Viromes of Agriculturally Important Insects in Alabama

By Aasma Sharma

A thesis submitted to the Graduate Faculty of Auburn University in partial fulfillment of the requirements for the Degree of Master of Science

Auburn, Alabama
April 21, 2025

Approved by
Dr. Kathleen M. Martin Assistant Professor of Entomology and Plant Pathology
Dr. Chaoyang Zhao, Research Entomologist, National Soil Dynamics Laboratory, USDA-ARS
Dr. Alana Jacobson, Associate Professor of Entomology and Plant Pathology
Dr. Scott Graham, Assistant Professor and Extension Specialist of Entomology and Plant Pathology
Dr. Jason Upton, Associate Professor of Department of Biological Sciences

Abstract

Insect vectors play a critical role in the transmission of viruses, posing significant threats to both agriculture and public health. The studies in this thesis, utilize RNA sequencing (RNA-seq) to characterize the viromes of two key insect species—Soybean thrips (*Neohydatothrips variabilis* Beach), and Tobacco thrips (*Frankliniella fusca* Hinds)—collected from different regions of Alabama. Analysis of *N. variabilis* specimens from Auburn (2023) identified 15 viruses spanning seven viral orders, including ten potential insect-infecting viruses, one plant-insect virus, one strictly plant-infecting virus, and three novel mycoviruses. In *F. fusca* collected from peanut and clover in Headland (2022–2023), 25 distinct viruses were detected, with greater viral diversity observed in thrips associated with clover. Bunyavirales was the most prevalent viral order, and *Orthotospovirus tomatomaculae* was consistently present across samples. These findings provide a comprehensive overview of viral diversity in key insect vectors at the state level, offering critical insights into virus-host interactions, viral evolution, and potential emerging threats to crops.

Keywords: RNA-seq, virome, tobacco thrips, soybean thrips

Acknowledgments

This journey has been made possible with the support of my professor, mentors, mother, and friends. I would like to express my deep gratitude to my supervisor, Dr. Kathleen Marteen, for her unwavering support and encouragement throughout my pursuit of this degree, as well as for motivating me to develop new skills, ask questions, and overcome challenges. I am thankful to Dr. Chaoyang Zhao, and Dr. Caesar Escalante for helping me to start learning the technical skills necessary for the project and always being available to answer my inquiries. I also appreciate Dr. Alana and Dr. Chaoyang for their assistance in collecting the field samples. My sincere thanks go to my committee members, Dr. Alana Jacobson, Dr. Chaoyang Zhao, Dr. Scott Graham, and Dr. Jason Upton. I am grateful to my undergraduate student, Cale Johnson, whom I mentored, for his support with my project. A special thank you to my lab mates, Rachel Livingston and Abdelaal Shehata, for their continuous help.

This degree has provided me with invaluable firsthand research experience that was not possible in my home country and was interrupted by COVID. I believe I have gained many skills through this program that have not only improved me professionally but also personally. I am forever thankful to everyone who played a direct or indirect role in my journey. Lastly, I am deeply grateful to my mother, who has been a constant source of motivation throughout my life.

Table of contents

List of Tables

Chapter 3

Supplemental chapter 1

List of Figures

Chapter 1

Chapter 2

Chapter 3

Supplemental Chapter 1

List of Abbreviations

NCBI: National Center for Biotechnology Information

ICTV: International Committee on Taxonomy of Viruses

TSWV: Tomato Spotted Wilt Virus

SVNV: Soybean Vein Necrosis Virus

Bp: base pair

RNA-seq: Ribo-Nucleic Acid Sequencing

TSWV: Tomato Spotted Wilt Virus

SVNV: Soybean Vein Necrosis Virus

Chapter 1

Introduction

Background

By definition, a vector is an organism that carries a disease-causing agent, like viruses, important bacteria, phytoplasmas, etc., and facilitates its spread among hosts (Pan, 2024). Specific examples of the major insect vectors that transmits agriculturally important pathogens are the Hemipterans (aphids, whiteflies, cicadas, spittlebugs, leafhoppers, assassin bugs, stink bugs, lygaeid bugs), Dipterans (Mosquitoes, sandflies, biting midges, blackflies), Coleopterans (Chrysomelidae), and Thysanopterans (e.g., thrips) (Wendimu & Gurmu, 2024) and Hymenopterans. Major vectors that transmit viruses in the agricultural system are aphids, whiteflies, leafhoppers, planthoppers, and thrips. The mechanisms for transmission include both active and passive means of moving pathogens from plant to plant. Passive transmission encompasses methods such as the movement of pollen or the movement of particles attached to hairs on the body (Card et al., 2007). Active transmission can be grouped into multiple categories, which will be discussed later.

The majority of the work presented in this thesis will focus on thrips (Thysanoptera) which are phytophagous, predaceous, or mycophagous insects with rasping and sucking types of mouthparts (Omkar, 2021). Morphologically, thrips are minute with a slender-bodied (length between 1/32 to 1/5 inch) and fringed wings. Thrips are considered hemimetabolous (incomplete metamorphosis) with four life stages between egg and larva with a life cycle of between 2-4 weeks depending on season and temperature. They are active during the summer months, with multiple generations per year (*Soybean Thrips*, n.d.). As important pests in the agriculture system, thrips feed on plants by inserting their tubular styles into cells and sucking out cellular contents (Hunter

& Ullman, 1989). While feeding, thrips first scrape the leaf surface, rupture the cell, and feed on the plant sap (Stafford et al., 2012). As they do not have separate food and salivary canals in their stylets, and they must salivate and ingest cell contents through the same feeding tube (Harrewijn et al., 1996). Thrips are efficient in transmitting viruses because this type of feeding can inject viruses into these cells (Stafford et al., 2012). They are actively feeding on the leaf's underside along the vein and cause feeding damage in the leaves, referred to as stippling (*Thrips Management Guidelines--UC IPM*, n.d.), and the leaf may also look crinkled and mottled. Thrips can transmit viruses of different genera such as *Tospovirus, Ilarvirus, Carmovirus, Sobemovirus,* and *Machlomovirus* (Jones, 2005). This thesis focuses on identifying viruses within insects, and it is expected that these insects will harbor both insect-specific viruses and agriculturally important viruses. This thesis will also primarily focus on two species of thrips: Tobacco thrips (*Frankliniella fusca* Hinds) and Soybean thrips (*Neohydatothrips variabilis* Beach). This study aims to identify the diversity of viruses found in these two species of thrips in Alabama: those that are insect-specific viruses, those that may be present by chance, and those that these insects transmit. This study was conducted with RNA-sequencing to obtain as thorough and complete a survey of all viruses present in the insects as possible.

Insect vectors are both hosts and vectors for viruses utilizing different mechanisms

Viruses have evolved a range of transmission strategies that influence their persistence and spread within host populations. At the broadest level, virus transmission can occur through horizontal or vertical pathways. Horizontal transmission involves the transfer of viruses between individuals of the same generation and can occur via direct contact, environmental exposure, or through biological vectors such as insects. In contrast, vertical transmission refers to the inheritance of viruses from parent to offspring, often through infected reproductive tissues, enabling viruses to persist within host lineages across generations (Y. Chen et al., 2006).

Viruses can be transmitted by insect vectors through either passive or active mechanisms. In passive transmission, the virus is mechanically transferred, often externally carried on the vector's mouthparts or body surface, without engaging in any biological interaction with the vector itself. This process lacks specificity and does not involve the vector's internal systems (Dietzgen et al., 2016). By comparison, active transmission requires a more co-evolved association between the virus and vector. Some viruses replicate within the vector, and many rely on specific interactions with vector tissues, such as receptor-mediated binding or utilization of vector proteins, to facilitate movement and delivery to the plant host. (Dietzgen et al., 2016).

Vector-mediated transmission, particularly by insects such as aphids, whiteflies, and thrips, is a major route for virus spread in agricultural systems. These vectors employ a range of mechanisms to transmit viruses, including non-circulative, circulative, and propagative modes, each defined by how the virus interacts with and moves through the insect's body. A detailed understanding of these transmission processes is essential for anticipating virus emergence and designing effective management strategies in crop protection.

Vertical vs Horizontal transmission of virus

Vertical transmission of viruses: The transmission of viruses occurs through the parent to progeny. Example: Rice stripe virus (RSV), a Tenuivirus, is primarily spread by the small brown planthopper (*Laodelphax striatellus)* and can also be transmitted to its offspring through the egg (The PLOS Pathogens Staff, 2014).

Horizontal transmission of viruses: In Horizontal transmission, the viruses spread among the individuals of the same species through different routes like direct contact, exposure to the environmental conditions, or through vectors like insects. According to International Committee on Taxonomy of Viruses, approx. 76% of the known plant viruses are transmitted by insect species in the order Hemiptera (*Common Types of Viruses of Floriculture Crops and Their Modes of Transmission*, 2014). This can occur through various mechanisms, categorized as non-persistent transmission, semi-persistent transmission, and persistent transmission. Each category is further divided into subcategories, where non-persistent and semi-persistent include helper component-mediated versus coat protein-mediated, and for persistent viruses, circulative propagative or non-propagative (Whitfield et al., 2015a). For all types of transmission, host proteins may be involved in transmission such as helper component proteins or receptors at the membranes of gut lumens (Hogenhout et al., 2008; Hýsková et al., 2024). In the following sections and in Figure 1, this thesis will cover each type of active transmission mechanisms and what is known about how it occurs in the insect vector.

Non-persistent transmission

It is non-circulative transmission; the virus is retained in the mouthpart (stylet) and does not reach to the gut of the insect. It is also called stylet borne transmission (Figure 3). It is retained from a few minutes to hours. It is further sub-classified into two types of transmission: capsid-only or helper-dependent. In capsid-only transmission, viral coat protein or conserved capsid surface domains are required for the transmission. For example: transmission of cucumber mosaic virus (CMV) by aphids (Pirone & Megahed, 1966). In helper-component mediated transmission, a helper protein is required to bind the virus to the style of the insect. For example: In the potyvirus, they encode for the helper protein; helper component-proteinase (HC-Pro) that acts as a bridge between potyvirus CP and aphid protein(s) in the stylet (Dietzgen et al., 2016).

Semi persistent transmission

It is non-circulative transmission the virus does not replicate in the hemolymph of virus. The virus reaches the foregut after ingestion and can be retained from hours to weeks before it is released back through the mouth. It also could be coat protein mediated or helper component mediated. Example of coat protein mediated transmission is: Transmission of lettuce infectious yellows virus (LIYV) transmission by whitefly vector *Bemisia tabaci* (A. Y. S. Chen et al., 2011; Stewart et al., 2010). Example of helper component mediated transmission is transmission of CaMV and other caulimoviruses. by *Myzus persicae* requires a helper component (Dietzgen et al., 2016).

Circulative transmission

Circulative transmission represents a persistent form of virus movement within the insect vectors and is further classified into non-propagative and propagative types. In both cases, the virus crosses internal barriers within the insect, such as the gut and salivary glands—and can remain in the vector’s body for its entire lifespan. However, whether the virus replicates within the insect distinguishes the two forms. The following sections will explore these circulative mechanisms in more detail, highlighting their differences and biological significance.

Circulative, non-propagative transmission

In this form of transmission, the virus enters the insect's digestive tract to reach the midgut or hindgut, depending on the virus. Although it does not replicate within these tissues, the virions bind to specific receptors on the gut epithelium and are transported across the gut barrier via transcytosis (endocytosis) to proceed through the vector's circulatory system (Dietzgen et al., 2016). Once in the hemolymph, the virus circulates until it reaches the salivary glands, which serve as a secondary barrier to transmission. To be successfully transmitted, the virus must cross the salivary gland membranes and accumulate in the saliva. From there, it can be introduced into plant tissues during subsequent feeding events. The latent period varies from days to weeks, as the virus needs to circulate the system until it arrives at the salivary glands before it can be transmitted. Example: Transmission of Cotton leaf dwarf virus(CLRDV) by aphids. In this, coat protein and coat protein-Read Through Domain (RTD) are believed to be necessary for the transmission (Brault et al., 2000; Bruyère et al., 1997; Dietzgen et al., 2016; Peter et al., 2008).

Circulative, propagative transmission

The virus reaches the mid-gut and hindgut, and the virus replicates across the insect tissues and reach to the salivary gland. The latent period could be up to a week to a couple of weeks. The transmission of *Orthotospovirus*es like Tomato spotted wilt virus and Soybean Mosaic viruses by tobacco thrips (*Frankliniella fusca* Hinds) and soybean thrips (*Neohydatothrips variabilis* Beach) follow this type of transmission. The virus is acquired by the larvae and is retained throughout the adult stage of insects. The virus replicates in insects and is released back through stylet. Unlike the non-propagative transmission, in propagative transmission, the titer of the virus will continue to increase in the insect throughout its lifetime until death.

RNA Sequencing as a Tool for Virus Discovery

In the past, conventional methods of virus discovery included the use of electron microscopy, serological methods like Enzyme-Linked Immunosorbent Assay (ELISA), Nucleic acid hybridization (Southern blot, northern blot, and dot blot), and other molecular methods like PCR, RT-PCR followed by cloning and sequencing, analysis of expressed sequence tag (EST) libraries, and microarray methods (Liu et al., 2011). All these techniques rely on previous knowledge of which viruses may be present and do not provide any insights into completely novel viruses. Due to this, there is limited knowledge of insect-infecting viruses, which in the past have been diagnosed based on physical symptoms present in the insect, such as deformed wing virus in honeybees (Lanzi et al., 2006). As there are examples of insect viruses that do not cause any symptoms (Burand et al., 2012) that can be measured in their hosts without extremely close examination, many viruses in insects could go unidentified. These unidentified viruses have the potential to pose a risk in the future if they continue to change and interact with other viruses in their environment (Lawrence et al., 2023). Moreover, these techniques do not have the capacity to monitor for mutation and what other unknown viruses may be present. RNA sequencing (RNA-seq) is a modern technology that identifies both known and unknown viruses. It is a high-throughput technique for analyzing all the transcripts present in a cell. RNA-seq is a recently developed method for transcriptome profiling that employs deep-sequencing technologies (Wang et al., 2009). The workflow for RNA-seq includes converting a population of RNA (total or fractionated, such as poly(A)+) into a library of cDNA fragments with adaptors attached to one (single-end sequencing) or both ends (paired-end sequencing). Then each molecule is sequenced in a high-throughput manner to obtain short sequences, where the read lengths typically range from 30 to 400 bp (Wang et al., 2009). For a virus diversity study, quality testing is conducted initially, followed by the potential mapping of the reads to the reference genome to eliminate host RNA. The unmapped reads are then assembled into larger sequences (contigs), and the contigs are further annotated (Figure 1, Figure 2).

