early	Extent	0.11	0.04	0.08	0.02	0.49
	Mid			0.10	0.14	0.16	1.47
2021	Early	Severity	0.17	0.37	0.27	0.47	0.82
	Mid			0.49	0.26	0.88	0.91
	Early	Extent	0.21	0.18	0.21	0.33	1.46
	Mid			0.32	0.25	1.11	2.37
2022	Early	Severity	0.21	0.40	0.27	0.59	0.87
	Mid			0.49	0.25	0.86	0.91
	Early	Extent	0.22	0.33	0.29	1.04	2.14
	Mid			0.22	0.20	0.54	1.91

Figure 3.1. Scale for SwMV symptom measurements used to rate disease severity, from 0 to 4. 0 corresponds to no symptoms observed; 1 corresponds to a few dots or blemishes but no clear symptoms, 2 corresponds to identifiable but few SwMV symptoms ($\leq 10\%$ leaf affected); 3 corresponds to moderate to strong SwMV symptoms (10-50% leaf affected); and 4 corresponds to severe SwMV symptoms ($> 50\%$ leaf affected).

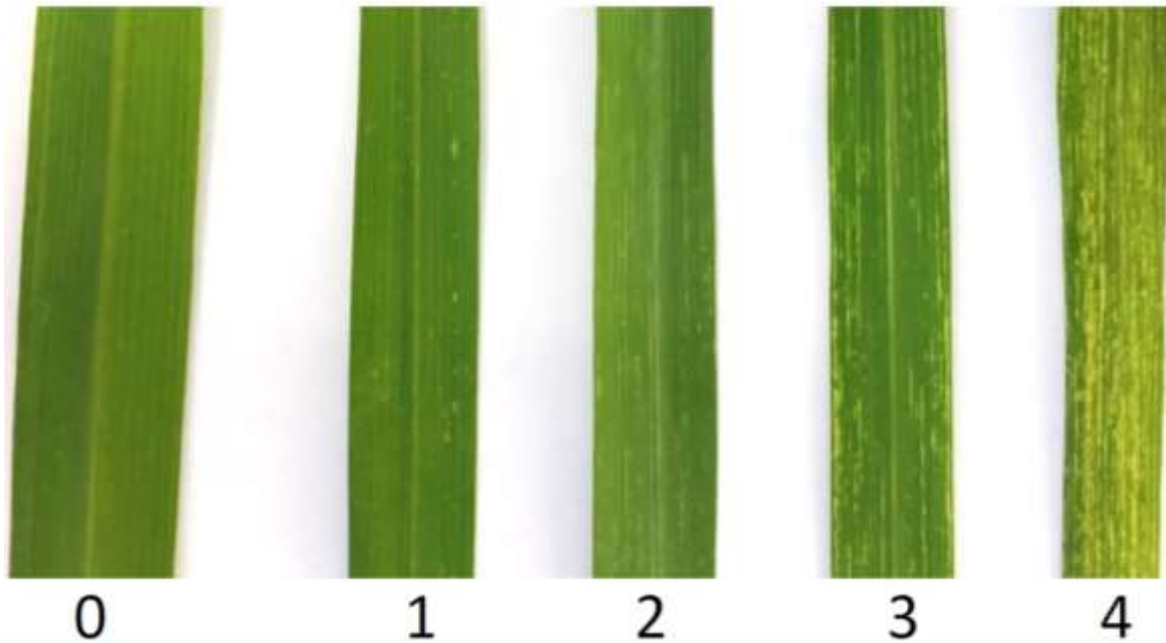


Figure 3.2. Patterns of disease prevalence, disease severity, and extent over time. (a) Disease prevalence was calculated based on the proportion of the 595 plants that were expressing clear signs of disease (severity score of two or greater) at the six time points. (b-c) Overall disease severity and extent metrics ($\pm$ 95% CIs) were estimated with repeated measures models using the mean phenotype scores by genotype ($n = 512$).

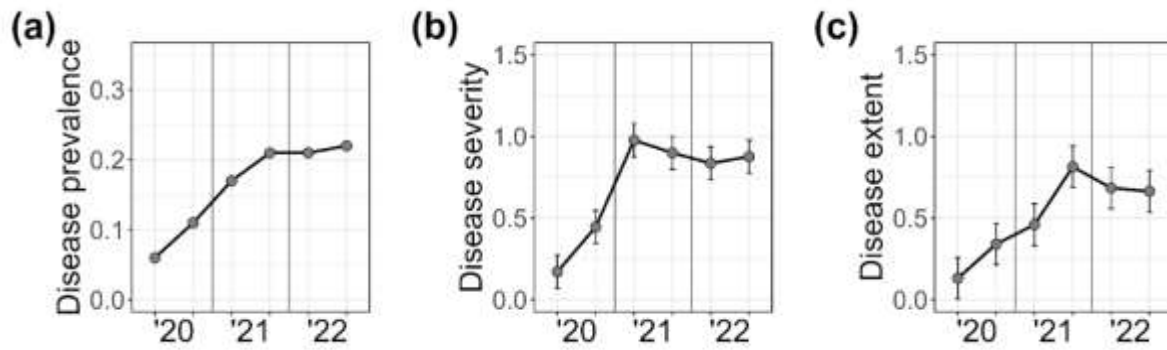
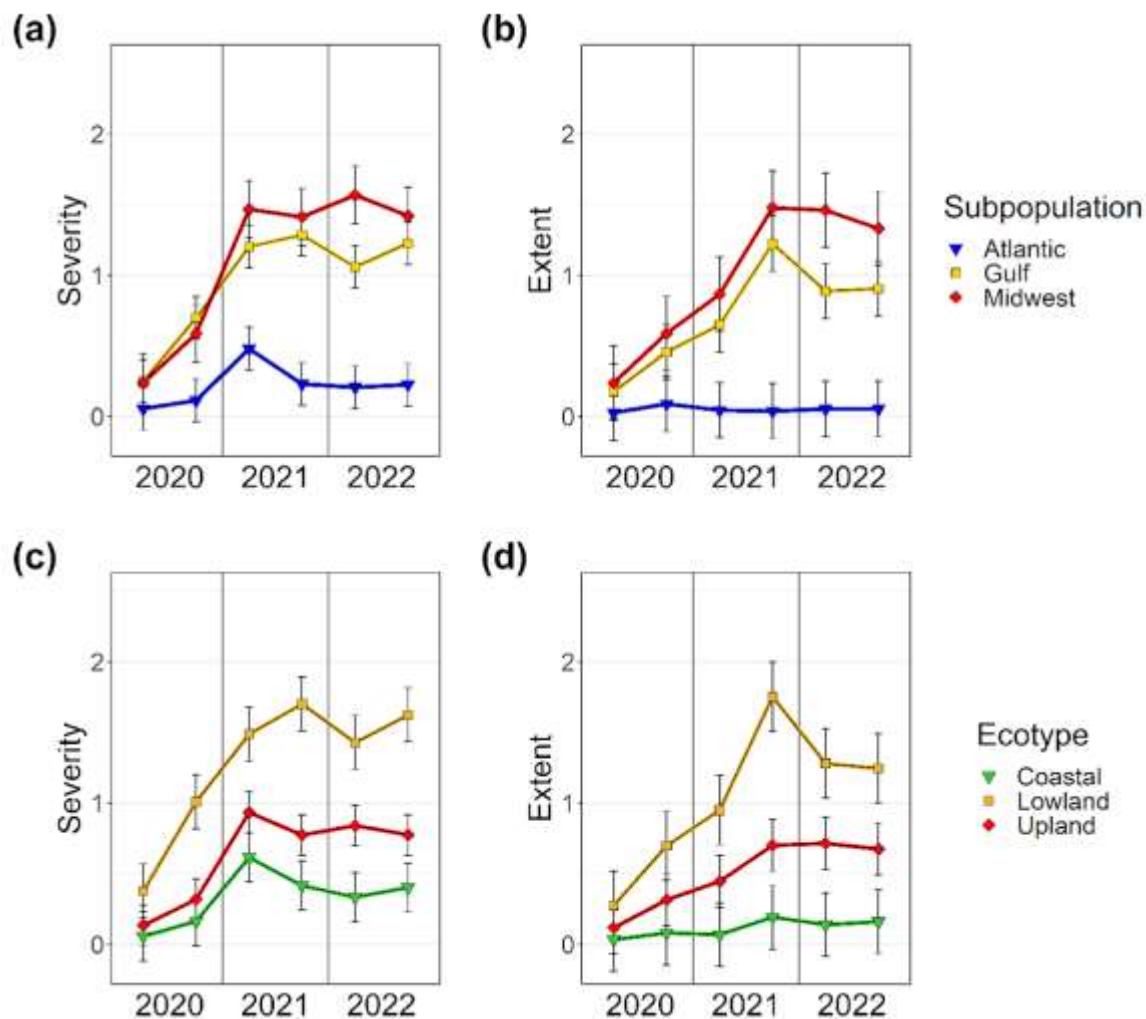


Figure 3.3. Variation in disease phenotype is best explained by ecological subgroup (subpopulation-ecotype affiliation). For all graphs, shapes represent model-estimated means for each disease phenotype at six time points between 2020 and 2022 (+/- 95% CI). (a-b) Estimates from the *Subpopulation* models of severity and extent. The Midwest subpopulation (red diamonds) generally had greater disease levels than Gulf (yellow squares) and Atlantic (blue triangles) genotypes. (c-d) Model estimates from the *Ecotype* models. The lowland ecotype (yellow squares) had greater disease levels than coastal (green triangles) and upland (red diamonds) ecotypes. (e-f) Estimates from the *Subgroup* models captured the most variation. Among the five subpopulation-ecotype subgroups, there were two divergent patterns. Gulf-lowland (orange triangles) and Midwest-upland (red diamonds) had consistently greater disease levels than Gulf-coastals (light-green squares), Atlantic-coastals (light-blue triangles), and Atlantic-uplands (purple circles).



(Figure 3.3, cont'd)

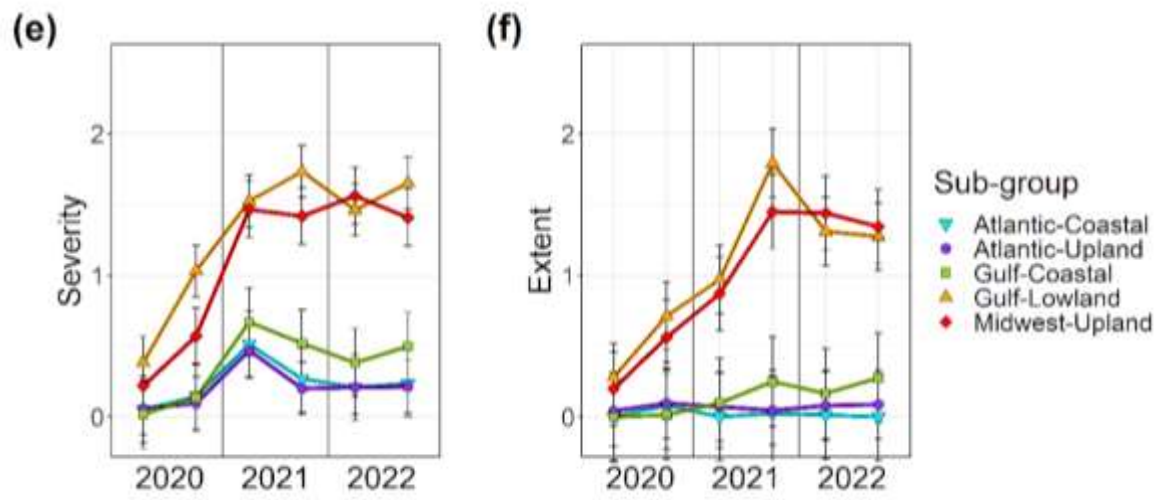


Figure 3.4. Manhattan plots for 12 GWAS of SwMV disease. Dashed line indicates the Bonferroni significance level ($\alpha = 0.05$) of 8.41. SNPs with red circles were identified as the most significant SNPs within LD blocks ($r^2 = 0.20$). Most SNPs were associated with increased disease expression. Black SNPs denote SNPs that were associated with reduced level of disease. Each quarter panel (a-d) represents a specific disease season and disease phenotype combination: (a) early-season severity, (b) mid-season severity, (c) early-season extent, and (d) mid-season extent. Roman numerals denote year: (i) 2020, (ii) 2021, and (iii) 2022.

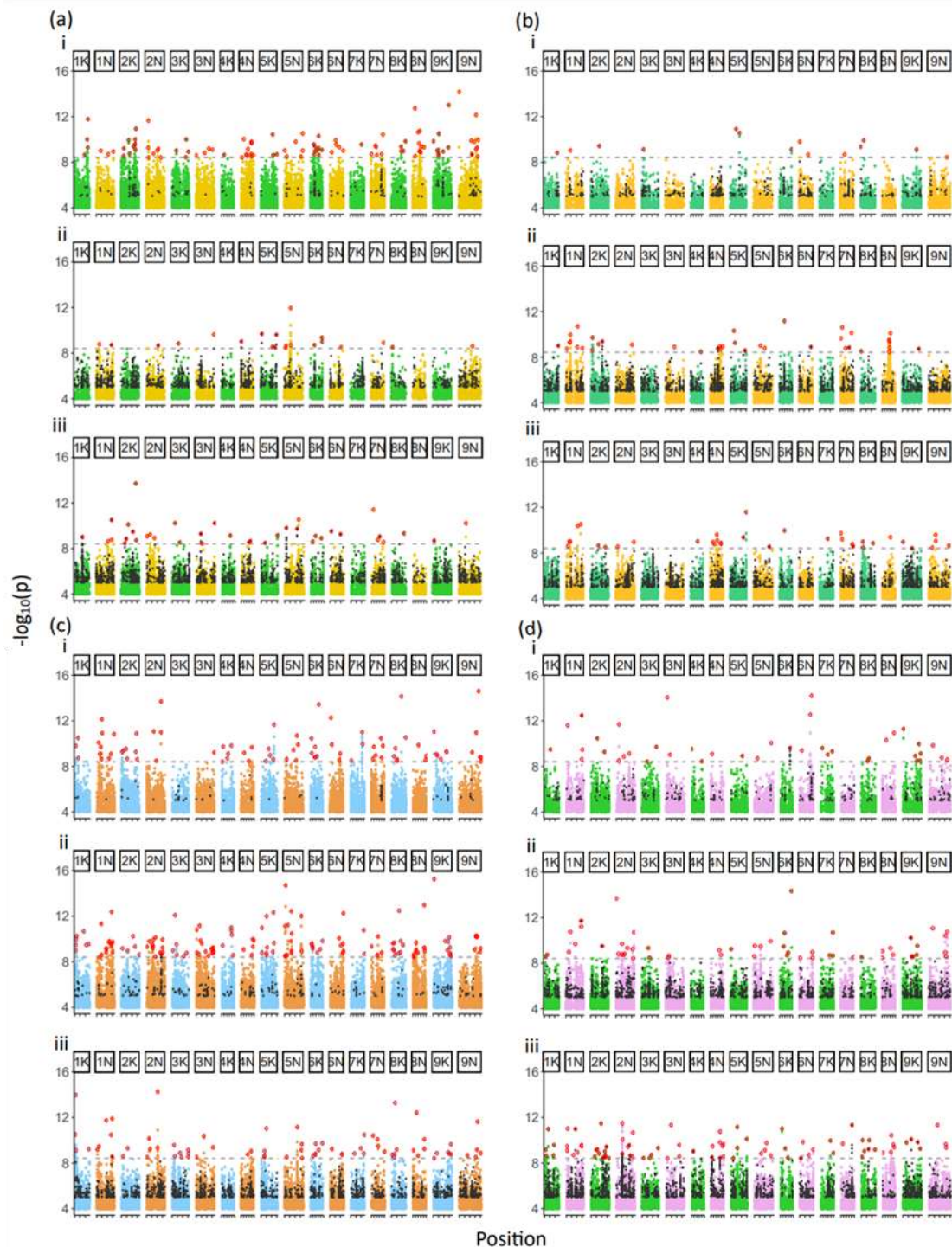


Figure 3.5. Chromosomal distribution of SNPs identified by linkage disequilibrium (LD) based pruning ($n = 649$). SNPs were pruned to the most significant SNP within LD regions of $r^2 = 0.2$.

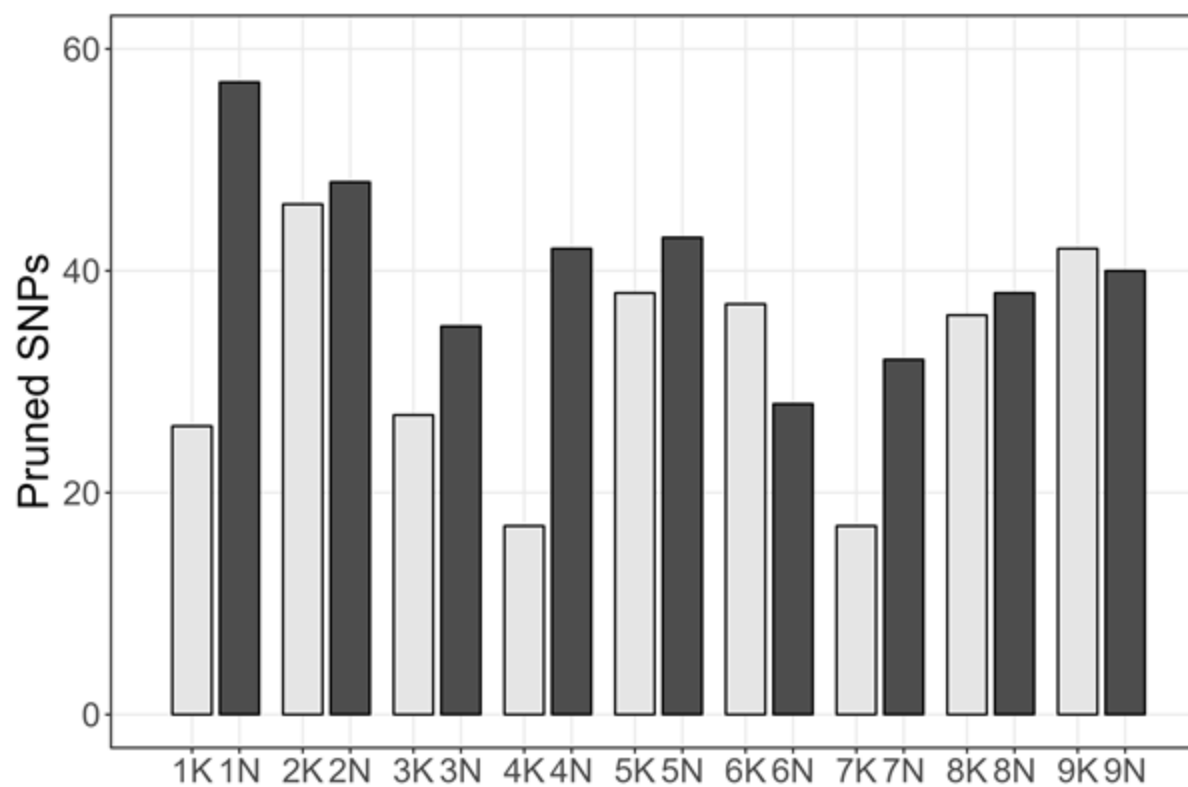


Figure 3.6. Hierarchically clustered heatmaps for alleles associated genotypes at (a) the 34 top candidate SNPs for disease severity and (b) the 60 top candidate SNPs for disease extent. Each column represents a unique genotype, and rows represents SNPs. Cell values of the matrix represent degree to which the genotype is similar (grey) or dissimilar (black) from other genotypes. Red = Midwest genotypes; yellow = Gulf-lowland; green = gulf-coastal; blue = Atlantics.

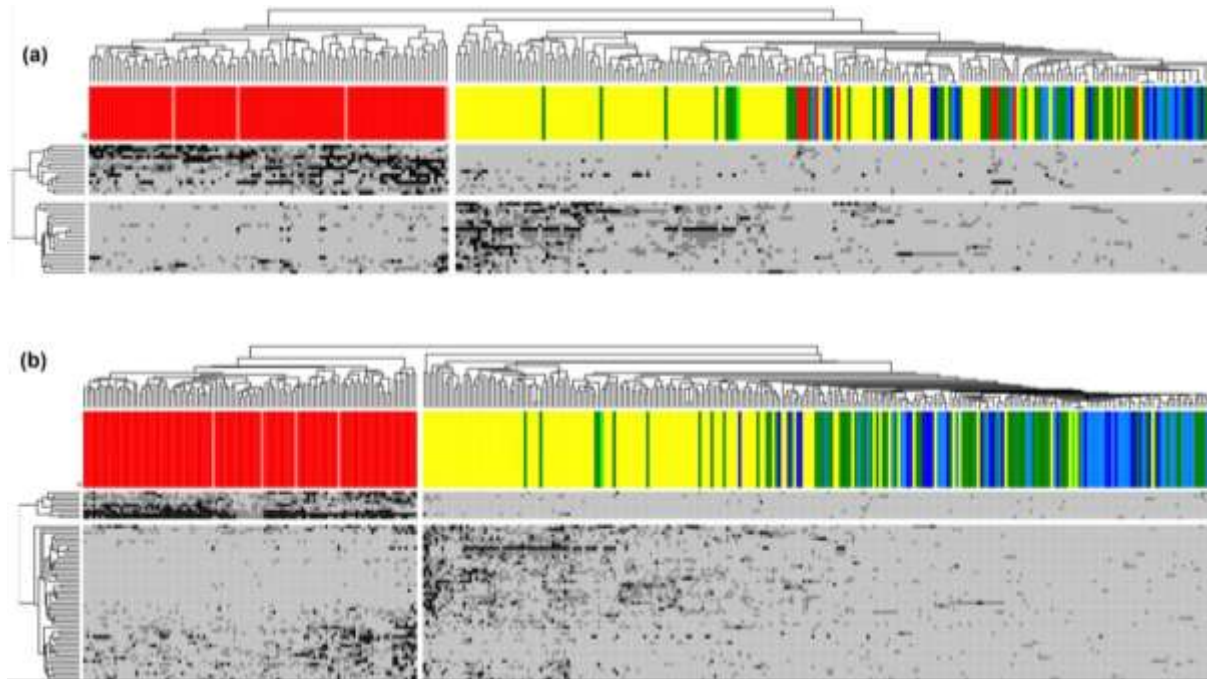


Figure 3.7. Gene ontology (GO) enrichment results for homologous genes in *Arabidopsis thaliana*. The pathways with the best support (red dots) for severity were generally associated with immune responses. For extent genes, the best supported pathways were related to signaling and communication.