Challenges of RNA-seq in Viral discovery compared to conventional virus detection methods

Although there are many advantages of virus discovery through RNA-seq, there are challenges as well. The presence of endogenous viral elements (EVEs) with significant similarity to the exogenous viruses, might impact the viral discovery by RNA-seq (Rozo-Lopez et al., 2023). This might result in false virus discovery. Similarly, the analysis of the transcripts is based on the comparison to already available viruses in the NCBI database (either the total nucleotide database or viral database specifically), it imposes a challenge to detect completely new viruses that have never been reported in the database resulting in the detection bias (Cobbin et al., 2021). Similarly, some viral constructs might be annotated to the genome of the insects or *vice versa*. This may be due to the presence of genetically integrated viral elements within the insect genome or the mis annotation of contigs as viral sequences when they are part of the insect's own genetic material (Brait et al., 2024). Similarly, large scale transcriptomics studies like RNA-seq are based on artificially constructed contigs, thus validation of these contigs may be necessary to determine the true sequences of viral genomes or genes present. There are multiple tools used in assembly and to test the quality of the viruses involved in RNA-seq studies. The length and quality of the contigs might be impacted by the choice of analysis tools. Once identified, RNA-seq analysis cannot provide information as to how the insect acquired the virus, requiring more future work to determine this. Although RNA-seq is becoming increasingly cost-effective, it still requires considerable expertise in bioinformatics for proper analysis and interpretation. Nevertheless, it is a powerful tool for revealing the virome diversity of major agricultural insect pests, many of which may pose significant economic threats in the future. The importance of early and accurate virus detection cannot be overstated. There have been several historical cases where delayed or inaccurate detection of viruses led to significant consequences. For instance, during the Nipah virus outbreak in Malaysia (September 1998 to May 1999), the respiratory virus was initially misidentified as Japanese encephalitis. As a result, early control measures were ineffective, leading to 265 cases of acute encephalitis with 105 deaths, a near collapse of Malaysia's billion-dollar pig-farming industry, and the spread of the outbreak to other parts of Malaysia and Singapore (Looi & Chua, 2007). This kind of scenario is highly plausible in the context of crop viral diseases, given the ever-changing nature of viruses. Therefore, having techniques that not only reveal the current virome status but also help forecast potential shifts in virus populations within agriculturally important insect vectors is essential (Wu et al., 2020).

Existing knowledge on virome on Agricultural vectors and Thrips species

The bioinformatic analysis of sequences from next-generation sequencing (NGS) technology has helped us identify RNA viruses and transcripts of DNA viruses over the years. Currently, there are more than 800 RNA viruses belonging to 39 families most of them are from *Iflaviridae, Dicistroviridae, Partitiviridae,* and *Rhabdoviridae* that are stored in the NCBI database from different agriculturally important insect orders like Hymenoptera, Hemiptera, Thysanoptera, Lepidoptera, Diptera, Coleoptera, and Orthoptera (Qi et al. 2023). A study done in the Midwest of the United States in 2018 found 181 viral-like sequences among which 155 were novel and arthropod infecting viruses, with similarity to plant and fungus-infecting viruses in Soybean thrips (Thekke-Veetil et al., 2020). Most of the viruses from this study that are the best understood are the plant viruses, which undergo different types of transmission (described below).

Two species of thrips in Alabama, Tobacco and Soybean thrips

Tobacco thrips, *Frankliniella fusca* Hinds, has a wide but not global distribution, primarily associated with crops like peanuts, cotton, and tobacco in temperate and subtropical areas (LaTora et al., 2022a). With a narrow body, females are usually brown or black, and the males are yellow, having eight segmented antennae ("Tobacco Thrips," 2025). It transmits the economically damaging plant virus *Orthotospovirus tomatoculae* (Tomato spotted wilt virus, TSWV) and is considered an early season vector of importance on peanuts (Culbreath et al., 2003) in Alabama compared to the later season and more well-known thrips vector *Franklinella occidentalis* (western flower thrips) (Culbreath et al., 2003). Due to TSWV, in Georgia, from the last 10 years' data (1996-2006), the estimated annual average losses are $12.3 million in peanut, $11.3 million in tobacco, and $9 million in tomato and pepper for a total of $326 million (*History and Economic Impact*, 2025). However, the diversity of viruses present in this study suggests that there is little that is understood about how many viruses are present in the ecosystem, both those that infect insects and those that are plant-infecting and insect-infecting.

Soybean thrips (*N. variabilis*) feed on the hosts: smartweed, cucumber (Beach, 1897), cotton, lima bean, soybean, and tomato (Nault & Speese Iii, 2002), kudzu plant (Thekke-Veetil et al., 2020b). It is an important vector of another orthotospovirus, *Orthotospovirus glycinecrovenae* (Soybean vein necrosis virus, SVNV) in soybeans (Zhou & Tzanetakis, 2013). In Alabama, SVNV was first reported in Limestone County in 2012 (Conner et al., 2013). The symptoms begin with yellowing along the leaf vein, which later turns reddish-brown as irregular lesions spread outward, eventually leading to tissue death (*Soybean Vein Necrosis Virus - Soybean Disease - Soybean Research & Information Network - SRIN*, n.d.). Currently, the leaf death does not impact yields as it occurs primarily later in the growing season, however, the incidence of SVNV in soybean are increasing over time (Sikora et al., 2018). To more fully characterize the viruses present in these two thrips species to fully understand the diversity, we conducted an RNA sequencing study.

Significance of This Research: Understanding Its Impact

The majority of plant-infecting viruses belonging to *Caulimovirus, Crinivirus, Luteovirus, Geminiviridae, Reovirus, Tospovirus,* and *Tenuivirus* are transmitted to their host plants by vectors

(Whitfield et al., 2015b). There is a significant amount of knowledge about these genera compared to insect-infecting viruses and other potential viruses that remain unidentified in insect systems. Insects are commonly infected with multiple viruses, including those that cause sublethal, asymptomatic, and latent infections (Liu et al., 2011). These insect-specific viruses might be crucial in controlling the population of insect species. As a result, their potential could be exploited as biocontrol agents.

There is little knowledge of the virus diversity study in different areas and their ecological impact. There is a knowledge gap in understanding the potential insect infecting viruses in the system of insect vectors. Unlike vector-borne or plant viruses, these insect-specific viruses (ISVs) are incapable of infecting vertebrates or plants and have been classified into families such as *Togaviridae*, *Reoviridae*, *Rhabdoviridae*, *Dicistroviridae*, *Flaviviridae*, *Iflaviridae*, *Aliusviridae*, *Chuviridae*, and *Negevirus* (Looi & Chua, 2007). Consequently, they have not been previously identified as significant insect vectors. Research focused on detecting these viruses in agriculturally important insect vectors has been limited. There is a lack of knowledge regarding whether the presence of other viruses might influence the transmission of plant-infecting viruses. Additionally, there is an unclear understanding of how the presence of some viruses might otherwise affect other viruses. Thus, this study offers a thorough examination of virus diversity in various insect species, primarily Soybean thrips and Tobacco thrips.

The previously described study in the Midwest on soybean thrips helps us uncover the diversity of viruses other than plant-infecting viruses, but this sequencing was done in a pooled population of Soybean thrips, and no specific state-wide or county-wide sequencing was conducted. This led to the identification of a variety of viruses; however, no real knowledge of how widespread they are and from which specific state they may have originated. Understanding the complexities of viral discovery at the state and local levels is crucial to understanding the specific problems in Alabama. Without this knowledge, we remain unaware of potential novel viruses that may pose economic threats to our crops. In Alabama, two particularly influential insect species, soybean thrips (*N. variabilis*), and tobacco thrips (*F. fusca*), are abundant in the summer months, infesting not just soybean and peanut plants but also clover and a variety of other weeds in the landscape. Alarmingly, research indicates that less than 1% of the viral diversity within these insect vectors has been examined through viral metagenomic analyses of environmental samples (Mokili et al., 2012). This considerable knowledge gap underscores the urgent need for a comprehensive study of the diversity of viruses—one that could provide insights to protect our crops and ensure economic stability in Alabama's agricultural sector. This study will explore RNA-seq as a tool for uncovering virus diversity in agriculturally important major insect vectors in Alabama and analyze and compare viral communities across different cropping system insects. Understanding this virus diversity is crucial, as it can help inform better pest control strategies and protect crop yields, ultimately supporting food security in the region.

Future Directions/Conclusion

Literature has found that the majority of viruses, around 60–99% of the contigs generated through different viral metagenomic studies, do not have significant homology to the existing known viruses (Mokili et al., 2012). This suggests that a multidisciplinary approach should be adopted in the future to enhance virus detection by accessing virome diversity. Virus discovery could be integrated with proteomics, metabolomics, and possibly the concept of epigenomics along with RNA-seq. Additionally, research should focus on developing a feasible RNA-seq equivalent that can assess real-time virome diversity in the field. These advancements will be critical in enabling early detection and risk assessment of emerging crop viruses, thereby informing timely and effective management strategies.

Reference

Baker, C., Blount, A., & Quesenberry, K. (1999). *Peanut Stunt Virus Infecting Perennial Peanuts in Florida and Georgia*. https://ccmedia.fdacs.gov/content/download/11401/file/pp395.pdf?utm_source=chatgpt.com

Barker, J., daSilva, L. L. P., & Crump, C. M. (2023). Mechanisms of bunyavirus morphogenesis and egress. *Journal of General Virology*, *104*(4). https://doi.org/10.1099/jgv.0.001845

Barkley, N. A., Pinnow, D. L., Wang, M. L., & Pederson, G. A. (2009). First Report of *Tomato spotted wilt virus* Infecting African Clover (*Trifolium tembense*) in Georgia. *Plant Disease*, *93*(2), 202–202. https://doi.org/10.1094/PDIS-93-2-0202C

Batson, J., Dudas, G., Haas-Stapleton, E., Kistler, A. L., Li, L. M., Logan, P., Ratnasiri, K., & Retallack, H. (2021). Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. *eLife, 10*, e68353. https://doi.org/10.7554/eLife.68353

Beach, A. M. (1897). Contributions to a knowledge of the Thripidae of Iowa. *The Proceedings of the Iowa Academy of Science.*, *3*, 214–227. https://doi.org/10.5962/bhl.part.14877

Brait, N., Hackl, T., Morel, C., Exbrayat, A., Gutierrez, S., & Lequime, S. (2024). A tale of caution: How endogenous viral elements affect virus discovery in transcriptomic data. *Virus Evolution*, *10*(1), vead088. https://doi.org/10.1093/ve/vead088

Brault, V., Mutterer, J., Scheidecker, D., Simonis, M. T., Herrbach, E., Richards, K., & Ziegler-Graff, V. (2000). Effects of Point Mutations in the Readthrough Domain of the Beet Western Yellows Virus Minor Capsid Protein on Virus Accumulation In Planta and on

Transmission by Aphids. *Journal of Virology*, *74*(3), 1140–1148. https://doi.org/10.1128/JVI.74.3.1140-1148.2000

Bruyère, A., Brault, V., Ziegler-Graff, V., Simonis, M.-T., Van Den Heuvel, J. F. J. M., Richards, K., Guilley, H., Jonard, G., & Herrbach, E. (1997). Effects of Mutations in the Beet Western Yellows Virus Readthrough Protein on Its Expression and Packaging and on Virus Accumulation, Symptoms, and Aphid Transmission. *Virology*, *230*(2), 323–334. https://doi.org/10.1006/viro.1997.8476

Burand, J. P., Kim, W., Afonso, C. L., Tulman, E. R., Kutish, G. F., Lu, Z., & Rock, D. L. (2012). Analysis of the genome of the sexually transmitted insect virus Helicoverpa zea nudivirus 2. *Viruses*, *4*(1), 28–61. https://doi.org/10.3390/v4010028

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009a). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009b). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009c). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Card, S. D., Pearson, M. N., & Clover, G. R. G. (2007). Plant pathogens transmitted by pollen. *Australasian Plant Pathology*, *36*(5), 455. https://doi.org/10.1071/AP07050

Chen, A. Y. S., Walker, G. P., Carter, D., & Ng, J. C. K. (2011). A virus capsid component mediates virion retention and transmission by its insect vector. *Proceedings of the National Academy of Sciences*, *108*(40), 16777–16782. https://doi.org/10.1073/pnas.1109384108

Chen, Y., Evans, J., & Feldlaufer, M. (2006). Horizontal and vertical transmission of viruses in the honey bee, Apis mellifera. *Journal of Invertebrate Pathology*, *92*(3), 152–159. https://doi.org/10.1016/j.jip.2006.03.010

Chen, Y.-M., Sadiq, S., Tian, J.-H., Chen, X., Lin, X.-D., Shen, J.-J., Chen, H., Hao, Z.-Y., Wille, M., Zhou, Z.-C., Wu, J., Li, F., Wang, H.-W., Yang, W.-D., Xu, Q.-Y., Wang, W., Gao, W.-H., Holmes, E. C., & Zhang, Y.-Z. (2022). RNA viromes from terrestrial sites across China expand environmental viral diversity. *Nature Microbiology*, *7*(8), 1312–1323. https://doi.org/10.1038/s41564-022-01180-2

Cobbin, J. C., Charon, J., Harvey, E., Holmes, E. C., & Mahar, J. E. (2021). Current challenges to virus discovery by meta-transcriptomics. *Current Opinion in Virology*, *51*, 48–55. https://doi.org/10.1016/j.coviro.2021.09.007

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015a). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015b). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Common types of viruses of floriculture crops and their modes of transmission. (2014, April 3). MSU Extension.

https://www.canr.msu.edu/news/common_types_of_viruses_of_floriculture_crops_and_their_modes_of_transmissi

Conner, K., Sikora, E. J., Zhang, L., & Burmester, C. (2013). First Report of *Soybean vein necrosis-associated virus* Affecting Soybeans in Alabama. *Plant Health Progress*, *14*(1), 49. https://doi.org/10.1094/PHP-2013-0729-03-BR

Culbreath, A. K., Todd, J. W., & Brown, S. L. (2003). E PIDEMIOLOGY AND M ANAGEMENT OF T OMATO S POTTED W ILT IN P EANUT. *Annual Review of Phytopathology*, *41*(1), 53–75. https://doi.org/10.1146/annurev.phyto.41.052002.095522

Dietzgen, R., Mann, K., & Johnson, K. (2016). Plant Virus–Insect Vector Interactions: Current and Potential Future Research Directions. *Viruses*, *8*(11), 303. https://doi.org/10.3390/v8110303

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Elena, S. F., Fraile, A., & García-Arenal, F. (2014). Evolution and Emergence of Plant Viruses. In *Advances in Virus Research* (Vol. 88, pp. 161–191). Elsevier. https://doi.org/10.1016/B978-0-12-800098-4.00003-9

Feng, Y., Krueger, E. N., Liu, S., Dorman, K., Bonning, B. C., & Miller, W. A. (2017). Discovery of Known and Novel Viral Genomes in Soybean Aphid by Deep Sequencing. *Phytobiomes Journal*, *1*(1), 36–45. https://doi.org/10.1094/PBIOMES-11-16-0013-R