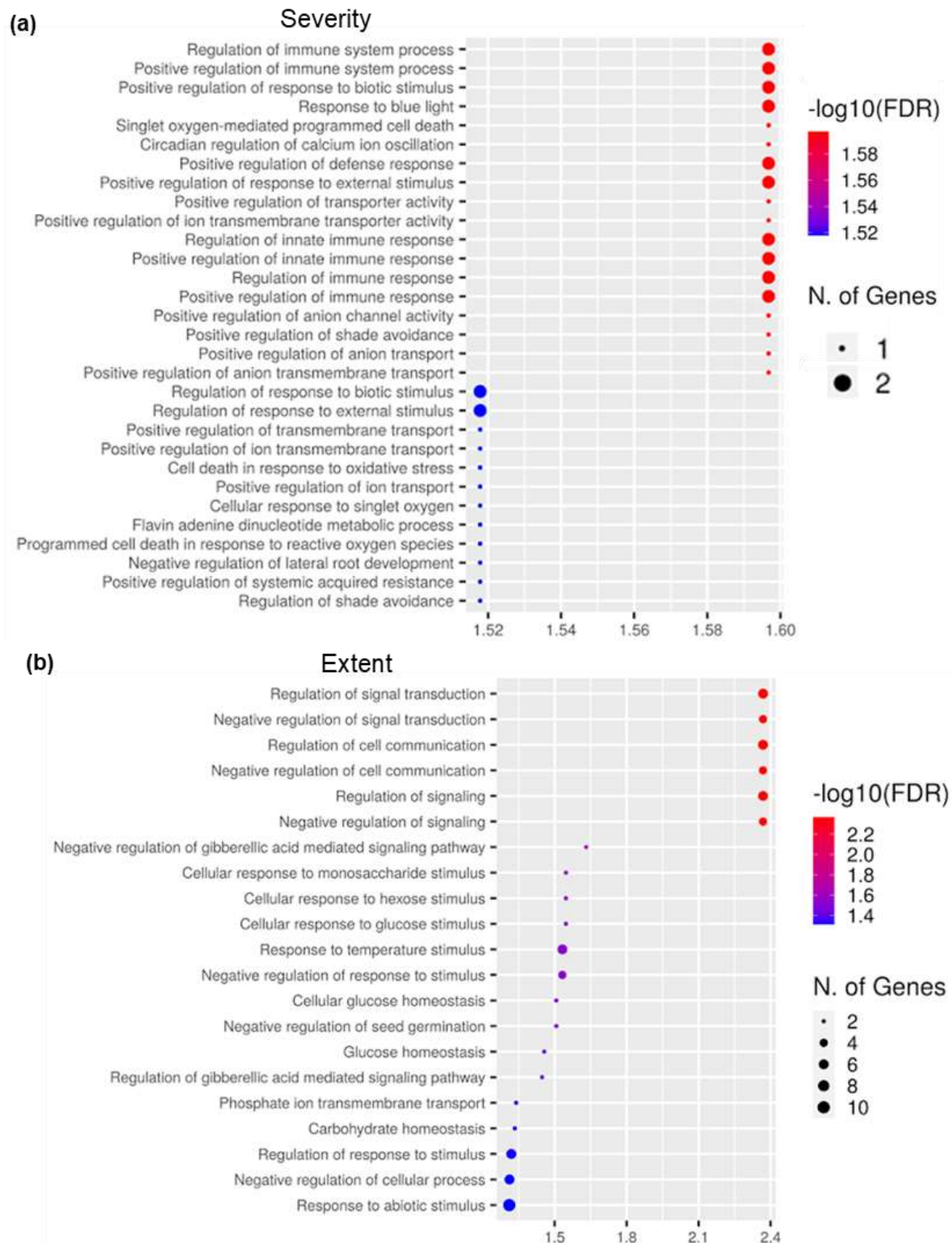


Figure S3.1. Three-year mean disease for genotypes at latitude of origin. Genotypes were grouped into relatively equal bins by mean severity (top) and extent (bottom).

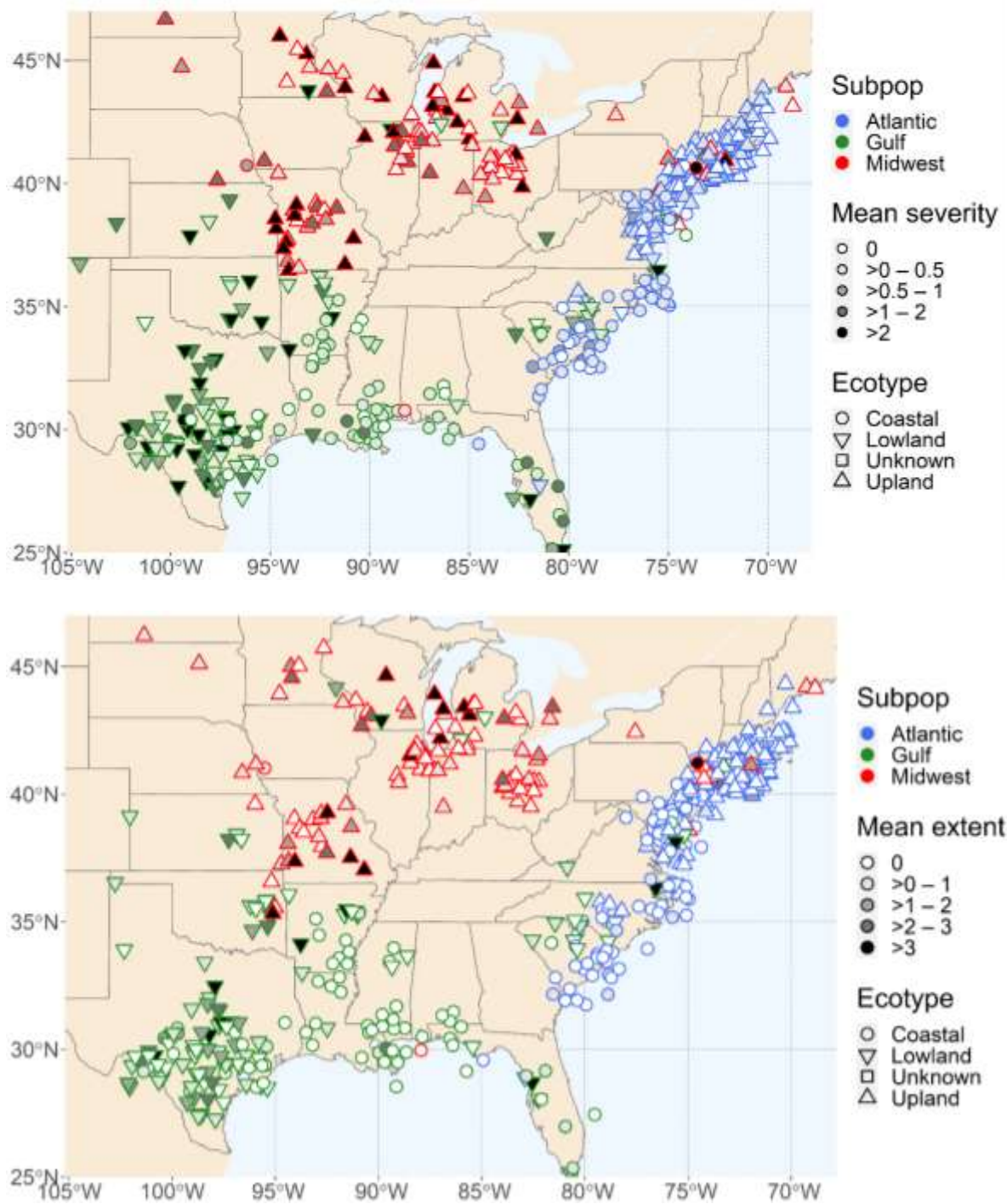


Table S3.1. Switchgrass gene candidates in association with 67 top SNP candidates and associated *Arabidopsis thaliana* and *Oryza sativa* gene homologies. The chromosome ('*Chr*') and position for each SNP are listed in the first column. For each top SNP, the highest $-\log_{10}p$ value is listed in the '*GWAS*' column for both disease phenotypes ('*Sev.*' = severity; '*Ext.*' = extent). Bolded $-\log_{10}p$ values indicate significance above the Bonferroni threshold, and numbers in parentheses indicate the number of times the SNP was identified as the most significant pruned SNP (out of a total of six GWAS for each disease phenotype). The '*P. virg.*' column lists switchgrass genes (without 'Pavir.' prefix) that were in +/-10kbp association with the 67 SNPs. 'NA's indicate that SNPs were in intergenic regions and not in association with known switchgrass genes. Homologous *Arabidopsis thaliana* and rice (*Oryza sativa*) are listed (without 'AT' and 'LOC_Os'), along with associated gene descriptions, in the last two columns (respectively).

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_) Description	Gene (LOC_Os_) Description		
		Sev.	Ext.					
1K 3,657, 171 52,614, 295	5.9 (0)	14.0 (3)	1KG041400	79000.1	histone acetyltransferase of the CBP family 1	02g04490.1	histone acetyltransferase HAC1	
			1KG045200	33410.2	SUPPRESSOR OF AUXIN RESISTANCE1	02g04970.1	expressed protein	
			1KG045000	16130.1	arabinose kinase	02g04840.1	GHMP kinases ATP-binding protein	
	4.93 (0)	9.6 (2)	1KG045100	35320.1	NA	02g04860.1	expressed protein	
			1KG111164	40830.3	RING-H2 finger C1A	02g52870.1	zinc finger, C3HC4 type domain containing protein	
			1KG111138	35370.1	S-locus lectin protein kinase family	02g52850.1	receptor-like protein kinase like protein	
			1KG111151	14040.1	phosphate transporter 3;1	02g52860.1	phosphate carrier protein, mitochondrial precursor	
			1KG111177	44835.2	YbaK/aminoacyl-tRNA synthetase-associated domain	02g52880.1	ybaK/prolyl-tRNA synthetase family	
			1KG111190	21910.1	Integrase-type DNA-binding superfamily	06g10780.1	AP2 domain containing protein	

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
1N 2,396,944		8.7 (1)	9.1 (1)	1KG518600	29380.1	protein kinase family protein / WD-40 repeat family	02g55340.1	WD domain and HEAT domain containing protein
				1NG042400	47120.1	BAX inhibitor 1	02g03280.2	transmembrane BAX inhibitor motif-containing protein
				1NG042300	44740.1	cyclin p4;1	02g03294.1	cyclin
	17,168,256	8.0 (0)	11.3 (2)	Intergenic	NA	NA	NA	NA
	36,080,195	7.7 (0)	11.8 (2)	1NG094800	03370.1	C2 calcium/lipid-binding and GRAM domain	02g10630.1	GRAM and C2 domains containing protein
				1NG239319	NA	NA	NA	NA
	44,379,380	10.7 (2)	6.4 (0)	Intergenic	NA	NA	NA	NA
	57,633,480	10.5 (2)	6.5 (0)	1NG449800	18280.1	tubby like protein 2	02g47640.1	OsFBT4 - F-box and tubby domain containing protein
				1NG449900	21620.1	Adenine nucleotide alpha hydrolases-like superfamily	02g47650.1	universal stress protein domain containing protein
				1NG450000	34530.1	cryptochrome-interacting basic-helix-loop-helix 1	02g47660.1	basic helix-loop-helix
	57,951,394	8.8 (1)	9.5 (1)	1NG450100	NA	NA	NA	NA
				Intergenic	NA	NA	NA	NA
	58,532,417	5.7 (0)	11.9 (2)	1NG459100	NA	NA	09g08300.1	expressed protein
				1NG364900	80050.1	adenine phosphoribosyl transferase 2	02g40010.1	phosphoribosyl transferase
	58,532,448	6.9 (0)	12.5 (3)	1NG459100	NA	NA	09g08300.1	expressed protein

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
2K	6,461,243	8.6 (1)	8.6 (1)	1NG515119	NA	NA	NA	NA
				2KG086000	09630.1	Ribosomal protein L4/L1 family	07g08330.1	ribosomal protein L4
	23,229,606	8.0 (0)	10.5 (2)	Intergenic	NA	NA	NA	NA
	29,755,895	9.1 (2)	8.8 (1)	2KG196100	10690.1	DNA GYRASE A	NA	NA
	44,171,097	9.4 (1)	8.5 (1)	2KG174500	26020.1	Protein phosphatase 2A regulatory B subunit family	07g17360.2	transposon protein
				2KG298402	NA	NA	NA	NA
				2KG298300	38450.1	cytochrome P450, family 735, subfamily A, polypeptide 1	09g23820.1	cytochrome P450 72A1

(Table S3.1, cont'd)

Chr. & Position		GWAS		<i>P. virg.</i>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>	
		-Log ₁₀ p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
2N	1,461,325	9.1 (1)	13.7 (3)	2KG420100	00840.1	DHHC-type zinc finger family	07g28470.1	coiled-coil domain-containing protein 111
				2NG203058	33230.1	Plant invertase/pectin methylesterase inhibitor superfamily	07g49100.1	pectinesterase
				2NG198672	NA	NA	NA	NA
	9,467,606	7.4 (0)	11.7 (3)	2KG499000	19855.1	Chaperonin-like RbcX	07g38230.1	expressed protein
				2NG096246	NA	NA	11g05710.1	expressed protein
	21,886,361	4.9 (0)	11.5 (2)	2NG030000	32300.1	uclacyanin 1	09g39940.1	plastocyanin-like domain containing protein
	23,157,134	4.3 (0)	10.1 (2)	2NG246909	28910.1	CAX interacting protein 4	09g01640.1	CAX-interacting protein 4
	26,682,817	8.8 (1)	11.1 (1)	Intergenic	NA	NA	NA	NA
	32,202,267	4.6 (0)	8.8 (2)	Intergenic	NA	NA	NA	NA
	43,230,915	5.2 (0)	9.4 (2)	2NG014500	NA	NA	NA	NA
				2NG326903	NA	NA	NA	NA
				2NG326600	38880.1	NA	09g21780.1	expressed protein
3K	10,218,159	5.0 (0)	12.1 (2)	3KG469300	27030.1	TOPLESS-related 3	08g06480.1	lissencephaly type-1-like homology motif
				3KG469327	NA	NA	11g35380.1	expressed protein
				3KG469354	NA	NA	NA	NA
				3KG151700	40030.2	Histone superfamily	03g27310.1	histone H3

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana			Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description	
		Sev.	Ext.						
3K	54,284,741	6.0 (0)	9.7 (2)	3KG446600	58140.1	phenylalanyl-tRNA synthetase class IIc family	12g34860.1	phenylalanyl-tRNA synthetase	
				3KG446800	31180.2	Class II aminoacyl-tRNA and biotin synthetases superfamily	02g46130.1	tRNA synthetases class II domain containing protein	
				3KG446700	03115.1	Mitochondrial substrate carrier family	12g34870.2	mitochondrial carrier protein	
				3KG447000	06350.1	dehydroquinase dehydratase, putative / shikimate dehydrogenase	12g34874.2	bifunctional 3-dehydroquinase dehydratase/shikimate dehydrogenase,chloroplast precursor	
				3KG446901	NA	NA	NA	NA	
4N	12,094,756	10.0 (1)	9.2 (1)	4KG017700	02010.1	secretory 1A	06g04450.1	Sec1 family transport protein	
				4NG080319	34160.1	Tetratricopeptide repeat (TPR)-like superfamily	12g17080.1	pentatricopeptide	
				4NG079600	19740.1	P-loop containing nucleoside triphosphate hydrolases superfamily	06g12160.1	AAA-type ATPase family	
	34,506,788	6.8 (0)	10.8 (2)	4NG215511	NA	NA	NA	NA	
				4NG215522	56860.1	GATA type zinc finger transcription factor family	06g37450.1	GATA zinc finger domain containing protein	
				4NG215500	09640.1	Translation elongation factor EF1B, gamma chain	06g37440.1	elongation factor 1-gamma	
	41,161,788	9.8 (1)	8.6 (1)	Intergenic	NA	NA	NA	NA	

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
5K	46,755,940	8.04 (0)	12.3 (2)	4NG163300	30980.1	SHAGGY-related protein kinase dZeta	06g35530.1	CGMC_GSK.8 - CGMC includes CDA, MAPK, GSK3, and CLKC kinases
				5KG468400	14470.1	NB-ARC domain-containing disease resistance	01g52280.1	NB-ARC domain containing protein
				5KG468407	14470.1		01g52380.1	expressed protein
				5KG468200	NA	NA	05g15200.1	zinc knuckle family
55,505,095		8.6 (2)	7.3 (0)	5KG653107	NA	NA	NA	NA
				5KG653500	10070.1	TBP-associated factor 12	01g63940.1	Transcription initiation factor TFIID subunit A containing protein
5N	2,006,974	6.2 (0)	11.1 (2)	5NG025608	NA	NA	NA	NA
				5NG102200	11050.1	glycosyl hydrolase 9C3	01g12070.1	endoglucanase precursor
				5NG102100	57050.1	Protein phosphatase 2C family	01g46760.1	protein phosphatase 2C
				5NG104708	26080.1	Protein phosphatase 2C family	09g15670.1	protein phosphatase 2C
				5NG107316	33355.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily	01g12020.1	LTPL18 - Protease inhibitor/seed storage/LTP family protein precursor

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
5N	25,605,278	12.0(1)	12.4(1)	5NG341900	16950.1	lipoamide dehydrogenase 1	01g23610.1	dihydrolipoyl dehydrogenase
				5NG224624	NA	NA	NA	NA
	25,734,241	9.0 (1)	9.5 (1)	Intergenic	NA	NA	NA	NA
	63,246,759	8.0 (0)	11.1(2)	5NG510000	29400.1	exocyst subunit exo70 E1	01g55799.1	exo70 exocyst complex subunit domain protein
				5NG584000	37570.1	HSP20-like chaperones superfam.	01g62300.1	SLT1
	63,269,914	7.1 (0)	12.0(2)	5NG584150	NA	NA	NA	NA
				5NG584100	59340.1	WUSCHEL related homeobox 2	01g62310.1	homeobox domain protein
	65,272,687	4.9 (0)	9.7 (2)	5NG110337	22540.1	Major facilitator superfamily	01g65100.1	peptide transporter
				5NG110338	22540.1	Major facilitator superfamily	01g65110.1	POT family
				5NG110336	NA	NA	NA	NA
6K	17,458,430	9.1 (2)	3.8 (0)	5NG499700	50890.1	alpha/beta-Hydrolases superfamily	01g54810.1	lipase
				6KG094412	07750.2	3'-5'-exoribonuclease family	12g21500.1	exosome complex exonuclease
	18,536,400	11.2(2)	10.7(3)	Intergenic	NA	NA	NA	NA
	28,493,021	6.8 (0)	8.7 (2)	5NG132200	13440.2	glyceraldehyde-3-phosphate dehydrogenase C2	08g03290.1	glyceraldehyde-3-phosphate dehydrogenase
				6KG140500	NA	NA	09g11900.1	expressed protein
				6KG140600	NA	NA	NA	NA
	32,391,257	10.3(1)	13.4(1)	6KG055500	11020.1	RING/FYVE/PHD zinc finger superfamily	08g29590.1	zinc finger, C3HC4 type family
				6KG055800	49400.1	Nucleic acid-binding, OB-fold-like	08g27090.1	ribosomal protein S17
	45,316,610	9.4 (1)	14.3(3)	Intergenic	NA	NA	NA	NA

(Table S3.1, cont'd)

Chr. & Position		GWAS		<i>P. virg.</i>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>	
		-Log ₁₀ p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
6N	4,400,775	9.5 (1)	8.4 (1)	6KG012680	10070.1	KH domain-containing protein	08g01930.1	KH domain-containing protein
				6NG034383	NA	NA	NA	NA
				6NG034400	54580.1	RNA-binding (RRM/RBD/RNP motifs) family	08g04440.1	RNA recognition motif containing protein
		8.9 (1)	9.9 (3)	6NG034500	11430.1	plastid developmental protein DAG	08g04450.3	DAG protein, chloroplast precursor
	8,225,071			6NG077100	23780.1	nuclear RNA polymerase D2A	08g07480.1	DNA-directed RNA polymerase subunit
				6NG077430	NA	NA	04g44640.1	ulp1 protease family, C-terminal catalytic domain containing protein
				6NG077300	NA	NA	04g54620.1	expressed protein
				6NG077400	39610.1	Protein of unknown function, DUF617	08g07500.1	DUF617 domain containing protein
	34,555,801	5.4 (0)	10.1 (2)	Intergenic	NA	NA	NA	NA
	51,687,153	4.7 (0)		6NG351900	51150.2	ATP binding microtubule motor family	08g43400.2	kinesin motor domain containing protein
7K	49,696,350	7.9 (0)	10.8 (3)	7KG422001	NA	NA	NA	NA
				7KG421800	NA	NA	04g55600.1	expressed protein
				7KG421901	NA	NA	NA	NA
				7KG422100	67200.1	Leucine-rich repeat protein kinase family	04g55620.1	receptor kinase