Fletcher, J., Tang, J., Blouin, A., Ward, L., MacDiarmid, R., & Ziebell, H. (2016). *Red clover vein mosaic virus*—A Novel Virus to New Zealand that is Widespread in Legumes. *Plant Disease*, *100*(5), 890–895. https://doi.org/10.1094/PDIS-04-15-0465-RE

Gómez, M., Martínez, D., Páez-Triana, L., Luna, N., De Las Salas, J. L., Hernández, C., Flórez, A. Z., Muñoz, M., & Ramírez, J. D. (2023). Characterizing viral species in mosquitoes (Culicidae) in the Colombian Orinoco: Insights from a preliminary metagenomic study. *Scientific Reports*, *13*(1), 22081. https://doi.org/10.1038/s41598-023-49232-9

Gordon, A. (2010). *FASTQ/A short reads pre-processing tools.* [Computer software]. http://hannonlab.cshl.edu/fastx_toolkit/

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011a). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011b). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Hajimorad, M. R., Domier, L. L., Tolin, S. A., Whitham, S. A., & Saghai Maroof, M. A. (2018). *Soybean mosaic virus*: A successful potyvirus with a wide distribution but restricted natural

host range. *Molecular Plant Pathology*, *19*(7), 1563–1579. https://doi.org/10.1111/mpp.12644

Harrewijn, P., Tjallingii, W. F., & Mollema, C. (1996). Electrical recording of plant penetration by western flower thrips. *Entomologia Experimentalis et Applicata*, *79*(3), 345–353. https://doi.org/10.1111/j.1570-7458.1996.tb00842.x

Hasiów-Jaroszewska, B., Boezen, D., & Zwart, M. P. (2021). Metagenomic Studies of Viruses in Weeds and Wild Plants: A Powerful Approach to Characterise Variable Virus Communities. *Viruses*, *13*(10), 1939. https://doi.org/10.3390/v13101939

History and Economic Impact. (2025, March 14). https://tswv.caes.uga.edu/usda-ramp-project/history-and-economic-impact.html?utm_source=chatgpt.com

Hogenhout, S. A., Ammar, E.-D., Whitfield, A. E., & Redinbaugh, M. G. (2008). Insect Vector Interactions with Persistently Transmitted Viruses. *Annual Review of Phytopathology*, *46*(1), 327–359. https://doi.org/10.1146/annurev.phyto.022508.092135

Hunter, W. B., & Ullman, D. E. (1989). Analysis of mouthpart movements during feeding of Frankliniella occidentalis (pergande) and F. schultzei trybom (Thysanoptera: Thripidae). *International Journal of Insect Morphology and Embryology*, *18*(2–3), 161–171. https://doi.org/10.1016/0020-7322(89)90024-X

Hýsková, V., Bělonožníková, K., Chmelík, J., Hoffmeisterová, H., Čeřovská, N., Moravec, T., & Ryšlavá, H. (2024). Potyviral Helper-Component Protease: Multifaced Functions and Interactions with Host Proteins. *Plants*, *13*(9), 1236. https://doi.org/10.3390/plants13091236

Jia, H., & Gong, P. (2019). A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. *Frontiers in Microbiology*, *10*, 1945. https://doi.org/10.3389/fmicb.2019.01945

Jones, D. R. (2005). *Plant viruses transmitted by thrips*. *113*, 119–157. https://doi.org/10.1007/s10658-005-2334-1

Junglen, S., & Drosten, C. (2013). Virus discovery and recent insights into virus diversity in arthropods. *Current Opinion in Microbiology*, *16*(4), 507–513. https://doi.org/10.1016/j.mib.2013.06.005

Käfer, S., Paraskevopoulou, S., Zirkel, F., Wieseke, N., Donath, A., Petersen, M., Jones, T. C., Liu, S., Zhou, X., Middendorf, M., Junglen, S., Misof, B., & Drosten, C. (2019). Re-assessing the diversity of negative strand RNA viruses in insects. *PLOS Pathogens*, *15*(12), e1008224. https://doi.org/10.1371/journal.ppat.1008224

Keough, S., Danielson, J., Marshall, J. M., Lagos-Kutz, D., Voegtlin, D. J., Srinivasan, R., & Nachappa, P. (2018). Factors Affecting Population Dynamics of Thrips Vectors of Soybean vein necrosis virus. *Environmental Entomology*, *47*(3), 734–740. https://doi.org/10.1093/ee/nvy021

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009a). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009b). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Lanzi, G., de Miranda, J. R., Boniotti, M. B., Cameron, C. E., Lavazza, A., Capucci, L., Camazine, S. M., & Rossi, C. (2006). Molecular and biological characterization of deformed wing virus of honeybees (Apis mellifera L.). *Journal of Virology*, *80*(10), 4998–5009. https://doi.org/10.1128/JVI.80.10.4998-5009.2006

LaTora, A. G., Lai, P.-C., Chen, Y.-J., Gautam, S., Abney, M. R., & Srinivasan, R. (2022a). *Frankliniella fusca* (Thysanoptera: Thripidae), The Vector of *Tomato Spotted Wilt Orthotospovirus* Infecting Peanut in the Southeastern United States. *Journal of Integrated Pest Management*, *13*(1), 3. https://doi.org/10.1093/jipm/pmab045

LaTora, A. G., Lai, P.-C., Chen, Y.-J., Gautam, S., Abney, M. R., & Srinivasan, R. (2022b). *Frankliniella fusca* (Thysanoptera: Thripidae), The Vector of *Tomato Spotted Wilt Orthotospovirus* Infecting Peanut in the Southeastern United States. *Journal of Integrated Pest Management*, *13*(1), 3. https://doi.org/10.1093/jipm/pmab045

Lawrence, P., Heung, M., Nave, J., Henkel, C., & Escudero-Pérez, B. (2023). The natural virome and pandemic potential: Disease X. *Current Opinion in Virology*, *63*, 101377. https://doi.org/10.1016/j.coviro.2023.101377

Liu, S., Vijayendr, D., & Bonning, B. C. (2011). *Next Generation Sequencing Technologies for Insect Virus Discovery*. *3*, 1849–1869. https://doi.org/10.3390/v3101849

Looi, L. M., & Chua, K. B. (2007). Lessons from the Nipah virus outbreak in Malaysia. *Malaysian Journal of Pathology*, *29*(2), 63–67.

Ma, Y., Marais, A., Lefebvre, M., Faure, C., & Candresse, T. (2020). Metagenomic analysis of virome cross-talk between cultivated Solanum lycopersicum and wild Solanum nigrum. *Virology*, *540*, 38–44. https://doi.org/10.1016/j.virol.2019.11.009

Marklewitz, M., Zirkel, F., Kurth, A., Drosten, C., & Junglen, S. (2015). Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. *Proceedings of the National Academy of Sciences*, *112*(24), 7536–7541. https://doi.org/10.1073/pnas.1502036112

Marklewitz, M., Zirkel, F., Rwego, I. B., Heidemann, H., Trippner, P., Kurth, A., Kallies, R., Briese, T., Lipkin, W. I., Drosten, C., Gillespie, T. R., & Junglen, S. (2013). Discovery of a Unique Novel Clade of Mosquito-Associated Bunyaviruses. *Journal of Virology*, *87*(23), 12850–12865. https://doi.org/10.1128/JVI.01862-13

McLaughlin, M. R. (1992). Virus Diseases and Stand Decline in a White Clover Pasture. *Plant Disease*, *76*(2), 158. https://doi.org/10.1094/PD-76-0158

Mokili, J. L., Rohwer, F., & Dutilh, B. E. (2012). Metagenomics and future perspectives in virus discovery. *Current Opinion in Virology*, *2*(1), 63–77. https://doi.org/10.1016/j.coviro.2011.12.004

Nault, B. A., & Speese Iii, J. (2002). Major insect pests and economics of fresh-market tomato in eastern Virginia. *Crop Protection*, *21*(5), 359–366. https://doi.org/10.1016/S0261-2194(01)00103-X

Ng, T. F. F., Duffy, S., Polston, J. E., Bixby, E., Vallad, G. E., & Breitbart, M. (2011). Exploring the Diversity of Plant DNA Viruses and Their Satellites Using Vector-Enabled

Metagenomics on Whiteflies. *PLoS ONE*, *6*(4), e19050. https://doi.org/10.1371/journal.pone.0019050

Omkar. (2021). *Polyphagous Pests of Crops*. Springer Singapore Pte. Limited.

Paliwal, Y. C. (1974). Some properties and thrip transmission of tomato spotted wilt virus in Canada. *Canadian Journal of Botany*, *52*(6), 1177–1182. https://doi.org/10.1139/b74-151

Pan, S. (2024, September 25). Insects as Vectors – Features, Types, Examples. *Insects as Vectors – Features, Types, Examples*. https://biologynotesonline.com/insects-as-vectors-features-types-examples/

Peter, K. A., Liang, D., Palukaitis, P., & Gray, S. M. (2008). Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. *Journal of General Virology*, *89*(8), 2037–2045. https://doi.org/10.1099/vir.0.83625-0

Pirone, T. P., & Megahed, E.-S. (1966). Aphid transmissibility of some purified viruses and viral RNA's. *Virology*, *30*(4), 631–637. https://doi.org/10.1016/0042-6822(66)90168-1

Power, A. G., & Mitchell, C. E. (2004). Pathogen Spillover in Disease Epidemics. *The American Naturalist*, *164*(S5), S79–S89. https://doi.org/10.1086/424610

Qi, Y.-H., Ye, Z.-X., Zhang, C.-X., Chen, J.-P., & Li, J.-M. (2023). Diversity of RNA viruses in agricultural insects. *Computational and Structural Biotechnology Journal*, *21*, 4312–4321. https://doi.org/10.1016/j.csbj.2023.08.036

Riley, D. G., Joseph, S. V., Srinivasan, R., & Diffie, S. (2011). Thrips Vectors of Tospoviruses. *Journal of Integrated Pest Management*, *2*(1), I1–I10. https://doi.org/10.1603/IPM10020

Roossinck, M. J., & García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Current Opinion in Virology*, *10*, 56–62. https://doi.org/10.1016/j.coviro.2015.01.005

Rosario, K., Capobianco, H., Ng, T. F. F., Breitbart, M., & Polston, J. E. (2014). RNA Viral Metagenome of Whiteflies Leads to the Discovery and Characterization of a Whitefly-Transmitted Carlavirus in North America. *PLoS ONE*, *9*(1), e86748. https://doi.org/10.1371/journal.pone.0086748

Rosario, K., Seah, Y., Marr, C., Varsani, A., Kraberger, S., Stainton, D., Moriones, E., Polston, J., Duffy, S., & Breitbart, M. (2015). Vector-Enabled Metagenomic (VEM) Surveys Using Whiteflies (Aleyrodidae) Reveal Novel Begomovirus Species in the New and OldWorlds. *Viruses*, *7*(10), 5553–5570. https://doi.org/10.3390/v7102895

Rozo-Lopez, P., Brewer, W., Käfer, S., Martin, M. M., & Parker, B. J. (2023). Untangling an insect's virome from its endogenous viral elements. *BMC Genomics*, *24*(1), 636. https://doi.org/10.1186/s12864-023-09737-z

Soybean Thrips. (n.d.). Purdue University - College of Agriculture. Retrieved March 14, 2025, from https://ag.purdue.edu/department/entm/extension/field-crops-ipm/soybean/soybean-thrips.html

Soybean Vein Necrosis Virus—Soybean Disease—Soybean Research & Information Network—SRIN. (n.d.). Soybean Research & Information Network. Retrieved March 29, 2025, from https://soybeanresearchinfo.com/soybean-disease/soybean-vein-necrosis-virus/

Srinivasan, R., Abney, M. R., Culbreath, A. K., Kemerait, R. C., Tubbs, R. S., Monfort, W. S., & Pappu, H. R. (2017). Three decades of managing Tomato spotted wilt virus in peanut in

southeastern United States. *Virus Research*, *241*, 203–212. https://doi.org/10.1016/j.virusres.2017.05.016

Stafford, C. A., Walker, G. P., & Ullman, D. E. (2012). Hitching a ride: Vector feeding and virus transmission. *Communicative & Integrative Biology*, *5*(1), 43–49. https://doi.org/10.4161/cib.18640

Stewart, L. R., Medina, V., Tian, T., Turina, M., Falk, B. W., & Ng, J. C. K. (2010). A Mutation in the *Lettuce Infectious Yellows Virus* Minor Coat Protein Disrupts Whitefly Transmission but Not *In Planta* Systemic Movement. *Journal of Virology*, *84*(23), 12165–12173. https://doi.org/10.1128/JVI.01192-10

The PLOS Pathogens Staff. (2014). Correction: Transovarial Transmission of a Plant Virus Is Mediated by Vitellogenin of Its Insect Vector. *PLoS Pathogens*, *10*(4), e1004141. https://doi.org/10.1371/journal.ppat.1004141

Thekke-Veetil, T., Lagos-Kutz, D., Domier, L. L., McCoppin, N. K., Hartman, G. L., & Clough, S. J. (2024). Exploring Virus Diversity in the Potato leafhopper (Empoasca fabae), an Economically Important Agricultural Pest. *Viruses*, *16*(8), 1305. https://doi.org/10.3390/v16081305

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020a). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020b). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly

Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020c). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thrips Management Guidelines—UC IPM. (n.d.). Retrieved April 6, 2025, from https://ipm.ucanr.edu/PMG/PESTNOTES/pn7429.html?utm_source=chatgpt.com

Tobacco thrips. (2025, February 27). *Tobacco Thrips*. https://www.biobee.com/pests/tobacco-thrips/

Umar, M., Tegg, R. S., Farooq, T., Thangavel, T., & Wilson, C. R. (2022). Abundance of Poleroviruses within Tasmanian Pea Crops and Surrounding Weeds, and the Genetic Diversity of TuYV Isolates Found. *Viruses, 14*(8), 1690. https://doi.org/10.3390/v14081690

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., Guzman, H., Wood, T. G., Popov, V., Gorchakov, R., González, A. V., Haddow, A. D., Watts, D. M., Da Rosa, A. P. A. T., Weaver, S. C., Lipkin, W. I., & Tesh, R. B. (2013). Negevirus: A Proposed New Taxon of Insect-Specific Viruses with Wide Geographic Distribution. *Journal of Virology*, *87*(5), 2475–2488. https://doi.org/10.1128/JVI.00776-12

Vasilakis, N., Guzman, H., Firth, C., Forrester, N. L., Widen, S. G., Wood, T. G., Rossi, S. L., Ghedin, E., Popov, V., Blasdell, K. R., Walker, P. J., & Tesh, R. B. (2014). Mesoniviruses

are mosquito-specific viruses with extensive geographic distribution and host range. *Virology Journal*, *11*(1), 97. https://doi.org/10.1186/1743-422X-11-97

Venkataraman, S., Prasad, B., & Selvarajan, R. (2018). RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. *Viruses*, *10*(2), 76. https://doi.org/10.3390/v10020076

Wamonje, F. O., Michuki, G. N., Braidwood, L. A., Njuguna, J. N., Musembi Mutuku, J., Djikeng, A., Harvey, J. J. W., & Carr, J. P. (2017). Viral metagenomics of aphids present in bean and maize plots on mixed-use farms in Kenya reveals the presence of three dicistroviruses including a novel Big Sioux River virus-like dicistrovirus. *Virology Journal*, *14*(1), 188. https://doi.org/10.1186/s12985-017-0854-x

Wang, Z., Gerstein, M., & Snyder, M. (2009). *RNA-Seq: A revolutionary tool for transcriptomics*. *10*, 57–63.