(Table S3.1, cont'd)

Chr. & Position	GWAS		<i>P. virg.</i>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>	
	-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
	Sev.	Ext.					
7N 6,588,815 23,394,556	11.4 (3) 8.8 (1)	10.4 (2) 10.1 (1)	Intergenic	NA	NA	NA	NA
			7KG0947 00	NA	NA	01g42860.1	inhibitor I family
			7NG1214 00	48280.1	cytochrome P450, family 71, subfamily A, polypeptide 25	08g01510.1	cytochrome P450
	8.8 (2)	5.0 (0)	7NG3283 00	19810.1	Glycosyl hydrolase family protein with chitinase insertion domain	04g30770.1	glycosyl hydrolase
			7NG4067 00	15210.1	ethylene responsive element binding factor 4	04g52090.1	AP2 domain containing protein
			7NG4065 00	65730.1	YELLOW STRIPE like 7	02g02450.1	transposon protein
	45,031,274	7.3 (0)	7KG4276 00	30190.1	H(+)-ATPase 2	04g56160.1	plasma membrane ATPase
			7NG4392 00	NA	NA	04g57040.1	expressed protein
49,758,308		9.0 (2)	7NG4394 00	08570.1	Pyruvate kinase family	04g58110.1	pyruvate kinase
8K 13,507,671	5.3 (0)	13.3 (2)	8KG1405 00	67730.1	beta-ketoacyl reductase 1	11g24484.1	estradiol 17-beta- dehydrogenase 12
			7NG0769 00	08440.1	Aluminum activated malate transporter family	04g34010.1	aluminum-activated malate transporter
			8KG1403 00	NA	NA	NA	NA
	29,175,263	6.1 (0)	8KG0766 05	NA	NA	01g34410.1	transposon protein, Pong sub-class
			7NG4445 00	63220.1	Calcium-dependent lipid- binding (CaLB domain) family	04g58570.1	C2 domain containing protein
37,593,423	9.34(1)	14.1 (1)	8KG2613 00	02630.1	Protein kinase superfamily	11g39420.1	jacalin-like lectin domain containing protein

(Table S3.1, cont'd)

Chr. & Position		GWAS		<i>P. virg.</i>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description	
		Sev.	Ext						
8K	46,479,439	9.3 (2)	6.9 (0)	8KG311816	58848.2	Disease resistance protein (CC-NBS-LRR class) family	11g39320.1	LZ-NBS-LRR class	
					8KG311818	NA	NA	01g06560.1	transcription factor HBP-1b
					7NG183100	08920.1	cryptochrome 1	04g37920.1	FAD binding domain of DNA photolyase domain containing protein
53,922,484			9.8 (1)	10.5 (1)	8KG251990	NA	NA	NA	NA
				7NG348800	43960.1	Nuclear transport factor 2 (NTF2) family protein with RNA binding (RRM-RBD-RNP motifs) domain	04g43150.2	nuclear transport factor 2	
8N	1,563,341	5.4 (0)	9.1 (2)	8NG025701	53000.1	2A phosphatase associated protein of 46 kD	12g04290.1	PP2A regulatory subunit TAP46	
					8NG026900	10190.1	Calcium-binding EF-hand family	12g04360.1	calmodulin-like protein 1
					8NG026901	NA	NA	NA	NA
					8NG026902	NA	NA	NA	NA
13,828,599		7.0 (0)		12.5 (2)	8NG154200	47710.1	Serine protease inhibitor (SERPIN) family	11g12420.1	serpin domain containing protein
					8KG089100	71910.1	NA	11g09150.1	expressed protein
					06350.1	dehydroquinase dehydratase, putative / shikimate dehydrogenase	12g34874.2	bifunctional 3-dehydroquinase dehydratase/shikimate dehydrogenase,chloroplast precursor	
	20,574,603	10.6 (1)	9.9 (1)	Intergenic	NA	NA	NA	NA	
	31,209,061	9.4 (1)	9.0 (1)	Intergenic	NA	NA	NA	NA	

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log10p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
8N	43,305,589	9.3 (1)	13.0(4)	8NG285104	NA	NA	11g41770.1	expressed protein
				8NG285300	55950.1	CRINKLY4 related 3	11g44250.1	protein kinase
				7NG379600	42160.1	zinc finger (ubiquitin-hydrolase) domain-containing protein	04g55480.2	BRCA1-associated
9K	1,934,790	9.9 (1)	11.3 (3)	9KG379084	66460.1	Glycosyl hydrolase superfamily	03g61280.1	OsMan05 - Endo-Beta-Mannanase
	31,516,881			9KG356200	NA	NA	NA	NA
	56,091,841			9KG092723	NA	NA	NA	NA
				9KG090500	56350.1	Pyruvate kinase family	01g16960.1	pyruvate kinase
				9KG090300	26410.1	Uncharacterised conserved protein UCP022280	03g20860.1	expressed protein
		7.3 (0)	10.2(2)	9KG090700	04030.3	Homeodomain-like superfamily	03g20900.1	Myb transcription factor
				9KG090400	56340.1	RING/U-box superfamily	03g20870.1	zinc finger, C3HC4 type domain containing protein
	60,296,765			9KG276800	NA	NA	10g10180.1	methyltransferase domain containing protein
	60,296,765			9KG514800	NA	NA	NA	NA
				9KG514900	68810.1	basic helix-loop-helix (bHLH) DNA-binding superfamily	03g15440.1	basic helix-loop-helix
		13.0 (1)	10.0 (1)	9KG515200	NA	NA	NA	NA
				8NG313300	55860.1	ubiquitin-protein ligase 1	04g40720.1	expressed protein
				9KG600601	29090.1	Ribonuclease H-like superfamily	NA	NA
				9KG600300	NA	NA	NA	NA
				9KG600500	54560.1	histone H2A 11	03g53190.1	Core histone H2A/H2B/H3/H4 domain containing protein

(Table S3.1, cont'd)

Chr. & Position		GWAS		P. virg.	Arabidopsis thaliana		Oryza sativa	
		-Log ₁₀ p (# sig.)		Gene (Pavir._)	Gene (AT_)	Description	Gene (LOC_Os_)	Description
		Sev.	Ext.					
9N	26,852,915	10.2 (2)	7.9 (0)	9KG616500	72490.1	NA	03g08290.1	expressed protein
				9NG325700	36950.1	Heavy metal transport/ detoxification superfamily	10g30450.1	heavy-metal-associated domain-containing protein
	63,109,619			9NG575114	NA	NA	NA	NA
	68,558,593	9.3 (1)	10.3 (2) 10.2 (1)	9KG101600	37640.1	calcium ATPase 2	03g42020.1	calcium-transporting ATPase, plasma membrane-type
				9NG686000	11180.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily	03g18030.1	Leucoantho-cyanidin dioxygenase
		10.0 (1)	14.6 (1)	9NG686100	58700.1	phosphatidylinositol- species phospholipase C4	03g18010.1	phospholipase C
	72,924,083			9NG853100	NA	NA	NA	NA
				9NG853000	58440.1	FAD/NAD(P)-binding oxidoreductase family	03g12910.1	squalene monooxygenase

CONCLUSIONS

Investigation of wild plant–virus interactions has been hampered by a lack of wild plant–virus model systems for use in investigation. In many ways, several of the research efforts outlined in my dissertation represent concrete steps toward the development of SwMV and switchgrass as a model system for wild plant–virus research. In chapter one, I demonstrate that SwMV disease is a remarkably consistent predictor of infection in tillers that can be used as a marker of SwMV disease in the field. In chapter two, I outline a procedure for controlled inoculation study with leafhoppers and a procedure for inducing senescence that results in symptomatic infection.

Yet there are many aspects of this emerging model system that are still ‘wild,’ and future investigations could elucidate more tractable and specific ways to conduct controlled inoculations. What are the most efficient ways to inoculate switchgrass? Are there cost-effective mechanical techniques? Are there cost-effective techniques to collect and rear both infectious and non-infectious leafhopper vectors for viral- and mock-inoculation treatments? Answering these fundamental questions would do much to expand the utility of SwMV and switchgrass as a model system.

To date, the work outlined here likely represents the most comprehensive characterization of wild plant-virus interactions. Together, the results of chapters one and two suggest that SwMV is generally a moderately pathogenic virus that coexists with hosts by way of multiple mechanisms: (1) as a nonuniformly distributed and localized disease, hosts continued to grow and form new tillers; and (2) in the context of a local fungal pathogen outbreak, SwMV conferred fitness benefits to infected plants. Together with the results from the third chapter, which suggest that SwMV disease dynamics depend on switchgrass ecological subgroups and are

influenced by a large number and diverse genes, these results may indicate that SwMV is not a recently emerged virus; rather, this host–virus combination may have been co-evolving for extended evolutionary time.

Viruses are the most prolific biological entities on earth (Cobián Güemes *et al.*, 2016), and their ubiquity in unmanaged plant communities is undeniable (Edgar *et al.*, 2022). Given the demonstrated ability of viruses to impact, alter, and influence organisms from all kingdoms of life, wild viruses are almost certainly influencing wild plants and the ecosystems they inhabit in profound and unappreciated ways. With the costs of molecular diagnostics decreasing, and with bioinformatic tools and resources expanding, we now have an unprecedented ability to ask and answer questions about these interactions. Future investigation will certainly yield exciting breakthroughs and new discoveries regarding the ecology and evolution of these interactions, which occur at microscopic scales yet have global implications.

LITERATURE CITED

- Agindotan BO, Gray ME, Hammond RW, Bradley CA. 2012. Complete genome sequence of Switchgrass mosaic virus, a member of a proposed new species in the genus Marafivirus. *Archives of Virology* 157: 1825–1830.
- Agindotan B, Okanu N, Oladeinde A, Voigt T, Long S, Gray M, Bradley C. 2013a. Detection of Switchgrass mosaic virus in *Miscanthus* and other grasses. *Canadian Journal of Plant Pathology* 35: 81–86.
- Agindotan BO, Prasifka JR, Gray ME, Dietrich CH, Bradley CA. 2013b. Transmission of Switchgrass mosaic virus by *Graminella aureovittata*. *Canadian Journal of Plant Pathology* 35: 384–389.
- Aguilar E, del Toro FJ, Figueira-Galán D, Hou W, Canto T, Tenllado F. 2020. Virus infection induces resistance to *Pseudomonas syringae* and to drought in both compatible and incompatible bacteria–host interactions, which are compromised under conditions of elevated temperature and CO₂ levels. *Journal of General Virology* 101: 122–135.
- Alexander HM, Bruns E, Schebor H, Malmstrom CM. 2017. Crop-associated virus infection in a native perennial grass: reduction in plant fitness and dynamic patterns of virus detection. *Journal of Ecology* 105: 1021–1031.
- Alexander HM, Steets JA, Ali A. 2020. Distribution of *Asclepias* asymptomatic virus and exploration of possible effects on the wild plant host, *Asclepias viridis*. *Plant Health Progress* 21: 54–59.
- Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* 19: 535–544.
- Belliure B, Janssen A, Sabelis MW. 2008. Herbivore benefits from vectoring plant virus through reduction of period of vulnerability to predation. *Oecologia* 156: 797–806.
- Bernardo P, Charles-Dominique T, Barakat M, Ortet P, Fernandez E, Filloux D, Hartnady P, Rebelo TA, Cousins SR, Mesleard F, *et al.* 2018. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *The ISME Journal* 12: 173–184.
- Biedermann R, Achtziger R, Nickel H, Stewart AJA. 2005. Conservation of grassland leafhoppers: a brief review. *Journal of Insect Conservation* 9: 229–243.
- Bowsher AW, Benucci GMN, Bonito G, Shade A. 2020. Pentatricopeptide. *Phytobiomes Journal*: PBIOMES-07-20-0051-R.
- Brown JKM, Rant JC. 2013. Fitness costs and trade-offs of disease resistance and their consequences for breeding arable crops. *Plant Pathology* 62: 83–95.

- Brunt AA, Stace-Smith R. 1978. Some hosts, properties and possible affinities of a labile virus from *Hypochoeris radicata* (Compositae). *Annals of Applied Biology* 90: 205–214.
- Bushell M, Sarnow P. 2002. Hijacking the translation apparatus by RNA viruses. *J Cell Biol* 158: 395–399.
- Butković A, González R, Rivarez MPS, Elena SF. 2021. A genome-wide association study identifies *Arabidopsis thaliana* genes that contribute to differences in the outcome of infection with two Turnip mosaic potyvirus strains that differ in their evolutionary history and degree of host specialization. *Virus Evolution* 7: veab063.
- Callaway A, Giesman-Cookmeyer D, Gillock ET, Sit TL, Lommel SA. 2001. The multifunctional capsid proteins of plant RNA viruses. *Annual Review of Phytopathology* 39: 419–460.
- Casler MD. 2012. Switchgrass breeding, genetics, and genomics. In: Monti A, ed. *Green Energy and Technology*. Switchgrass: a valuable biomass crop for energy. London: Springer London, 29–53.
- Casler MD, Vogel KP, Taliaferro CM, Ehlke NJ, Berdahl JD, Brummer EC, Kallenbach RL, West CP, Mitchell RB. 2007. Latitudinal and longitudinal adaptation of Switchgrass populations. *Crop Science; Madison* 47: 2249–2260.
- Chang H-X, Brown PJ, Lipka AE, Domier LL, Hartman GL. 2016. Genome-wide association and genomic prediction identifies associated loci and predicts the sensitivity of Tobacco ringspot virus in soybean plant introductions. *BMC Genomics* 17: 153.
- Che Z, Yan H, Liu H, Yang H, Du H, Yang Y, Liu B, Yu D. 2020. Genome-wide association study for soybean mosaic virus SC3 resistance in soybean. *Molecular Breeding* 40: 69.
- Cobián Güemes AG, Youle M, Cantú VA, Felts B, Nulton J, Rohwer F. 2016. Viruses as winners in the game of life. *Annual Review of Virology* 3: 197–214.
- Cooper I, Jones RAC. 2006. Chapter 1 - Wild plants and viruses: under-investigated ecosystems. In: Maramorosch K, Shatkin AJ and Thresh JM, eds. *Plant Virus Epidemiology*. *Advances in Virus Research* 67: 1–47. Academic Press.
- Cortese LM, Honig J, Miller C, Bonos SA. 2010. Genetic diversity of twelve switchgrass populations using molecular and morphological markers. *BioEnergy Research* 3: 262–271.
- Crosslin JM, Thomas PE, Brown CR. 1999. Distribution of tobacco rattle virus in tubers of resistant and susceptible potatoes and systemic movement of virus into daughter plants. *American Journal of Potato Research* 76: 191.
- Crouch JA, Beirn LA, Cortese LM, Bonos SA, Clarke BB. 2009. Anthracnose disease of switchgrass caused by the novel fungal species *Colletotrichum navitas*. *Mycological Research* 113: 1411–1421.

- Cubry P, Pidon H, Ta KN, Tranchant-Dubreuil C, Thuillet A-C, Holzinger M, Adam H, Kam H, Chrestin H, Ghesquière A, et al. 2020. Genome wide association study pinpoints key agronomic QTLs in African rice *Oryza glaberrima*. *Rice* 13: 66.
- Culver JN, Padmanabhan MS. 2007. Virus-induced disease: altering host physiology one interaction at a time. *Annual Review of Phytopathology* 45: 221–243.
- Danial DL, Parlevliet JE. 1995. Effects of nitrogen fertilization on disease severity and infection type of yellow rust on wheat genotypes varying in quantitative resistance. *Journal of Phytopathology* 143: 679–681.
- Dawson WO, Garnsey SM, Tatineni S, Folimonova SY, Harper SJ, Gowda S. 2013. Citrus tristeza virus-host interactions. *Frontiers in Microbiology* 4.
- Dechaine JM, Burger JC, Chapman MA, Seiler GJ, Brunick R, Knapp SJ, Burke JM. 2009. Fitness effects and genetic architecture of plant-herbivore interactions in sunflower crop-wild hybrids. *The New Phytologist* 184: 828–841.
- Demirjian C, Vaillau F, Berthomé R, Roux F. 2023. Genome-wide association studies in plant pathosystems: success or failure? *Trends in Plant Science* 28: 471–485.
- Desmedt W, Vanholme B, Kyndt T. 2021. Chapter 5 - Plant defense priming in the field: a review. In: Maienfisch P, Mangelinckx S, eds. *Recent Highlights in the Discovery and Optimization of Crop Protection Products*. Academic Press, 87–124.
- Devadas R, Simpfendorfer S, Backhouse D, Lamb DW. 2014. Effect of stripe rust on the yield response of wheat to nitrogen. *The Crop Journal* 2: 201–206.
- Ding P, Ding Y. 2020. Stories of salicylic acid: a plant defense hormone. *Trends in Plant Science* 25: 549–565.
- Dreher TW, Miller WA. 2006. Translational control in positive strand RNA plant viruses. *Virology* 344: 185–197.
- Edgar RC, Taylor J, Lin V, Altman T, Barbera P, Meleshko D, Lohr D, Novakovsky G, Buchfink B, Al-Shayeb B, et al. 2022. Petabase-scale sequence alignment catalyses viral discovery. *Nature* 602: 142–147.
- Fernandez AR, Sáez A, Quintero C, Gleiser G, Aizen MA. 2021. Intentional and unintentional selection during plant domestication: herbivore damage, plant defensive traits and nutritional quality of fruit and seed crops. *New Phytologist* 231: 1586–1598.
- Ferri MB, Costes E, Quiot J-B, Dosba F. 2002. Systemic spread of Plum pox virus (PPV) in Mariana plum GF 8-1 in relation to shoot growth. *Plant Pathology* 51: 142–148.
- Fraile A, García-Arenal F. 2010. Chapter 1 - The coevolution of plants and viruses: resistance and pathogenicity. In: Carr JP, Loebeinstein G, eds. *Natural and Engineered Resistance to Plant Viruses, Part II. Advances in Virus Research* 76: 1–32. Academic Press.

- Fukuhara T, Tabara M, Koiwa H, Takahashi H. 2020. Effect of asymptomatic infection with southern tomato virus on tomato plants. *Archives of Virology* 165: 11–20.
- Funayama S, Hikosaka K, Yahara T. 1997. Effects of virus infection and growth irradiance on fitness components and photosynthetic properties of *Eupatorium makinoi* (Compositae). *American Journal of Botany* 84: 823–829.
- Funayama S, Terashima I, Yahara T. 2001. Effects of virus infection and light environment on population dynamics of *Eupatorium makinoi* (Asteraceae). *American Journal of Botany* 88: 616–622.
- Geyer CJ. 2021. Aster: Aster Models. R package.
- Geyer CJ, Wagenius S, Shaw RG. 2007. Aster models for life history analysis. *Biometrika* 94: 415–426.
- Gibbs A. 1980. A plant virus that partially protects its wild legume host against herbivores. *Intervirology* 13: 42–47.
- Grabowski PP, Evans J, Daum C, Deshpande S, Barry KW, Kennedy M, Ramstein G, Kaeppler SM, Buell CR, Jiang Y, *et al.* 2017. Genome-wide associations with flowering time in switchgrass using exome-capture sequencing data. *New Phytologist* 213: 154–169.
- Hanley ME, Lamont BB, Fairbanks MM, Rafferty CM. 2007. Plant structural traits and their role in anti-herbivore defence. *Perspectives in Plant Ecology, Evolution and Systematics* 8: 157–178.
- Hasiów-Jaroszewska B, Boezen D, Zwart MP. 2021. Metagenomic studies of viruses in weeds and wild plants: a powerful approach to characterise variable virus communities. *Viruses* 13: 1939.
- He Z, Webster S, He SY. 2022. Growth–defense trade-offs in plants. *Current Biology* 32: R634–R639.
- Herrera CM. 2009. *Multiplicity in Unity: Plant Subindividual Variation and Interactions with Animals*. Chicago: University of Chicago Press.
- Herrera CM. 2017. The ecology of subindividual variability in plants: patterns, processes, and prospects. *Web Ecology* 17: 51–64.
- Hoffman L, Chaves LM, Weibel EN, Mayton HS, Bonos SA. 2016. Impact of growing environment on anthracnose severity of switchgrass cultivars and clones. *Plant Disease* 100: 2034–2042.
- Honjo MN, Emura N, Kawagoe T, Sugisaka J, Kamitani M, Nagano AJ, Kudoh H. 2020. Seasonality of interactions between a plant virus and its host during persistent infection in a natural environment. *The ISME Journal* 14: 506–518.