Wendimu, G. Y., & Gurmu, A. K. (2024). Insect Vectors of Plant Viruses: Host Interactions, Their Effects, and Future Opportunities. *2024*, *2024*(1). https://doi.org/10.1155/aia/6006985

Whitfield, A. E., Falk, B. W., & Rotenberg, D. (2015a). Insect vector-mediated transmission of plant viruses. *Virology*, *479–480*, 278–289. https://doi.org/10.1016/j.virol.2015.03.026

Whitfield, A. E., Falk, B. W., & Rotenberg, D. (2015b). *Insect vector-mediated transmission of plant viruses*. *479*, 278–289.

Wu, H., Pang, R., Cheng, T., Xue, L., Zeng, H., Lei, T., Chen, M., Wu, S., Ding, Y., Zhang, J., Shi, M., & Wu, Q. (2020). Abundant and Diverse RNA Viruses in Insects Revealed by RNA-Seq Analysis: Ecological and Evolutionary Implications. *mSystems*, *5*(4), e00039-20. https://doi.org/10.1128/mSystems.00039-20

Xie, Y.-H., Li, Z.-Y., Dong, K., & Zhang, H.-R. (2012). Changes in the species composition of thrips on *Trifolium repens* (Fabales). *Acta Phytopathologica et Entomologica Hungarica*, *47*(1), 61–67. https://doi.org/10.1556/APhyt.47.2012.1.7

Zhou, J., Kantartzi, S. K., Wen, R.-H., Newman, M., Hajimorad, M. R., Rupe, J. C., & Tzanetakis, I. E. (2011). Molecular characterization of a new tospovirus infecting soybean. *Virus Genes*, *43*(2), 289–295. https://doi.org/10.1007/s11262-011-0621-9

Zhou, J., & Tzanetakis, I. E. (2013). Epidemiology of Soybean vein necrosis-associated virus. *Phytopathology®*, *103*(9), 966–971. https://doi.org/10.1094/PHYTO-12-12-0322-R

Figures

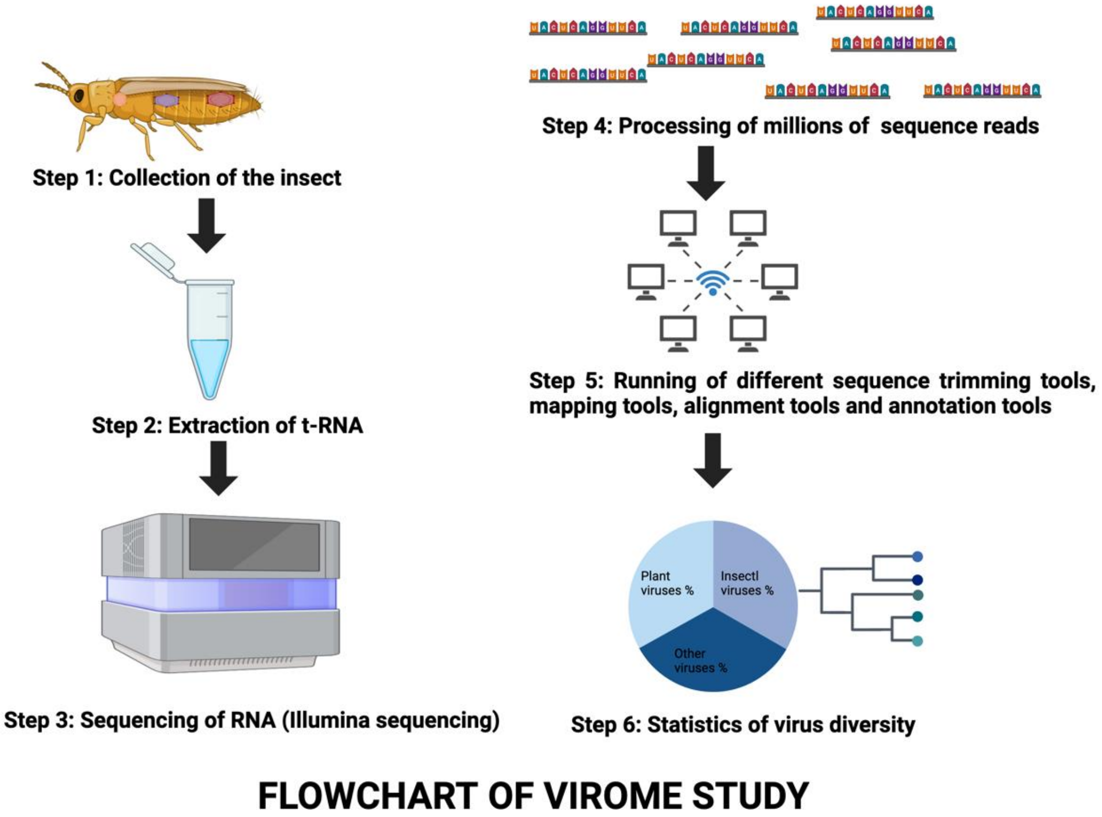


Figure 1. Flowchart of the virome study through RNA-seq analysis

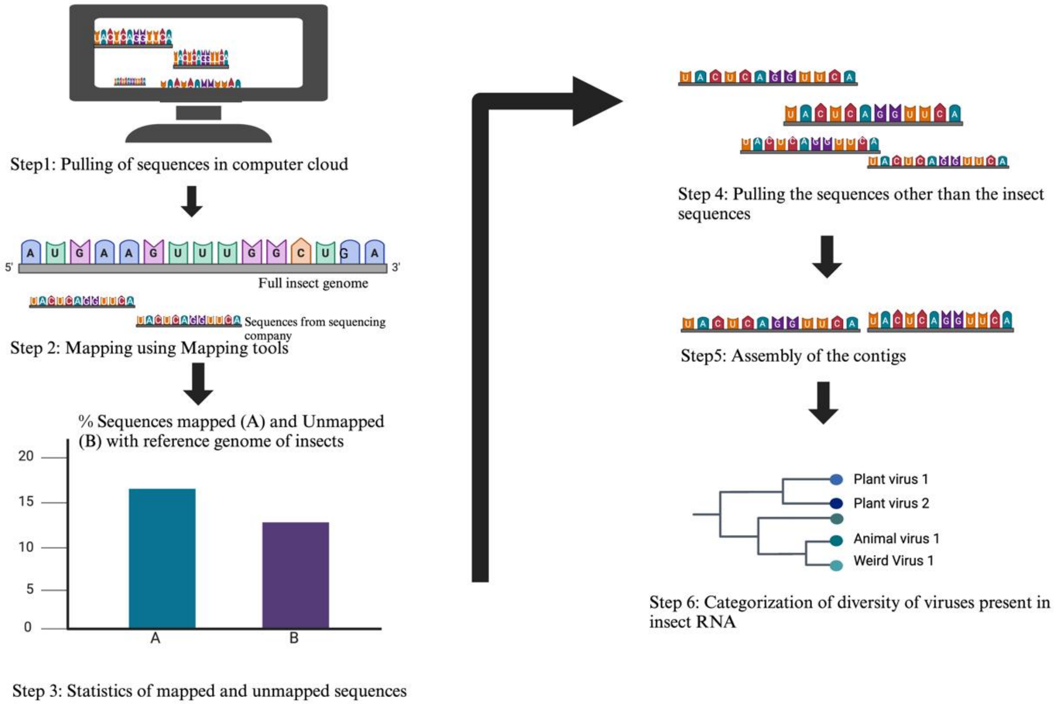
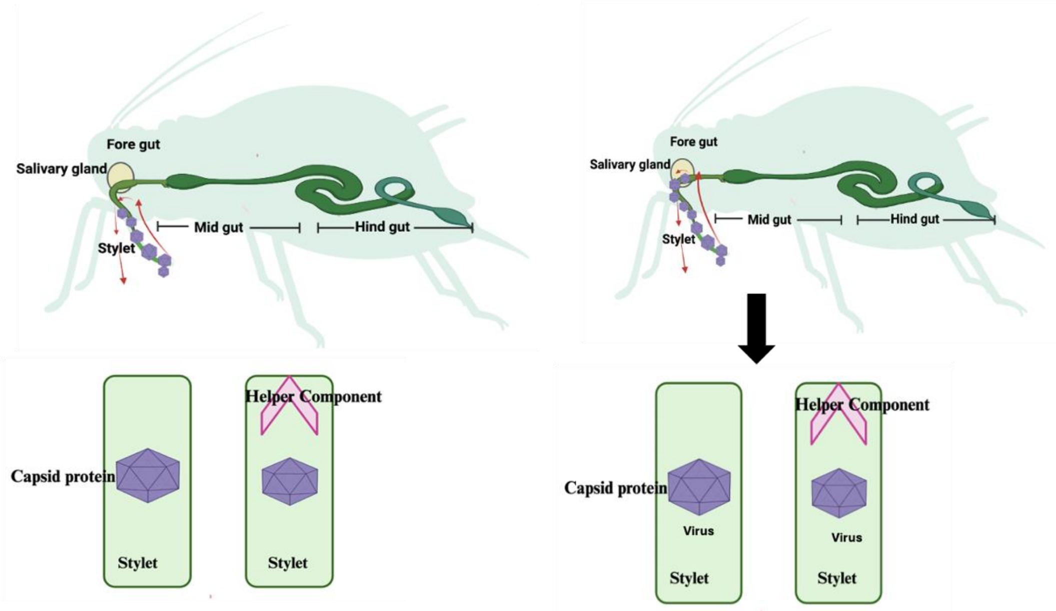
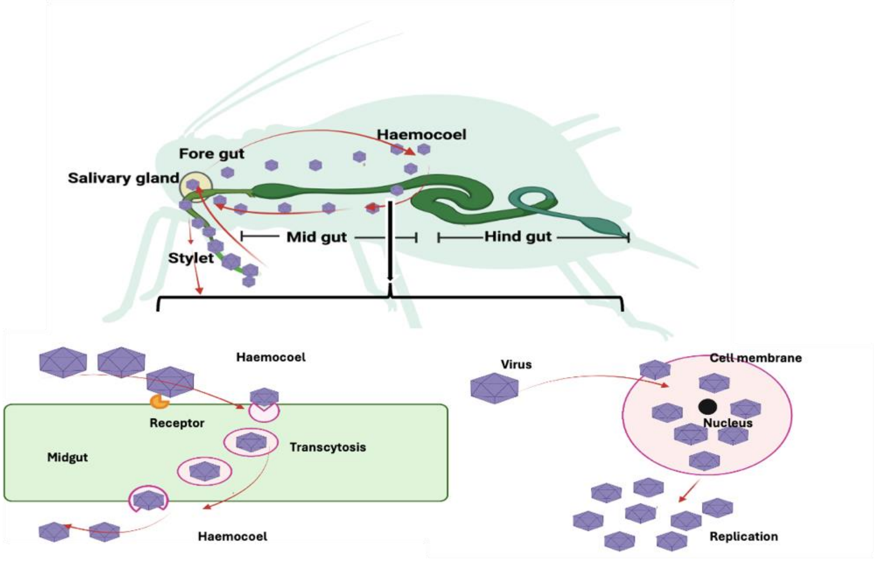


Figure 2. Flowchart of Bioinformatic analysis of RNA-seq data received from Sequencing company. Each step represents the flow of the sequence analysis



A. Non-Persistent, Non-circulative transmission

B. Semi-Persistent, Non-circulative transmission



i. Non-Propagative

ii. Propagative

C. Persistent transmission: i. Non-propagative ii. Propagative

Figure 3. Modes of virus transmission in insect vector A) Non-Persistent, Non-circulative transmission B) Semi-Persistent, Non-circulative transmission C) Persistent transmission: i. non-propagative ii. Propagative

Chapter 2

Metagenomic Analysis of Soybean Thrips Reveals Viruses of Insects, Plants, and Fungi in Alabama

Sharma A.*, Zhao C.**, Martin K. M.*

*Department of Entomology and Plant Pathology, Auburn University, Auburn, Alabama

** USDA- ARS

Abstract

Soybean thrips (*Neohydatothrips variabilis* Beach; Thysanoptera, Thripidae) is a cosmopolitan pest and an important vector for many viral diseases including *Orthotospovirus glycinecrovenae*, Soybean vein necrosis virus (SVNV). A previous meta-transcriptome study analyzed the total RNA of *N. variabilis* from a sample pooled from eight states in the Midwest USA. To discover the viruses present in insects in Alabama, we utilized RNA-seq to sequence total RNA of *N. variabilis* collected in Auburn, Alabama, in 2023. Fifteen viruses were discovered through metagenomic analysis. We characterized fifteen viruses belonging to Bunyavirales (4), Durnavirales (1), Ghabrivirales (1), Mononegavirales (1), Patatavirales (1), Wolframvirales (3) and unclassified orders (4). The best matches to viruses that we found were discovered previously from USA, Japan, Korea, and Spain. Out of the fifteen viruses ten are potential insect infecting viruses, one virus is both plant and insect infecting virus, one is strictly plant-infecting virus, and three are potential mycoviruses. The three mycoviruses are novel with less than 80% Amino acid identity. This study helped us to characterize the viruses at the state level, which will help us to understand the diversity, evolution, and eventually, potential crop threats in the future.

Key words: RNA-seq, Soybean thrips, virome, novel viruses, transcriptomics, viruses

Introduction

Soybean thrips (*Neohydratothrips variabilis* Beach, Thysanoptera, Thripidae) are significant phytophagous thrips species with a host range that includes soybean, cotton, lima bean, tomato, cucumber, and smartweed, as well as the kudzu plant. They are important agricultural pests in Alabama, where soybean (*Glycine max* L. Merr) is cultivated on approximately 360,000 acres, contributing an estimated economic value of $19.6 million. After Brazil, the United States ranks as the second-largest exporter of soybeans globally. In both the United States and Canada, the most economically damaging virus transmitted by *N. variabilis* in a persistent, propagative manner is the Soybean vein necrosis orthotospovirus (*Orthotospovirus glycininecrovenae*, SVNV) which produces symptoms like leaf chlorosis, leaf necrosis, and vein clearing. However, sequencing studies suggest that the viral associations of this insect extend beyond known plant pathogens, indicating a broader vector potential (Thekke-Veetil et al., 2020b). Sequencing total RNA from insects to identify viruses has been instrumental in the identification of both known and novel viruses that are present in insects. It has been successfully used to study the viromes in different insect species such as aphids, thrips (Feng et al., 2017; Wamonje et al., 2017) and whiteflies (Ng et al., 2011; Rosario et al., 2014, 2015) in addition to soybean thrips (Thekke-Veetil et al., 2020b). In soybean thrips, viruses were characterized through high-throughput sequencing (HTS) from a pooled sample collected from eight states in the Midwestern USA identified 181 virus-like sequences were characterized with 155 being novel viruses (Thekke-Veetil et al., 2020b). Although impressive in its identification of novel viruses, it does not answer the question of what diversity might look like in soybean thrips at a single-state level.