- Ivanov KI, Mäkinen K. 2012. Coat proteins, host factors and plant viral replication. *Current Opinion in Virology* 2: 712–718.
- Jin E, Mendis GP, Sutherland JW. 2019. Integrated sustainability assessment for a bioenergy system: A system dynamics model of switchgrass for cellulosic ethanol production in the U.S. Midwest. *Journal of Cleaner Production* 234: 503–520.
- Jones R a. C, Fribourg CE. 1979. Host plant reactions some properties, and serology of wild potato mosaic virus [isolated from *Solanum chancayense* growing in the Peruvian coastal desert]. *Phytopathology (USA)*.
- Jridi C, Martin J-F, Marie-Jeanne V, Labonne G, Blanc S. 2006. Distinct viral populations differentiate and evolve independently in a single perennial host plant. *Journal of Virology* 80: 2349–2357.
- Kenaley SC, Bergstrom GC, Montes Ortiz ZK, Van Wallendael A, Lowry DB, Bonnette JE, Juenger TE. 2019. First report of the head smut fungus *Tilletia maclaganii* affecting switchgrass in Texas. *Plant Disease* 103: 578.
- Kenaley SC, Quan M, Aime MC, Bergstrom GC. 2018. New insight into the species diversity and life cycles of rust fungi (Pucciniales) affecting bioenergy switchgrass (*Panicum virgatum*) in the Eastern and Central United States. *Mycological Progress* 17: 1251–1267.
- Kominek P, Glasa M, Kominkova M. 2009. Analysis of multiple virus-infected grapevine plant reveals persistence but uneven virus distribution. *Acta Virologica* 53: 281–285.
- Lacroix C, Seabloom EW, Borer ET. 2017. Environmental nutrient supply directly alters plant traits but indirectly determines virus growth rate. *Frontiers in Microbiology* 8.
- Lee BD, Neri U, Roux S, Wolf YI, Camargo AP, Krupovic M, Simmonds P, Kyrpides N, Gophna U, Dolja VV, *et al.* 2023. Mining metatranscriptomes reveals a vast world of viroid-like circular RNAs. *Cell* 186: 646-661.e4.
- Lovell JT, MacQueen AH, Mamidi S, Bonnette J, Jenkins J, Napier JD, Sreedasyam A, Healey A, Session A, Shu S, *et al.* 2021. Genomic mechanisms of climate adaptation in polyploid bioenergy switchgrass. *Nature* 590: 438–444.
- Lowry DB, Behrman KD, Grabowski P, Morris GP, Kiniry JR, Juenger TE. 2014. Adaptations between ecotypes and along environmental gradients in *Panicum virgatum*. *The American Naturalist* 183: 682–692.
- Lowry DB, Lovell JT, Zhang L, Bonnette J, Fay PA, Mitchell RB, Lloyd-Reilley J, Boe AR, Wu Y, Rouquette FM, *et al.* 2019. QTL × environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. *Proceedings of the National Academy of Sciences* 116: 12933–12941.

- Lowry DB, Taylor SH, Bonnette J, Aspinwall MJ, Asmus AL, Keitt TH, Tobias CM, Juenger TE. 2015. QTLs for biomass and developmental traits in switchgrass (*Panicum virgatum*). *BioEnergy Research* 8: 1856–1867.
- Lowry DB, Willis JH. 2010. A Widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLOS Biology* 8: e1000500.
- Maclot F, Debie V, Malmstrom CM, Filloux D, Roumagnac P, Eck M, Tamisier L, Blouin AG, Candresse T, Massart S. 2023. Long-term anthropogenic management and associated loss of plant diversity deeply impact virome richness and composition of Poaceae communities. *Microbiology Spectrum* 0: e04850-22.
- Malmstrom CM, Alexander HM. 2016. Effects of crop viruses on wild plants. *Current Opinion in Virology* 19: 30–36.
- Malmstrom CM, Bigelow P, Trębicki P, Busch AK, Friel C, Cole E, Abdel-Azim H, Phillippo C, Alexander HM. 2017. Crop-associated virus reduces the rooting depth of non-crop perennial native grass more than non-crop-associated virus with known viral suppressor of RNA silencing (VSR). *Virus Research* 241: 172–184.
- Malmstrom CM, Busch AK, Cole EA, Trębicki P, Bernardo P, Brown AK, Landis DA, Werling BP. 2022. Emerging wild virus of native grass bioenergy feedstock is well-established in the midwestern USA and associated with premature stand senescence. *GCB Bioenergy* 14: 463–480.
- Malmstrom CM, Hughes CC, Newton LA, Stoner CJ. 2005a. Virus infection in remnant native bunchgrasses from invaded California grasslands. *The New Phytologist* 168: 217–230.
- Malmstrom CM, McCullough AJ, Johnson HA, Newton LA, Borer ET. 2005b. Invasive annual grasses indirectly increase virus incidence in California native perennial bunchgrasses. *Oecologia* 145: 153–164.
- Malmstrom CM, Stoner CJ, Brandenburg S, Newton LA. 2006. Virus infection and grazing exert counteracting influences on survivorship of native bunchgrass seedlings competing with invasive exotics. *Journal of Ecology* 94: 264–275.
- Márquez LM, Redman RS, Rodriguez RJ, Roossinck MJ. 2007. A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science* 315: 513–515.
- Mauck KE, Moraes CMD, Mescher MC. 2010. Deceptive chemical signals induced by a plant virus attract insect vectors to inferior hosts. *Proceedings of the National Academy of Sciences* 107: 3600–3605.
- McMillan C, Weiler J. 1959. Cytogeography of *Panicum virgatum* in central North America. *American Journal of Botany* 46: 590–593.

- van Molken T, de Caluwe H, Hordijk CA, Leon-Reyes A, Snoeren TAL, van Dam NM, Stuefer JF. 2012. Virus infection decreases the attractiveness of white clover plants for a non-vectoring herbivore. *Oecologia* 170: 433–444.
- Montes N, Vijayan V, Pagán I. 2020. Trade-offs between host tolerances to different pathogens in plant–virus interactions. *Virus Evolution* 6.
- Morozov SYu, Solovyev AG. 2003. Triple gene block: modular design of a multifunctional machine for plant virus movement. *Journal of General Virology* 84: 1351–1366.
- Morris GP, Grabowski PP, Borevitz JO. 2011. Genomic diversity in switchgrass (*Panicum virgatum*): from the continental scale to a dune landscape. *Molecular Ecology* 20: 4938–4952.
- Murphy AM, Jiang S, Elderfield JAD, Pate AE, Halliwell C, Glover BJ, Cunniffe NJ, Carr JP. 2023. Biased pollen transfer by bumblebees favors the paternity of virus-infected plants in cross-pollination. *iScience* 26: 106116.
- Mushegian A, Shipunov A, Elena SF. 2016. Changes in the composition of the RNA virome mark evolutionary transitions in green plants. *BMC Biology* 14: 68.
- Muthukumar V, Melcher U, Pierce M, Wiley GB, Roe BA, Palmer MW, Thapa V, Ali A, Ding T. 2009. Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. *Virus Research* 141: 169–173.
- Nam MH, Jeong SK, Lee YS, Choi JM, Kim HG. 2006. Effects of nitrogen, phosphorus, potassium and calcium nutrition on strawberry anthracnose. *Plant Pathology* 55: 246–249.
- Oliveira Silva A de, Aliyeva-Schnorr L, Wirsal SGR, Deising HB. 2022. Fungal pathogenesis-related cell wall biogenesis, with emphasis on the maize anthracnose fungus *Colletotrichum graminicola*. *Plants* 11: 849.
- Orians CM, Jones CG. 2001. Plants as resource mosaics: a functional model for predicting patterns of within-plant resource heterogeneity to consumers based on vascular architecture and local environmental variability. *Oikos* 94: 493–504.
- Pagán I, Garcia-Arenal F, Fraile A. 2016. Evolution of the interactions of viruses with their plant hosts. In: Weaver S, Denison M, Roossinck M, Vignuzzi M, eds. *Virus Evolution: Current Research and Future Directions*. Caister Academic Press, 127–153.
- Pagny G, Paulstephenraj PS, Poque S, Sicard O, Cosson P, Eyquard J-P, Caballero M, Chague A, Gourdon G, Negrel L, *et al.* 2012. Family-based linkage and association mapping reveals novel genes affecting Plum pox virus infection in *Arabidopsis thaliana*. *New Phytologist* 196: 873–886.

- Porter CL. 1966. An analysis of variation between upland and lowland switchgrass, *Panicum virgatum* L., in central Oklahoma. *Ecology* 47: 980–992.
- Prendeville HR, Ye X, Morris TJ, Pilson D. 2012. Virus infections in wild plant populations are both frequent and often unapparent. *American Journal of Botany* 99: 1033–1042.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* 38: 904–909.
- Privé F, Luu K, Blum MGB, McGrath JJ, Vilhjálmsson BJ. 2020. Efficient toolkit implementing best practices for principal component analysis of population genetic data. *Bioinformatics* 36: 4449–4457.
- R Core Team. 2021. R: A language and environment for statistical computing.
- Randles JW, Davies C, Hatta T, Gould AR, Francki RIB. 1981. Studies on encapsidated viroid-like RNA I. Characterization of velvet tobacco mottle virus. *Virology* 108: 111–122.
- Randles JW, Harrison BD, Roberts IM. 1976. Nicotiana velutina mosaic virus: purification, properties and affinities with other rod-shaped viruses. *Annals of Applied Biology* 84: 193–204.
- Randles JW, Steger G, Riesner D. 1982. Structural transitions in viroid-like RNAs associated with cadang-cadang disease, velvet tobacco mottle virus, and Solanum nodiflorum mottle virus. *Nucleic Acids Research* 10: 5569–5586.
- Ravindran MS, Bagchi P, Cunningham CN, Tsai B. 2016. Opportunistic intruders: how viruses orchestrate ER functions to infect cells. *Nature Reviews Microbiology* 14: 407–420.
- Rodríguez-Nevado C, Montes N, Pagán I. 2017. Ecological factors affecting infection risk and population genetic diversity of a novel potyvirus in its native wild ecosystem. *Frontiers in Plant Science* 8.
- Roossinck MJ. 2011. The good viruses: viral mutualistic symbioses. *Nature Reviews Microbiology* 9: 99–108.
- Roossinck MJ. 2012. Plant virus metagenomics: biodiversity and ecology. *Annual Review of Genetics* 46: 359–369.
- Roossinck MJ. 2015a. A new look at plant viruses and their potential beneficial roles in crops: Beneficial viruses for crops. *Molecular Plant Pathology* 16: 331–333.
- Roossinck MJ. 2015b. Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. In: Schultz-Cherry S, ed. *Journal of Virology* 89: 6532–6535.