Broader studies have not resolved the specific viral diversity present in soybean thrips (*N. variabilis*) populations at the local scale in Alabama, limiting our ability to detect emerging viruses or understand regional variation. To address this gap, we performed metagenomic RNA sequencing on *N. variabilis* collected from a single site in Auburn, Alabama. By narrowing the geographic scope, this study enables a detailed characterization of the viral community carried by local thrips populations, including both known and novel viruses. These findings provide a foundation for understanding virus-vector associations, potential transmission dynamics, and future risks to regional crop systems.

Method

Collection and identification of Soybean thrips (*Neohydatothrips variablilis* Beach)

Soybean thrips (*N. variablilis* Beach) was collected from soybeans from the Old Rotation and Cullar's rotation in Auburn during September and October 2023. To minimize bias and sampling error, random rows of soybeans were selected upon arrival at each site. Starting at the beginning of two parallel rows of soybeans, a sweep collection with 100 sweeps was performed while simultaneously walking 100 steps. At the end of the 100 sweeps, the net contents were emptied into large plastic Ziploc bags and placed on ice until identification. Upon arrival at the lab, each collection bag was taken from the container of ice for identification. The contents of the bags were emptied into large, covered petri dishes where leaves and tissue were examined for *N. variabilis.* They were stored in ethanol at -20^{o}C until extraction.

RNA Extraction – TriZol Method

Soybean thrips (83) were moved from their original container into a sterile 1.7mL tube containing 250 uL of TriZol and ground thoroughly. An additional 750 uL of TriZol was added while grinding and then incubated for 3 minutes using a rotating mixer. Samples were spun down at 12,000 x g at 4^{o}C for 15 minutes. TriZol was collected while carefully avoiding the insect debris and placed in a new sterile 1.7mL tube. Then, 200 uL of chloroform was added to the sample and shaken vigorously for 15 seconds before incubating at room temperature for 5 minutes. Samples were then spun down at 12,000 x g at 4°C for 10 minutes. The aqueous phase (top layer) was carefully collected and placed in a new sterile 1.7mL tube with 500 uL of isopropanol. Samples were then shaken gently by hand and incubated at room temperature for 10 minutes. After incubation, the samples were spun down at 12,000 x g at 4°C for 10 minutes. The supernatant was removed, taking caution not to disturb the small white RNA pellet at the bottom. The pellet was washed with 500 uL of 80% ethanol and spun down at 7,500 x g at 4°C for 5 minutes. All ethanol was removed, and after two minutes at room temperature, the pellet was redissolved in 10 uL of molecular-grade water. Sample concentrations were checked on the Nanodrop and stored in a -20°C freezer.

RNA Ribodepletion

The Illumina Ribo-Zero Plus rRNA Depletion Kit was utilized for the removal of ribosomal RNA (rRNA) from total RNA samples. Then, the total RNA sample was incubated with Ribo-Zero Plus solution targeting both cytoplasmic and mitochondrial rRNA. The mixture was then subjected to hybridization, followed by binding to magnetic beads. Subsequently, the rRNA-depleted supernatant containing non-rRNA transcripts was carefully collected and quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific). The rRNA-depleted RNA samples were stored at -80°C until the quality was checked. RNA samples were thawed on ice and gently mixed by pipetting. ScreenTape assays were prepared according to the manufacturer's instructions. Then RNA samples were diluted to a final concentration of 50 ng/µL using molecular grade water. Subsequently, 1 µL of each diluted RNA sample was mixed with 5 µL of RNA ScreenTape sample buffer. The mixture was briefly vortexed and centrifuged to ensure proper mixing. The Agilent

TapeStation instrument (Agilent Technologies) was used to determine RNA quality following the manufacturer's instructions. All subsequent steps were conducted following the manufacturer’s instructions, and the software automatically calculated RNA integrity numbers (RINs) based on the electropherogram analysis.

Sequencing

Total RNA samples were sent to the North Carolina State Genomic Sciences Laboratory (Raleigh, NC, USA) for Illumina RNA library preparation and sequencing. Before constructing the library, RNA integrity, purity, and concentration were evaluated again using an Agilent 4200 Tapestation with RNA Tape (Agilent Technologies, USA). Messenger RNA (mRNA) was isolated using oligo-dT beads from the NEBNExt Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, USA). cDNA libraries for Illumina sequencing were generated using the NEBNext Ultra II Directional RNA Library Prep Kit (NEB) and NEB Next Multiplex Oligos for Illumina (NEB) following the manufacturer’s protocol. Then mRNA was chemically fragmented and primed with random oligos for first-strand cDNA synthesis, followed by second-strand synthesis using dUTPs to retain strand orientation. The resulting double-stranded cDNA was purified, end-repaired, and subjected to “A-tailing” for adaptor ligation. After ligation, sequential AMPure XP bead isolation (Beckman Coulter, USA) was used to select library fragments of 400-550 bp (including adapters). Library enrichment and indexing for each sample were conducted through PCR amplification, after which the amplified fragments were purified and assessed for quality and concentration using an Agilent 4200 Tapestation (DNA Tape, Agilent Technologies, USA). The quantified libraries were then pooled in equimolar amounts for clustering and sequencing on an Illumina NovaSeq 6000 DNA sequence, utilizing an S4 150x2 PE sequencing reagent kit (Illumina, USA). The Real Time Analysis (RTA) software package was employed to generate raw base call (bcl) files, which were subsequently de-multiplexed into fastq files for data submission.

Bioinformatic Analysis

Once received, the quality of the sequences was checked using FastQC (Andrews, S., n.d.). Phred33 Score was used to measure the base quality of sequences. Any sequences or bp that had less than 28 Phred score were eliminated to ensure >99.9% of the base call accuracy. The poor-quality sequences were trimmed to 1-150 bp to get rid of the poor-quality base pairs using Trim sequences tools (Gordon, 2010). For further virome analysis, sequences with a quality score of 36 were processed through the bioinformatic pipeline (Figure 1). The reference genome of *N. variabilis* is not available, following previous study the reads were mapped to the full reference genome of Tobacco thrips, *Frankliniella fusca* Hinds (GCA-029379305.1) (Thekke-Veetil et al., 2020b) using Bowtie2 tools (Langmead et al., 2009a). The unaligned reads were assembled using Trinity (Grabherr et al., 2011a). These contigs were blasted using NCBI BLAST+ blastn against NCBI-NT 2023 universal database (Camacho et al., 2009a; Cock et al., 2015a) (Figure1). The sorting of the contigs was done using Microsoft Excel. The virus annotation was sorted based on the key words “virus”, “viridae”, “phage” and the longest contig with high amino acid identity to the best match was considered. Then, the NCBI ORF Finder was used to predict the start of viral genes that have not been characterized previously. For the viruses whose polymerases (RNA-dependent, RNA polymerase, RdRp) was discovered, the conserved domain was predicted using NCBI Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml).

Result

Bioinformatics statistics

The sequence statistics suggests the total sequences are paired end 1235537868, with total bases 186.5 Gbp, sequence length 151 bp and GC 58%. Out of 1235537868 paired end sequences the overall alignment rate to the reference genome is 29.21%. The total contigs generated are 439109. The length 0-1000 bp are 357971, and >1000 bp are 81138. The total trinity 'genes' are 311,643, the total trinity transcripts are 439,109, and GC percentage is 47.50%. Based on all transcripts, the Contig N10 is 12,588, Contig N20 is 8,890, Contig N30: 6,528, Contig N40: 4,607, and Contig N50: 2,998. The Median contig length is 371 bp, the average contig length is 1003.32 bp, and the total assembled bases were 440,568,337. Based on only the longest isoform per 'gene', the Contig N10 is 7,833, Contig N20 is 3,952, Contig N30: 1,963, Contig N40: 1,116, and Contig N50: 740. The median contig length is 331 bp, the average contig length is 590.82 bp, and the total assembled bases are 184,123,743. The total viral contigs were 3971 (Table 1).

Fifteen potential viruses were discovered

There are 18 unique contigs annotated to 15 different viruses: Bunyavirales (4), Durnavirales (1), Mononegavirales (1), Wolframvirales (3), Ghabrivirales (1), Patatavirales (1), and Other unclassified orders (4) (Figure 2). 46% of viruses were RNA viruses, 21% were DNA viruses, and the rest 46% had the unclassified genome (Figure 2, Figure 3). In this study, it is expected that the closest matches to the viruses discovered from *N. variabilis*, will be similar to the larger Midwest study conducted in 2018 (Thekke-Veetil et al., 2020).

Potential insect-infecting viruses

Virus belonging to Bunyavirales. Three viruses belonging to the order Bunyavirales were identified in this study, each isolated from the host soybean thrips (*N. variabilis*). The first, Soybean thrips bunya-like virus 9, was characterized from sequence PQ492259 (1139 bp), which aligns from 282 bp to 1420 bp and shares 99.649% amino acid identity with its best match—the nucleocapsid protein gene of soybean thrips bunya-like virus 9 strain STN1BV9B (MW033656.1, 2411 bp), originally isolated from Soybean thrips in the Midwest USA in 2018 (Thekke-Veetil et al., 2020). This represents the first report of this virus in *N. variabilis* in Alabama, and no other sequences are currently reported in NCBI. Additionally, two fragments of Soybean thrips-associated tospovirus 1 were identified: PQ492260 (1063 bp), aligning from 2421 bp to 3482 bp with 95.58% amino acid identity to strain STN1 (MT195545.1, 4065 bp), and PQ492261 (485 bp), aligning from 4638 bp to 5122 bp with 94.02% amino acid identity to the RNA-dependent RNA polymerase gene of strain STN1 (MT195544.1, 8917 bp). Both fragments match sequences previously isolated from Soybean thrips in the Midwest USA in 2018 (Thekke-Veetil et al., 2020), and no other relatives are reported at NCBI. Lastly, the RNA-dependent RNA polymerase (RdRp) of Soybean thrips chu-like virus 6 was represented by PQ492267 (452 bp), aligning from 3 bp to 454 bp with 85.398% amino acid identity to soybean thrips chu-like virus 5 strain STN1CV5 (MT293124.1), also originally found from soybean thrips (*N. variabilis*) in the Midwest in 2018. As with the others, no additional relatives for this virus are reported in NCBI.

Virus belonging to Durnavirales

The RNA-dependent RNA polymerase (RdRp) of soybean thrips partiti-like virus 5 RNA-dependent RNA polymerase gene, partial cds (PQ492258,1355 bp) has an alignment range 1 bp to 1355 bp and 99.63% amino acid identity to the best match soybean thrips partiti-like virus 5 isolate STN1PV5 RNA-dependent RNA polymerase gene (MT648425.1, 1375bp) previously found from soybean thrips (*N. variabilis*) from Midwest, USA in 2018 (Thekke-Veetil et al., 2020b). There are no other relatives reported at NCBI.

Virus belonging to Ghabrivirales

The putative capsid protein of Soybean thrips-associated totivirus 1 nonfunctional capsid protein gene, partial sequence (PQ492268, 402 bp) that has an alignment range 485 bp to 886 bp and 99.0 % Amino acid identity to the best match soybean thrips-associated totivirus 1 strain STN1 (MT293124.1, 5662 bp) that was previously found from soybean thrips (*N. variabilis*) from Midwest USA in 2018 (Thekke-Veetil et al., 2020b). There are no other relatives reported at NCBI.

Virus belonging to Mononegavirales

The RNA-dependent RNA polymerase (RdRp) of Soybean thrips rhabdo-like virus 3 nonfunctional RNA-dependent RNA polymerase gene, partial sequence (PQ492262, 2981 bp), 969 bp) that has an alignment range 10932 bp to 11840 bp and 98.65% Amino acid identity to Soybean thrips rhabdo-like virus 3 strain STN1RV3 (MW023861.1, 12697 bp). The best match was previously found from soybean thrips (*N. variabilis*) from Midwest USA in 2018 (Thekke-Veetil et al., 2020b). There are no other relatives reported at NCBI.

Virus belonging to unclassified orders

The predicted ORF2 (PQ492257, 2981 bp) of the Soybean thrips virus 2 NS3-like protein gene, complete cds has an alignment range 143 bp to 2653 bp and 99.49% Amino acid identity to Soybean thrips virus 2 strain STN1ThV2 NS3-like protein gene (MW023856.1, 3071 bp) that was previously found from soybean thrips (*N. variabilis*) from Midwest USA in 2018 (Thekke-Veetil et al., 2020b). There are no other relatives reported at NCBI.

The predicted ORF1 (PQ492263, 654 bp) has the alignment range 1273 bp to 1926 bp and 94.8% amino acid identity to the best match soybean thrips virus 4 strain STN1ThV4 putative capsid protein and hypothetical protein genes (MW033627.1, 3002 bp). The predicted ORF9 (PQ492264, 545 bp) has an alignment range 1 bp to 545 bp and 95.41% Amino acid identity to the best match Soybean thrips virus 4 strain STN1ThV4 NS5-like protein gene, complete cds (MW033625.1, 3279 bp). Both best matches were previously found from soybean thrips (*N. variabilis*) from Mid-west USA in 2018 (Thekke-Veetil et al., 2020b). There are no other relatives reported at NCBI.

The two genes predicted ORF2 (PQ492265, 639 bp) and ORF3(PQ492266, 354 bp) of Aphis glycines virus 2 was discovered. The ORF2 has an alignment range from 4084bp to 4721bp and amino acid identity 98.74% and ORF3 has an alignment range of 1219 bp to 1572 bp and 98.84% Amino acid identity respectively to Aphis glycines virus 2 strain STN1AV2 (MW033624.1). The best match was previously been found from the soybean thrips (*N. variabilis*) from Mid-west USA in 2018 (Thekke-Veetil et al., 2020b).

Potential plant and insect infecting viruses

Virus belonging to order Bunyavirales

The study discovered different segments of Soybean vein necrosis virus (SVNV) isolate Auburn. One segment matches the nucleoprotein gene of the SVNV nonfunctional nucleocapsid gene, partial sequence (PQ492254, 349 bp), has an alignment range of 34 bp to 382 bp and 93.98% amino acid identity to the Soybean vein necrosis virus isolate Maryland 2 nucleoprotein gene (JQ946870.1, 834bp) that was isolated in the USA in 2011 from soybean.

Soybean vein necrosis virus M segment sequence (PQ492255, 520 bp) has an alignment range 1063bp to 1578bp and 97.11% amino acid identity to the best match Soybean vein necrosis virus isolate TN segment M (NC_055180.1, 4955 bp). The 213 bp annotates the portion of gene junction (1063 bp-1276 bp), and 302 bp matches the glycoprotein precursor (1276 bp-1578 bp).

The segment L of the Soybean vein necrosis virus sequence (PQ492256, 588 bp) corresponding to the RNA-dependent RNA polymerase (RdRp) has 773 bp to 1360 bp and 96.77% amino acid identity to the best match Soybean vein necrosis virus isolate TN segment L (NC_055179.1, 9010 bp). This sequence was characterized in 2011 in the Mid-west and Mid-south parts of the United States (Zhou et al., 2011) with an alignment range from 773 bp to 1360 bp.

Potential plant infecting viruses

The study discovered the polyprotein gene of Soybean mosaic virus (PQ492253, 2256 bp) with an alignment range from 3721 bp to 5976 bp with 99.29% amino acid identity to Soybean mosaic virus isolate Andong (OQ161635.1, 9594 bp) that was isolated in South Korea in 2020 from *Glycine max*.

Potential mycoviruses

The RdRp of three novel mycoviruses (dsDNA) belonging to the order Wolframvirales with the best match <80% were discovered. The RdRp of Aspergillus lentulus Narnavirus 1 sequence (PQ492251, 955 bp) has an alignment range of 1091 bp to 2044 bp and 80.39% amino acid identity to Aspergillus lentulus Narnavirus 1 J-YC genomic RNA, segment 1 (LC553702.1, 2074 bp) that was found in Japan (Date NA) from the host Aspergillus lentulus IFM63547 The RdRp of Magnaporthe oryzae narnavirus 3 nonfunctional RNA-dependent RNA polymerase gene, partial sequence (PQ492250, 798 bp) has an alignment range 1088bp to 1998bp and 78% Amino acid identity to Magnaporthe oryzae narnavirus 2 J-YC genomic RNA (LC553714.1, 3074 bp) that was isolated in Japan (Date NA) in host Magnaporthe oryzae APU10-199A.

The RdRp of Erysiphe necator associated narnavirus 49 nonfunctional RNA-dependent RNA polymerase gene, partial sequence (PQ492252, 769 bp) has an alignment range 988 bp to 1704 bp and 79.84% Amino acid identity to Erysiphe necator associated narnavirus 49 isolate PMS8_DN15148 (2300 bp, MN605462.1) that was first time isolated in Spain in 2018 from host *Erysiphe necator*.

Three out of five viral RdRp have a conserved catalytic domain

We initially hypothesized that all viruses for which only the RNA-dependent RNA polymerase (RdRp) was detected would share a conserved catalytic domain—a hallmark of their replication machinery. Surprisingly, this was not the case. Only three out of the five viruses exhibited the conserved catalytic domain, highlighting unexpected diversity among these RdRp-only viruses. The three viruses that did retain this conserved domain were Magnaporthe oryzae narnavirus 3 (PQ492250, 798 bp), Aspergillus lentulus narnavirus 1 (PQ492251, 955 bp), and soybean thrips partiti-like virus 5 (PQ492258, 1355 bp) (Figure 4).

Discussion

The study expected to discover different potential insect-infecting viruses through metagenomic analysis of the total RNA of the soybean thrips. The best match of all insect viruses (PQ492261, PQ492260, PQ492261, PQ492267, PQ492258, PQ492268, PQ492262, PQ492257, PQ492264, PQ492263, PQ492265, PQ492266) has been discovered before in USA in 2018 from a Midwest survey (RNA-seq analysis of soybean thrips from 8 different states) (Thekke-Veetil et al., 2020b). This suggests that these viruses could be common viruses in soybean thrips sampled from across the United States without respect to region. These might be common viruses in soybean thrips from other areas as well in USA; for that, we needed to conduct more metagenomic studies on the virome of soybean thrips in different regions in USA. Potential modes of transmission of these viruses in the thrips population might include oral uptake, exposure to contaminated substrates, or vertical transmission from parent to offspring.

The study discovered the best match of viruses from South Korea (OQ161635.1), Japan (LC553702.1, LC553714.1), and Spain (MN605462.1). Therefore, this suggests that the virus diversity is not limited to any specific country but is also found across countries. We expected majority of the viruses that were discovered would be RNA viruses. We observed 46% of the viruses as RNA viruses. Over the past decade, advances in sequencing technologies—particularly next-generation sequencing and metagenomics—combined with phylogenetic analyses, have proven highly effective in uncovering a diverse array of novel RNA viruses in insects (Junglen & Drosten, 2013; Marklewitz et al., 2013, 2015; Vasilakis et al., 2013, 2014). Additionally, this study identified that 23% of the detected viruses were DNA viruses, which may represent transcripts from open reading frames of DNA virus genomes. Notably, 31% of the viral sequences remain taxonomically unclassified, highlighting a significant knowledge gap that warrants further investigation. Increased attention should also be given to RNA viruses, as their higher mutation rates—compared to DNA viruses—can lead to rapid evolution and potentially greater ecological and pathogenic impacts (Sanjuán & Domingo-Calap, 2016).

Although the study initially aimed to identify viruses from a wide range of orders, the majority of detected viruses were found to belong to the order Bunyavirales. This group comprises arthropod-borne, spherical, enveloped RNA viruses capable of infecting both plants and animals, including insects. Notably, Bunyavirales continues to gain attention due to its expanding potential to impact global agriculture and public health, posing a growing threat to the world economy (Ye & Yan, 2024). This dominance may be further amplified by their capacity for rapid mutation and evolution. Given this, continued research is essential to better understand the viruses identified in this study. In fact, there is evidence that insect-infecting viruses within Bunyavirales can evolve into new genotypes within just two weeks in other insect species (Beaty et al., 1985), underscoring the urgency of deeper investigation into their genetic diversity and potential impact.

While some thrips species are known to feed on fungi, the detection of potential mycoviruses in this study was unexpected. One possibility is that these viruses were present in the insect's gut at the time of sampling or adhered to the insect's surface. In nature, herbivorous insects can acquire viruses through sap-sucking and foliage feeding (Ma et al., 2021), suggesting multiple possible routes of viral acquisition. It is also plausible, though highly unlikely, these viruses may exist in both insect and fungal hosts, a cross-kingdom association that has yet to be fully explored.

The best matches of SVNV for S, M and L segments were previously discovered in the USA. This result was expected because SVNV is the prevalent viral disease that infects soybean and replicates in soybean thrips. The discovery of Soybean mosaic viruses, SMV, (2256bp, PQ492253) is surprising because soybean thrips do not transmit SMV, as it is transmitted by aphids in a non-persistent manner and through seeds (Hajimorad et al., 2018). The occurrence of SMV in the system of *N. variabilis* suggests that the thrips might have ingested it through feeding on soybeans that were infected with SMV. Previous research suggests that the intestinal tract is a major route of entry for external viruses when insects accidentally ingest contaminated materials (Ma et al., 2021).

Beyond plant-associated viruses, this study also detected viral sequences from the order Durnavirales, expanding the known viral diversity associated with *N. variabilis*. Viruses from this order have previously been identified in the viromes of marine invertebrates across multiple seas (Zhang et al., 2022), as well as in a wide range of hosts and environments, including animal stool samples, birds, reptiles, fish, invertebrates, microbial communities, and environmental sources (Sadiq et al., 2024). Some families within this order are thought to include bacteriophages, particularly those associated with the gastrointestinal systems of terrestrial mammals and birds. In our study, we also identified a virus belonging to Durnavirales that may represent a bacteriophage. Given its potential to interact with microbial communities in the insect intestinal system, such a virus could be explored for its biocontrol potential. Modulating the gut microbiota of *N. variabilis* could, in turn, affect digestion and potentially influence vector competence. Prior research has shown that viruses can alter insect feeding behavior, fecundity, and lifespan (Shalileh et al., 2016; Stafford et al., 2011; Stafford-Banks et al., 2014), supporting the idea that insect-associated viruses may play a role in regulating host biology. These findings point to an intriguing, though still largely unexplored, avenue for understanding virus–insect interactions and their potential applications in population management.

Conclusion

This study identified a diverse set of viruses associated with soybean thrips (*N. variabilis*) including potential plant-infecting viruses, insect-specific viruses, dual plant-insect viruses, and putative mycoviruses. The presence of these viral types suggests multiple acquisition routes, such as feeding or vertical transmission. Notably, the majority of identified viruses belonged to the order Bunyavirales, highlighting its prominence in the thrips virome. Additionally, RNA viruses were more prevalent than DNA viruses, and a significant portion of sequences remained taxonomically unclassified, indicating a high level of unexplored viral diversity. These findings contribute valuable insight into the viral community structure associated with *N. variabilis* and underscore the complexity of virus–insect–environment interactions.

References

Baker, C., Blount, A., & Quesenberry, K. (1999). *Peanut Stunt Virus Infecting Perennial Peanuts in Florida and Georgia*. https://ccmedia.fdacs.gov/content/download/11401/file/pp395.pdf?utm_source=chatgpt.com

Barker, J., daSilva, L. L. P., & Crump, C. M. (2023). Mechanisms of bunyavirus morphogenesis and egress. *Journal of General Virology*, *104*(4). https://doi.org/10.1099/jgv.0.001845

Barkley, N. A., Pinnow, D. L., Wang, M. L., & Pederson, G. A. (2009). First Report of *Tomato spotted wilt virus* Infecting African Clover (*Trifolium tembense*) in Georgia. *Plant Disease*, *93*(2), 202–202. https://doi.org/10.1094/PDIS-93-2-0202C

Batson, J., Dudas, G., Haas-Stapleton, E., Kistler, A. L., Li, L. M., Logan, P., Ratnasiri, K., & Retallack, H. (2021). Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. *eLife*, *10*, e68353. https://doi.org/10.7554/eLife.68353

Beach, A. M. (1897). Contributions to a knowledge of the Thripidae of Iowa. *The Proceedings of the Iowa Academy of Science.*, *3*, 214–227. https://doi.org/10.5962/bhl.part.14877

Brait, N., Hackl, T., Morel, C., Exbrayat, A., Gutierrez, S., & Lequime, S. (2024). A tale of caution: How endogenous viral elements affect virus discovery in transcriptomic data. *Virus Evolution*, *10*(1), vead088. https://doi.org/10.1093/ve/vead088

Brault, V., Mutterer, J., Scheidecker, D., Simonis, M. T., Herrbach, E., Richards, K., & Ziegler-Graff, V. (2000). Effects of Point Mutations in the Readthrough Domain of the Beet Western Yellows Virus Minor Capsid Protein on Virus Accumulation In Planta and on

Transmission by Aphids. *Journal of Virology*, *74*(3), 1140–1148. https://doi.org/10.1128/JVI.74.3.1140-1148.2000

Bruyère, A., Brault, V., Ziegler-Graff, V., Simonis, M.-T., Van Den Heuvel, J. F. J. M., Richards, K., Guilley, H., Jonard, G., & Herrbach, E. (1997). Effects of Mutations in the Beet Western Yellows Virus Readthrough Protein on Its Expression and Packaging and on Virus Accumulation, Symptoms, and Aphid Transmission. *Virology*, *230*(2), 323–334. https://doi.org/10.1006/viro.1997.8476

Burand, J. P., Kim, W., Afonso, C. L., Tulman, E. R., Kutish, G. F., Lu, Z., & Rock, D. L. (2012). Analysis of the genome of the sexually transmitted insect virus Helicoverpa zea nudivirus 2. *Viruses*, *4*(1), 28–61. https://doi.org/10.3390/v4010028

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009a). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009b). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009c). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Card, S. D., Pearson, M. N., & Clover, G. R. G. (2007). Plant pathogens transmitted by pollen. *Australasian Plant Pathology*, *36*(5), 455. https://doi.org/10.1071/AP07050

Chen, A. Y. S., Walker, G. P., Carter, D., & Ng, J. C. K. (2011). A virus capsid component mediates virion retention and transmission by its insect vector. *Proceedings of the National Academy of Sciences*, *108*(40), 16777–16782. https://doi.org/10.1073/pnas.1109384108

Chen, Y., Evans, J., & Feldlaufer, M. (2006). Horizontal and vertical transmission of viruses in the honey bee, Apis mellifera. *Journal of Invertebrate Pathology*, *92*(3), 152–159. https://doi.org/10.1016/j.jip.2006.03.010

Chen, Y.-M., Sadiq, S., Tian, J.-H., Chen, X., Lin, X.-D., Shen, J.-J., Chen, H., Hao, Z.-Y., Wille, M., Zhou, Z.-C., Wu, J., Li, F., Wang, H.-W., Yang, W.-D., Xu, Q.-Y., Wang, W., Gao, W.-H., Holmes, E. C., & Zhang, Y.-Z. (2022). RNA viromes from terrestrial sites across China expand environmental viral diversity. *Nature Microbiology*, *7*(8), 1312–1323. https://doi.org/10.1038/s41564-022-01180-2

Cobbin, J. C., Charon, J., Harvey, E., Holmes, E. C., & Mahar, J. E. (2021). Current challenges to virus discovery by meta-transcriptomics. *Current Opinion in Virology*, *51*, 48–55. https://doi.org/10.1016/j.coviro.2021.09.007

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015a). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015b). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Common types of viruses of floriculture crops and their modes of transmission. (2014, April 3). MSU Extension.

https://www.canr.msu.edu/news/common_types_of_viruses_of_floriculture_crops_and_their_modes_of_transmissi

Conner, K., Sikora, E. J., Zhang, L., & Burmester, C. (2013). First Report of *Soybean vein necrosis-associated virus* Affecting Soybeans in Alabama. *Plant Health Progress*, *14*(1), 49. https://doi.org/10.1094/PHP-2013-0729-03-BR

Culbreath, A. K., Todd, J. W., & Brown, S. L. (2003). E PIDEMIOLOGY AND M ANAGEMENT OF T OMATO S POTTED W ILT IN P EANUT. *Annual Review of Phytopathology*, *41*(1), 53–75. https://doi.org/10.1146/annurev.phyto.41.052002.095522

Dietzgen, R., Mann, K., & Johnson, K. (2016). Plant Virus–Insect Vector Interactions: Current and Potential Future Research Directions. *Viruses*, *8*(11), 303. https://doi.org/10.3390/v8110303

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Elena, S. F., Fraile, A., & García-Arenal, F. (2014). Evolution and Emergence of Plant Viruses. In *Advances in Virus Research* (Vol. 88, pp. 161–191). Elsevier. https://doi.org/10.1016/B978-0-12-800098-4.00003-9

Feng, Y., Krueger, E. N., Liu, S., Dorman, K., Bonning, B. C., & Miller, W. A. (2017). Discovery of Known and Novel Viral Genomes in Soybean Aphid by Deep Sequencing. *Phytobiomes Journal*, *1*(1), 36–45. https://doi.org/10.1094/PBIOMES-11-16-0013-R

Fletcher, J., Tang, J., Blouin, A., Ward, L., MacDiarmid, R., & Ziebell, H. (2016). *Red clover vein mosaic virus*—A Novel Virus to New Zealand that is Widespread in Legumes. *Plant Disease*, *100*(5), 890–895. https://doi.org/10.1094/PDIS-04-15-0465-RE