- Roossinck MJ, Saha P, Wiley GB, Quan J, White JD, Lai H, Chavarría F, Shen G, Roe BA. 2010. Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology* 19: 81–88.
- Roscher C, Schumacher J, Gubsch M, Lipowsky A, Weigelt A, Buchmann N, Schmid B, Schulze E-D. 2012. Using plant functional traits to explain diversity–productivity relationships. *PLOS ONE* 7: e36760.
- Rossi EA, Ruiz M, Bonamico NC, Balzarini MG. 2020. Genome-wide association study of resistance to Mal de Río Cuarto disease in maize. *Agronomy Journal* 112: 4624–4633.
- Rubio B, Cosson P, Caballero M, Revers F, Bergelson J, Roux F, Schurdi-Levraud V. 2019. Genome-wide association study reveals new loci involved in *Arabidopsis thaliana* and Turnip mosaic virus (TuMV) interactions in the field. *New Phytologist* 221: 2026–2038.
- Ryskamp M, Geyer CJ. 2023. Aster-model-for-three-year-switchgrass-fitness: Initial Release (v1.0.0).
- Sakschewski B, von Bloh W, Boit A, Poorter L, Peña-Claros M, Heinke J, Joshi J, Thonicke K. 2016. Resilience of Amazon forests emerges from plant trait diversity. *Nature Climate Change* 6: 1032–1036.
- Sallinen S, Norberg A, Susi H, Laine A-L. 2020. Intraspecific host variation plays a key role in virus community assembly. *Nature Communications* 11: 5610.
- Sanderson MA. 2008. Upland switchgrass yield, nutritive value, and soil carbon changes under grazing and clipping. *Agronomy Journal* 100: 510–516.
- Schmid CJ, Clarke BB, Murphy JA. 2018. Potassium nutrition affects anthracnose on annual bluegrass. *Agronomy Journal* 110: 2171–2179.
- Schneider RJ, Mohr I. 2003. Translation initiation and viral tricks. *Trends in Biochemical Sciences* 28: 130–136.
- Schrotenboer AC, Allen MS, Malmstrom CM. 2011. Modification of native grasses for biofuel production may increase virus susceptibility. *GCB Bioenergy* 3: 360–374.
- Schuh MK, Bahlai CA, Malmstrom CM, Landis DA. 2019. Effect of switchgrass ecotype and cultivar on establishment, feeding, and development of fall armyworm (Lepidoptera: noctuidae). *Journal of Economic Entomology* 112: 440–449.
- Serba DD, Uppalapati SR, Mukherjee S, Krom N, Tang Y, Mysore KS, Saha MC. 2015. Transcriptome profiling of rust resistance in switchgrass using RNA-Seq analysis. *The Plant Genome* 8: plantgenome2014.10.0075.
- Shates TM, Sun P, Malmstrom CM, Dominguez C, Mauck KE. 2019. Addressing research needs in the field of plant virus ecology by defining knowledge gaps and developing wild dicot study systems. *Frontiers in Microbiology* 9.

- Shaw RG, Geyer CJ, Wagenius S, Hangelbroek HH, Etterson JR. 2008. Unifying life-history analyses for inference of fitness and population growth. *The American Naturalist* 172: E35–E47.
- Shrestha V, Chhetri HB, Kainer D, Xu Y, Hamilton L, Piasecki C, Wolfe B, Wang X, Saha M, Jacobson D, *et al.* 2022. The genetic architecture of nitrogen use efficiency in switchgrass (*Panicum virgatum* L.). *Frontiers in Plant Science* 13.
- Siefert A, Violle C, Chalmandrier L, Albert CH, Taudiere A, Fajardo A, Aarssen LW, Baraloto C, Carlucci MB, Cianciaruso MV, *et al.* 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters* 18: 1406–1419.
- Sitonik C, Suresh LM, Beyene Y, Olsen MS, Makumbi D, Oliver K, Das B, Bright JM, Mugo S, Crossa J, *et al.* 2019. Genetic architecture of maize chlorotic mottle virus and maize lethal necrosis through GWAS, linkage analysis and genomic prediction in tropical maize germplasm. *Theoretical and Applied Genetics* 132: 2381–2399.
- Skotnicki ML, Mackenzie AM, Torronen M, Brunt AA, Gibbs AJ. 1992. Cardamine chlorotic fleck virus, a new carmovirus from the Australian Alps. *Australasian Plant Pathology* 21: 120–122.
- Stam R, McDonald BA. 2018. When resistance gene pyramids are not durable—the role of pathogen diversity. *Molecular Plant Pathology* 19: 521–524.
- Stanton-Geddes J, Shaw RG, Tiffin P. 2012. Interactions between soil habitat and geographic range location affect plant fitness. *PLOS ONE* 7: e36015.
- Stewart CL, Pyle JD, Jochum CC, Vogel KP, Yuen GY, Scholthof K-BG. 2015. Multi-year pathogen survey of biofuel switchgrass breeding plots reveals high prevalence of infections by *Panicum mosaic virus* and its satellite virus. *Phytopathology* 105: 1146–1154.
- Susi H, Filloux D, Frilander MJ, Roumagnac P, Laine A-L. 2019. Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ* 7: e6140.
- Susi H, Laine A-L. 2021. Agricultural land use disrupts biodiversity mediation of virus infections in wild plant populations. *New Phytologist* 230: 2447–2458.
- Sutherland J, Bell T, Trexler RV, Carlson JE, Lasky JR. 2022. Host genomic influence on bacterial composition in the switchgrass rhizosphere. *Molecular Ecology* 31: 3934–3950.
- Thivierge K, Nicaise V, Dufresne PJ, Cotton S, Laliberté J-F, Gall OL, Fortin MG. 2005. Plant virus RNAs. Coordinated recruitment of conserved host functions by (+) ssRNA Viruses during early infection events. *Plant Physiology* 138: 1822–1827.

- Thottappilly G, van LENT JWM, Rossel HW, Sehgal OP. 1992. Rottboellia yellow mottle virus, a new sobemovirus affecting Rottboellia cochinchinensis (Itch grass) in Nigeria. *Annals of Applied Biology* 120: 405–415.
- Torrico AK, Salazar SM, Kirschbaum DS, Conci VC. 2018. Yield losses of asymptomatic strawberry plants infected with Strawberry mild yellow edge virus. *European Journal of Plant Pathology* 150: 983–990.
- Uppalapati SR, Serba DD, Ishiga Y, Szabo LJ, Mittal S, Bhandari HS, Bouton JH, Mysore KS, Saha MC. 2013. Characterization of the rust fungus, Puccinia emaculata, and evaluation of genetic variability for rust resistance in switchgrass populations. *BioEnergy Research* 6: 458–468.
- USDA. 2011. Fact sheet for release of Cave-In-Rock switchgrass (*Panicum virgatum* L.).
- VanWallendael A, Benucci GMN, Costa PB da, Fraser L, Sreedasyam A, Fritschi F, Juenger TE, Lovell JT, Bonito G, Lowry DB. 2022. Host genotype controls ecological change in the leaf fungal microbiome. *PLOS Biology* 20: e3001681.
- VanWallendael A, Bonnette J, Juenger TE, Fritschi FB, Fay PA, Mitchell RB, Lloyd-Reilley J, Rouquette Jr FM, Bergstrom GC, Lowry DB. 2020. Geographic variation in the genetic basis of resistance to leaf rust between locally adapted ecotypes of the biofuel crop switchgrass (*Panicum virgatum*). *New Phytologist* 227: 1696–1708.
- Vincent SJ, Coutts BA, Jones RAC. 2014. Effects of Introduced and indigenous viruses on native plants: exploring their disease causing potential at the agro-ecological interface. *PLOS ONE* 9: e91224.
- Vogel KP, Hopkins AA, Moore KJ, Johnson KD, Carlson IT. 1996. Registration of ‘Shawnee’ switchgrass. *Crop Science* 36: 1736.
- Wallner AM, Molano-Flores B, Dietrich CH. 2013. Using Auchenorrhyncha (Insecta: Hemiptera) to develop a new insect index in measuring North American tallgrass prairie quality. *Ecological Indicators* 25: 58–64.
- Wang Y, Vik JO, Omholt SW, Gjuvsland AB. 2013. Effect of regulatory architecture on broad versus narrow sense heritability. *PLoS Computational Biology* 9: e1003053.
- Webster CG, Coutts BA, Jones R a. C, Jones MGK, Wylie SJ. 2007. Virus impact at the interface of an ancient ecosystem and a recent agroecosystem: studies on three legume-infecting potyviruses in the southwest Australian floristic region. *Plant Pathology* 56: 729–742.
- Wei K, Gibbs A, Mackenzie A. 2012. Clitoria yellow mottle virus: a tobamovirus from Northern Australia. *Australasian Plant Disease Notes* 7: 59–61.
- Westwood JH, Mccann L, Naish M, Dixon H, Murphy AM, Stancombe MA, Bennett MH, Powell G, Webb AAR, Carr JP. 2013. A viral RNA silencing suppressor interferes with

- abscisic acid-mediated signalling and induces drought tolerance in *Arabidopsis thaliana*. *Molecular Plant Pathology* 14: 158–170.
- Wetzel WC, Thaler JS. 2016. Does plant trait diversity reduce the ability of herbivores to defend against predators? The plant variability–gut acclimation hypothesis. *Current Opinion in Insect Science* 14: 25–31.
- Williams CJ, Li Z, Harvey N, Lea RA, Gurd BJ, Bonafiglia JT, Papadimitriou I, Jacques M, Croci I, Stensvold D, *et al.* 2021. Genome wide association study of response to interval and continuous exercise training: the Predict-HIIT study. *Journal of Biomedical Science* 28: 37.
- Wu X, Valli A, García JA, Zhou X, Cheng X. 2019. The tug-of-war between plants and viruses: great progress and many remaining questions. *Viruses* 11: 203.
- Xu P, Chen F, Mannas JP, Feldman T, Sumner LW, Roossinck MJ. 2008. Virus infection improves drought tolerance. *New Phytologist* 180: 911–921.
- Yang S, Li J, Zhang X, Zhang Q, Huang J, Chen J-Q, Hartl DL, Tian D. 2013. Rapidly evolving R genes in diverse grass species confer resistance to rice blast disease. *Proceedings of the National Academy of Sciences* 110: 18572–18577.
- Yoo RH, Lee S-W, Lim S, Zhao F, Igori D, Baek D, Hong J-S, Lee S-H, Moon JS. 2017. Complete genome analysis of a novel umbravirus-polerovirus combination isolated from *Ixeridium dentatum*. *Archives of Virology* 162: 3893–3897.
- Zhang L, MacQueen A, Weng X, Behrman KD, Bonnette J, Reilley JL, Rouquette FM, Fay PA, Wu Y, Fritschi FB, *et al.* 2022. The genetic basis for panicle trait variation in switchgrass (*Panicum virgatum*). *Theoretical and Applied Genetics* 135: 2577–2592.
- Zhang Y, Zalapa JE, Jakubowski AR, Price DL, Acharya A, Wei Y, Brummer EC, Kaeppler SM, Casler MD. 2011. Post-glacial evolution of *Panicum virgatum*: centers of diversity and gene pools revealed by SSR markers and cpDNA sequences. *Genetica* 139: 933.
- Zhang J, Zhou J-M. 2010. Plant immunity triggered by microbial molecular signatures. *Molecular Plant* 3: 783–793.