Gómez, M., Martínez, D., Páez-Triana, L., Luna, N., De Las Salas, J. L., Hernández, C., Flórez, A. Z., Muñoz, M., & Ramírez, J. D. (2023). Characterizing viral species in mosquitoes (Culicidae) in the Colombian Orinoco: Insights from a preliminary metagenomic study. *Scientific Reports*, *13*(1), 22081. https://doi.org/10.1038/s41598-023-49232-9

Gordon, A. (2010). *FASTQ/A short reads pre-processing tools.* [Computer software]. http://hannonlab.cshl.edu/fastx_toolkit/

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011a). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011b). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Hajimorad, M. R., Domier, L. L., Tolin, S. A., Whitham, S. A., & Saghai Maroof, M. A. (2018). *Soybean mosaic virus*: A successful potyvirus with a wide distribution but restricted natural

host range. *Molecular Plant Pathology*, *19*(7), 1563–1579. https://doi.org/10.1111/mpp.12644

Harrewijn, P., Tjallingii, W. F., & Mollema, C. (1996). Electrical recording of plant penetration by western flower thrips. *Entomologia Experimentalis et Applicata*, *79*(3), 345–353. https://doi.org/10.1111/j.1570-7458.1996.tb00842.x

Hasiów-Jaroszewska, B., Boezen, D., & Zwart, M. P. (2021). Metagenomic Studies of Viruses in Weeds and Wild Plants: A Powerful Approach to Characterise Variable Virus Communities. *Viruses*, *13*(10), 1939. https://doi.org/10.3390/v13101939

History and Economic Impact. (2025, March 14). https://tswv.caes.uga.edu/usda-ramp-project/history-and-economic-impact.html?utm_source=chatgpt.com

Hogenhout, S. A., Ammar, E.-D., Whitfield, A. E., & Redinbaugh, M. G. (2008). Insect Vector Interactions with Persistently Transmitted Viruses. *Annual Review of Phytopathology*, *46*(1), 327–359. https://doi.org/10.1146/annurev.phyto.022508.092135

Hunter, W. B., & Ullman, D. E. (1989). Analysis of mouthpart movements during feeding of Frankliniella occidentalis (pergande) and F. schultzei trybom (Thysanoptera: Thripidae). *International Journal of Insect Morphology and Embryology*, *18*(2–3), 161–171. https://doi.org/10.1016/0020-7322(89)90024-X

Hýsková, V., Bělonožníková, K., Chmelík, J., Hoffmeisterová, H., Čeřovská, N., Moravec, T., & Ryšlavá, H. (2024). Potyviral Helper-Component Protease: Multifaced Functions and Interactions with Host Proteins. *Plants*, *13*(9), 1236. https://doi.org/10.3390/plants13091236

Jia, H., & Gong, P. (2019). A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. *Frontiers in Microbiology*, *10*, 1945. https://doi.org/10.3389/fmicb.2019.01945

Jones, D. R. (2005). *Plant viruses transmitted by thrips*. *113*, 119–157. https://doi.org/10.1007/s10658-005-2334-1

Junglen, S., & Drosten, C. (2013). Virus discovery and recent insights into virus diversity in arthropods. *Current Opinion in Microbiology*, *16*(4), 507–513. https://doi.org/10.1016/j.mib.2013.06.005

Käfer, S., Paraskevopoulou, S., Zirkel, F., Wieseke, N., Donath, A., Petersen, M., Jones, T. C., Liu, S., Zhou, X., Middendorf, M., Junglen, S., Misof, B., & Drosten, C. (2019). Re-assessing the diversity of negative strand RNA viruses in insects. *PLOS Pathogens*, *15*(12), e1008224. https://doi.org/10.1371/journal.ppat.1008224

Keough, S., Danielson, J., Marshall, J. M., Lagos-Kutz, D., Voegtlin, D. J., Srinivasan, R., & Nachappa, P. (2018). Factors Affecting Population Dynamics of Thrips Vectors of Soybean vein necrosis virus. *Environmental Entomology*, *47*(3), 734–740. https://doi.org/10.1093/ee/nvy021

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009a). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009b). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Lanzi, G., de Miranda, J. R., Boniotti, M. B., Cameron, C. E., Lavazza, A., Capucci, L., Camazine, S. M., & Rossi, C. (2006). Molecular and biological characterization of deformed wing virus of honeybees (Apis mellifera L.). *Journal of Virology*, *80*(10), 4998–5009. https://doi.org/10.1128/JVI.80.10.4998-5009.2006

LaTora, A. G., Lai, P.-C., Chen, Y.-J., Gautam, S., Abney, M. R., & Srinivasan, R. (2022a). *Frankliniella fusca* (Thysanoptera: Thripidae), The Vector of *Tomato Spotted Wilt Orthotospovirus* Infecting Peanut in the Southeastern United States. *Journal of Integrated Pest Management*, *13*(1), 3. https://doi.org/10.1093/jipm/pmab045

LaTora, A. G., Lai, P.-C., Chen, Y.-J., Gautam, S., Abney, M. R., & Srinivasan, R. (2022b). *Frankliniella fusca* (Thysanoptera: Thripidae), The Vector of *Tomato Spotted Wilt Orthotospovirus* Infecting Peanut in the Southeastern United States. *Journal of Integrated Pest Management*, *13*(1), 3. https://doi.org/10.1093/jipm/pmab045

Lawrence, P., Heung, M., Nave, J., Henkel, C., & Escudero-Pérez, B. (2023). The natural virome and pandemic potential: Disease X. *Current Opinion in Virology*, *63*, 101377. https://doi.org/10.1016/j.coviro.2023.101377

Liu, S., Vijayendr, D., & Bonning, B. C. (2011). *Next Generation Sequencing Technologies for Insect Virus Discovery*. *3*, 1849–1869. https://doi.org/10.3390/v3101849

Looi, L. M., & Chua, K. B. (2007). Lessons from the Nipah virus outbreak in Malaysia. *Malaysian Journal of Pathology*, *29*(2), 63–67.

Ma, Y., Marais, A., Lefebvre, M., Faure, C., & Candresse, T. (2020). Metagenomic analysis of virome cross-talk between cultivated Solanum lycopersicum and wild Solanum nigrum. *Virology*, *540*, 38–44. https://doi.org/10.1016/j.virol.2019.11.009

Marklewitz, M., Zirkel, F., Kurth, A., Drosten, C., & Junglen, S. (2015). Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. *Proceedings of the National Academy of Sciences*, *112*(24), 7536–7541. https://doi.org/10.1073/pnas.1502036112

Marklewitz, M., Zirkel, F., Rwego, I. B., Heidemann, H., Trippner, P., Kurth, A., Kallies, R., Briese, T., Lipkin, W. I., Drosten, C., Gillespie, T. R., & Junglen, S. (2013). Discovery of a Unique Novel Clade of Mosquito-Associated Bunyaviruses. *Journal of Virology*, *87*(23), 12850–12865. https://doi.org/10.1128/JVI.01862-13

McLaughlin, M. R. (1992). Virus Diseases and Stand Decline in a White Clover Pasture. *Plant Disease*, *76*(2), 158. https://doi.org/10.1094/PD-76-0158

Mokili, J. L., Rohwer, F., & Dutilh, B. E. (2012). Metagenomics and future perspectives in virus discovery. *Current Opinion in Virology*, *2*(1), 63–77. https://doi.org/10.1016/j.coviro.2011.12.004

Nault, B. A., & Speese Iii, J. (2002). Major insect pests and economics of fresh-market tomato in eastern Virginia. *Crop Protection*, *21*(5), 359–366. https://doi.org/10.1016/S0261-2194(01)00103-X

Ng, T. F. F., Duffy, S., Polston, J. E., Bixby, E., Vallad, G. E., & Breitbart, M. (2011). Exploring the Diversity of Plant DNA Viruses and Their Satellites Using Vector-Enabled

Metagenomics on Whiteflies. *PLoS ONE*, *6*(4), e19050. https://doi.org/10.1371/journal.pone.0019050

Omkar. (2021). *Polyphagous Pests of Crops*. Springer Singapore Pte. Limited.

Paliwal, Y. C. (1974). Some properties and thrip transmission of tomato spotted wilt virus in Canada. *Canadian Journal of Botany*, *52*(6), 1177–1182. https://doi.org/10.1139/b74-151

Pan, S. (2024, September 25). Insects as Vectors – Features, Types, Examples. *Insects as Vectors – Features, Types, Examples*. https://biologynotesonline.com/insects-as-vectors-features-types-examples/

Peter, K. A., Liang, D., Palukaitis, P., & Gray, S. M. (2008). Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. *Journal of General Virology*, *89*(8), 2037–2045. https://doi.org/10.1099/vir.0.83625-0

Pirone, T. P., & Megahed, E.-S. (1966). Aphid transmissibility of some purified viruses and viral RNA's. *Virology*, *30*(4), 631–637. https://doi.org/10.1016/0042-6822(66)90168-1

Power, A. G., & Mitchell, C. E. (2004). Pathogen Spillover in Disease Epidemics. *The American Naturalist*, *164*(S5), S79–S89. https://doi.org/10.1086/424610

Qi, Y.-H., Ye, Z.-X., Zhang, C.-X., Chen, J.-P., & Li, J.-M. (2023). Diversity of RNA viruses in agricultural insects. *Computational and Structural Biotechnology Journal*, *21*, 4312–4321. https://doi.org/10.1016/j.csbj.2023.08.036

Riley, D. G., Joseph, S. V., Srinivasan, R., & Diffie, S. (2011). Thrips Vectors of Tospoviruses. *Journal of Integrated Pest Management*, *2*(1), I1–I10. https://doi.org/10.1603/IPM10020

Roossinck, M. J., & García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Current Opinion in Virology*, *10*, 56–62. https://doi.org/10.1016/j.coviro.2015.01.005

Rosario, K., Capobianco, H., Ng, T. F. F., Breitbart, M., & Polston, J. E. (2014). RNA Viral Metagenome of Whiteflies Leads to the Discovery and Characterization of a Whitefly-Transmitted Carlavirus in North America. *PLoS ONE*, *9*(1), e86748. https://doi.org/10.1371/journal.pone.0086748

Rosario, K., Seah, Y., Marr, C., Varsani, A., Kraberger, S., Stainton, D., Moriones, E., Polston, J., Duffy, S., & Breitbart, M. (2015). Vector-Enabled Metagenomic (VEM) Surveys Using Whiteflies (Aleyrodidae) Reveal Novel Begomovirus Species in the New and OldWorlds. *Viruses*, *7*(10), 5553–5570. https://doi.org/10.3390/v7102895

Rozo-Lopez, P., Brewer, W., Käfer, S., Martin, M. M., & Parker, B. J. (2023). Untangling an insect's virome from its endogenous viral elements. *BMC Genomics*, *24*(1), 636. https://doi.org/10.1186/s12864-023-09737-z

Soybean Thrips. (n.d.). Purdue University - College of Agriculture. Retrieved March 14, 2025, from https://ag.purdue.edu/department/entm/extension/field-crops-ipm/soybean/soybean-thrips.html

Soybean Vein Necrosis Virus—Soybean Disease—Soybean Research & Information Network—SRIN. (n.d.). Soybean Research & Information Network. Retrieved March 29, 2025, from https://soybeanresearchinfo.com/soybean-disease/soybean-vein-necrosis-virus/

Srinivasan, R., Abney, M. R., Culbreath, A. K., Kemerait, R. C., Tubbs, R. S., Monfort, W. S., & Pappu, H. R. (2017). Three decades of managing Tomato spotted wilt virus in peanut in

southeastern United States. *Virus Research*, *241*, 203–212. https://doi.org/10.1016/j.virusres.2017.05.016

Stafford, C. A., Walker, G. P., & Ullman, D. E. (2012). Hitching a ride: Vector feeding and virus transmission. *Communicative & Integrative Biology*, *5*(1), 43–49. https://doi.org/10.4161/cib.18640

Stewart, L. R., Medina, V., Tian, T., Turina, M., Falk, B. W., & Ng, J. C. K. (2010). A Mutation in the *Lettuce Infectious Yellows Virus* Minor Coat Protein Disrupts Whitefly Transmission but Not *In Planta* Systemic Movement. *Journal of Virology*, *84*(23), 12165–12173. https://doi.org/10.1128/JVI.01192-10

The PLOS Pathogens Staff. (2014). Correction: Transovarial Transmission of a Plant Virus Is Mediated by Vitellogenin of Its Insect Vector. *PLoS Pathogens*, *10*(4), e1004141. https://doi.org/10.1371/journal.ppat.1004141

Thekke-Veetil, T., Lagos-Kutz, D., Domier, L. L., McCoppin, N. K., Hartman, G. L., & Clough, S. J. (2024). Exploring Virus Diversity in the Potato leafhopper (Empoasca fabae), an Economically Important Agricultural Pest. *Viruses*, *16*(8), 1305. https://doi.org/10.3390/v16081305

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020a). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020b). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly

Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020c). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thrips Management Guidelines—UC IPM. (n.d.). Retrieved April 6, 2025, from https://ipm.ucanr.edu/PMG/PESTNOTES/pn7429.html?utm_source=chatgpt.com

Tobacco thrips. (2025, February 27). *Tobacco Thrips*. https://www.biobee.com/pests/tobacco-thrips/

Umar, M., Tegg, R. S., Farooq, T., Thangavel, T., & Wilson, C. R. (2022). Abundance of Poleroviruses within Tasmanian Pea Crops and Surrounding Weeds, and the Genetic Diversity of TuYV Isolates Found. *Viruses,* *14*(8), 1690. https://doi.org/10.3390/v14081690

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., Guzman, H., Wood, T. G., Popov, V., Gorchakov, R., González, A. V., Haddow, A. D., Watts, D. M., Da Rosa, A. P. A. T., Weaver, S. C., Lipkin, W. I., & Tesh, R. B. (2013). Negevirus: A Proposed New Taxon of Insect-Specific Viruses with Wide Geographic Distribution. *Journal of Virology*, *87*(5), 2475–2488. https://doi.org/10.1128/JVI.00776-12

Vasilakis, N., Guzman, H., Firth, C., Forrester, N. L., Widen, S. G., Wood, T. G., Rossi, S. L., Ghedin, E., Popov, V., Blasdell, K. R., Walker, P. J., & Tesh, R. B. (2014). Mesoniviruses

are mosquito-specific viruses with extensive geographic distribution and host range. *Virology Journal*, *11*(1), 97. https://doi.org/10.1186/1743-422X-11-97

Venkataraman, S., Prasad, B., & Selvarajan, R. (2018). RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. *Viruses*, *10*(2), 76. https://doi.org/10.3390/v10020076

Wamonje, F. O., Michuki, G. N., Braidwood, L. A., Njuguna, J. N., Musembi Mutuku, J., Djikeng, A., Harvey, J. J. W., & Carr, J. P. (2017). Viral metagenomics of aphids present in bean and maize plots on mixed-use farms in Kenya reveals the presence of three dicistroviruses including a novel Big Sioux River virus-like dicistrovirus. *Virology Journal*, *14*(1), 188. https://doi.org/10.1186/s12985-017-0854-x

Wang, Z., Gerstein, M., & Snyder, M. (2009). *RNA-Seq: A revolutionary tool for transcriptomics*. *10*, 57–63.

Wendimu, G. Y., & Gurmu, A. K. (2024). Insect Vectors of Plant Viruses: Host Interactions, Their Effects, and Future Opportunities. *2024*, *2024*(1). https://doi.org/10.1155/aia/6006985

Whitfield, A. E., Falk, B. W., & Rotenberg, D. (2015a). Insect vector-mediated transmission of plant viruses. *Virology*, *479–480*, 278–289. https://doi.org/10.1016/j.virol.2015.03.026

Whitfield, A. E., Falk, B. W., & Rotenberg, D. (2015b). *Insect vector-mediated transmission of plant viruses*. *479*, 278–289.

Wu, H., Pang, R., Cheng, T., Xue, L., Zeng, H., Lei, T., Chen, M., Wu, S., Ding, Y., Zhang, J., Shi, M., & Wu, Q. (2020). Abundant and Diverse RNA Viruses in Insects Revealed by RNA-Seq Analysis: Ecological and Evolutionary Implications. *mSystems*, *5*(4), e00039-20. https://doi.org/10.1128/mSystems.00039-20

Xie, Y.-H., Li, Z.-Y., Dong, K., & Zhang, H.-R. (2012). Changes in the species composition of thrips on *Trifolium repens* (Fabales). *Acta Phytopathologica et Entomologica Hungarica*, *47*(1), 61–67. https://doi.org/10.1556/APhyt.47.2012.1.7

Zhou, J., Kantartzi, S. K., Wen, R.-H., Newman, M., Hajimorad, M. R., Rupe, J. C., & Tzanetakis, I. E. (2011). Molecular characterization of a new tospovirus infecting soybean. *Virus Genes*, *43*(2), 289–295. https://doi.org/10.1007/s11262-011-0621-9

Zhou, J., & Tzanetakis, I. E. (2013). Epidemiology of Soybean vein necrosis-associated virus. *Phytopathology®*, *103*(9), 966–971. https://doi.org/10.1094/PHYTO-12-12-0322-R

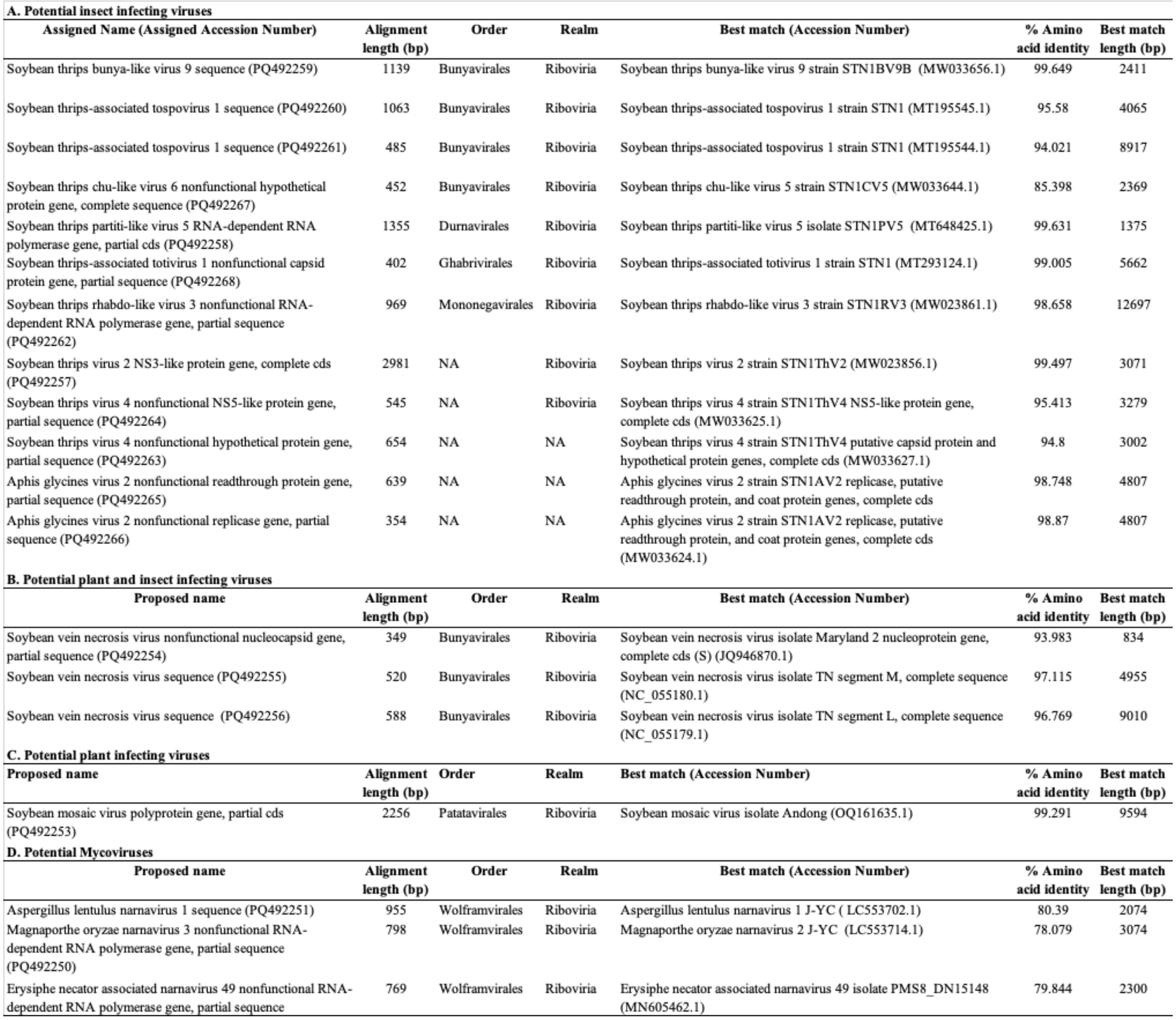
Tables

Table 1. Sequence statistics of the reads generated RNA-seq analysis of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama in 2023

	Statistics
Number of total reads	1235537868
Total number of forward and reverse reads	617768934
Length range of the read before trimming (bp)	151
GC% before trimming	57
Number of bases (Gbp)	186.5
Sequences matched to the reference genome Tobacco thrips *Frankliniella fusca* (GCA_029379305.1) %	29.21
Sequences unmatched to the reference genome %	70.79
Number of trinity transcripts	439109
GC% of Trinity transcripts	47.5
Average contig length (bp)	1003.32
Average assembled base for the contig	440568337
Contig N10	12588
Contig N20	8890
Contig N30	6528

Contig N50	2998
Number of viral contigs	3971

Table 2. Diversity of viruses discovered through RNA-seq analysis of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama in 2023



A. Potential insect infecting viruses

Assigned Name (Assigned Accession Number)	Alignment length (bp)	Order	Realm	Best match (Accession Number)	% Amino acid identity	Best match length (bp)
Soybean thrips bunya-like virus 9 sequence (PQ492259)	1139	Bunyavirales	Riboviria	Soybean thrips bunya-like virus 9 strain STN1BV9B (MW033656.1)	99.649	2411
Soybean thrips-associated tospovirus 1 sequence (PQ492260)	1063	Bunyavirales	Riboviria	Soybean thrips-associated tospovirus 1 strain STN1 (MT195545.1)	95.58	4065
Soybean thrips-associated tospovirus 1 sequence (PQ492261)	485	Bunyavirales	Riboviria	Soybean thrips-associated tospovirus 1 strain STN1 (MT195544.1)	94.021	8917
Soybean thrips chu-like virus 6 nonfunctional hypothetical protein gene, complete sequence (PQ492267)	452	Bunyavirales	Riboviria	Soybean thrips chu-like virus 5 strain STN1CV5 (MW033644.1)	85.398	2369
Soybean thrips partiti-like virus 5 RNA-dependent RNA polymerase gene, partial cds (PQ492258)	1355	Durnavirales	Riboviria	Soybean thrips partiti-like virus 5 isolate STN1PV5 (MT648425.1)	99.631	1375
Soybean thrips-associated totivirus 1 nonfunctional capsid protein gene, partial sequence (PQ492268)	402	Ghabrivirales	Riboviria	Soybean thrips-associated totivirus 1 strain STN1 (MT293124.1)	99.005	5662
Soybean thrips rhabdo-like virus 3 nonfunctional RNA-dependent RNA polymerase gene, partial sequence (PQ492262)	969	Mononegavirales	Riboviria	Soybean thrips rhabdo-like virus 3 strain STN1RV3 (MW023861.1)	98.658	12697
Soybean thrips virus 2 NS3-like protein gene, complete cds (PQ492257)	2981	NA	Riboviria	Soybean thrips virus 2 strain STN1ThV2 (MW023856.1)	99.497	3071
Soybean thrips virus 4 nonfunctional NS5-like protein gene, partial sequence (PQ492264)	545	NA	Riboviria	Soybean thrips virus 4 strain STN1ThV4 NS5-like protein gene, complete cds (MW033625.1)	95.413	3279
Soybean thrips virus 4 nonfunctional hypothetical protein gene, partial sequence (PQ492263)	654	NA	NA	Soybean thrips virus 4 strain STN1ThV4 putative capsid protein and hypothetical protein genes, complete cds (MW033627.1)	94.8	3002
Aphis glycines virus 2 nonfunctional readthrough protein gene, partial sequence (PQ492265)	639	NA	NA	Aphis glycines virus 2 strain STN1AV2 replicase, putative readthrough protein, and coat protein genes, complete cds	98.748	4807
Aphis glycines virus 2 nonfunctional replicase gene, partial sequence (PQ492266)	354	NA	NA	Aphis glycines virus 2 strain STN1AV2 replicase, putative readthrough protein, and coat protein genes, complete cds (MW033624.1)	98.87	4807

B. Potential plant and insect infecting viruses

Proposed name	Alignment length (bp)	Order	Realm	Best match (Accession Number)	% Amino acid identity	Best match length (bp)
Soybean vein necrosis virus nonfunctional nucleocapsid gene, partial sequence (PQ492254)	349	Bunyavirales	Riboviria	Soybean vein necrosis virus isolate Maryland 2 nucleoprotein gene, complete cds (S) (JQ946870.1)	93.983	834
Soybean vein necrosis virus sequence (PQ492255)	520	Bunyavirales	Riboviria	Soybean vein necrosis virus isolate TN segment M, complete sequence (NC_055180.1)	97.115	4955
Soybean vein necrosis virus sequence (PQ492256)	588	Bunyavirales	Riboviria	Soybean vein necrosis virus isolate TN segment L, complete sequence (NC_055179.1)	96.769	9010

C. Potential plant infecting viruses

Proposed name	Alignment length (bp)	Order	Realm	Best match (Accession Number)	% Amino acid identity	Best match length (bp)
Soybean mosaic virus polyprotein gene, partial cds (PQ492253)	2256	Patatavirales	Riboviria	Soybean mosaic virus isolate Andong (OQ161635.1)	99.291	9594

D. Potential Mycoviruses

Proposed name	Alignment length (bp)	Order	Realm	Best match (Accession Number)	% Amino acid identity	Best match length (bp)
Aspergillus lentulus narnavirus 1 sequence (PQ492251)	955	Wolframvirales	Riboviria	Aspergillus lentulus narnavirus 1 J-YC (LC553702.1)	80.39	2074
Magnaporthe oryzae narnavirus 3 nonfunctional RNA-dependent RNA polymerase gene, partial sequence (PQ492250)	798	Wolframvirales	Riboviria	Magnaporthe oryzae narnavirus 2 J-YC (LC553714.1)	78.079	3074
Erysiphe necator associated narnavirus 49 nonfunctional RNA-dependent RNA polymerase gene, partial sequence	769	Wolframvirales	Riboviria	Erysiphe necator associated narnavirus 49 isolate PMS8_DN15148 (MN605462.1)	79.844	2300

Figures

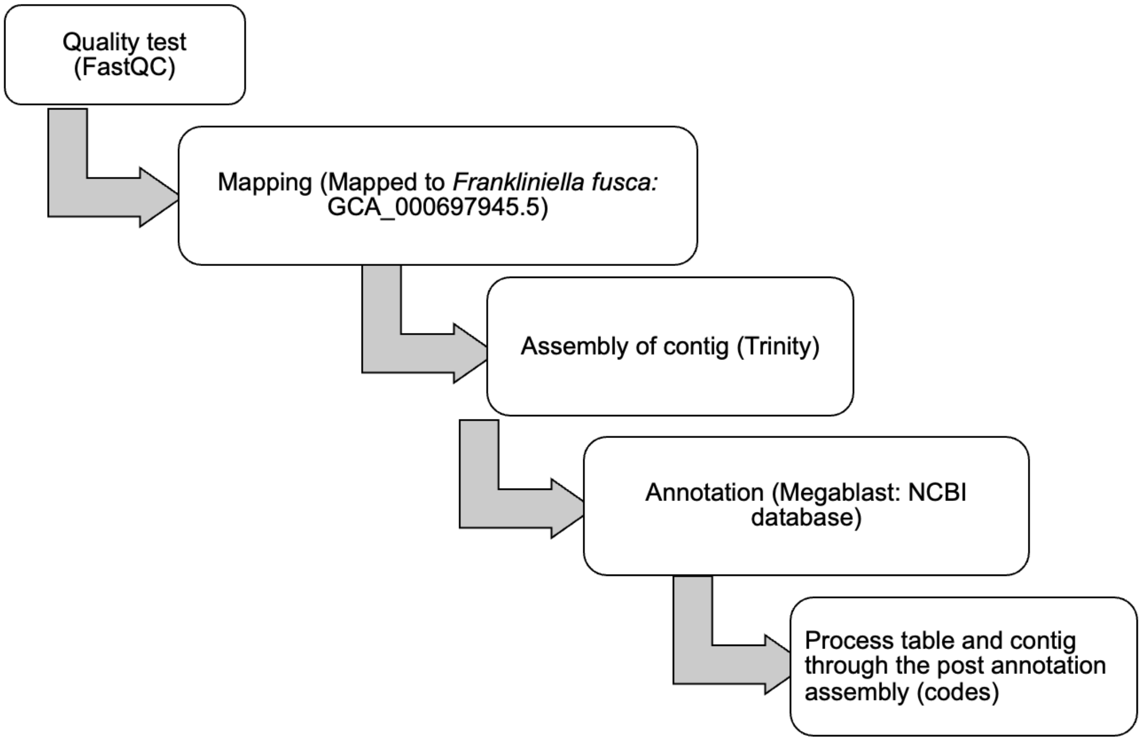


Figure 1. Bioinformatic pipeline to analyze the RNA-seq data of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama in 2023

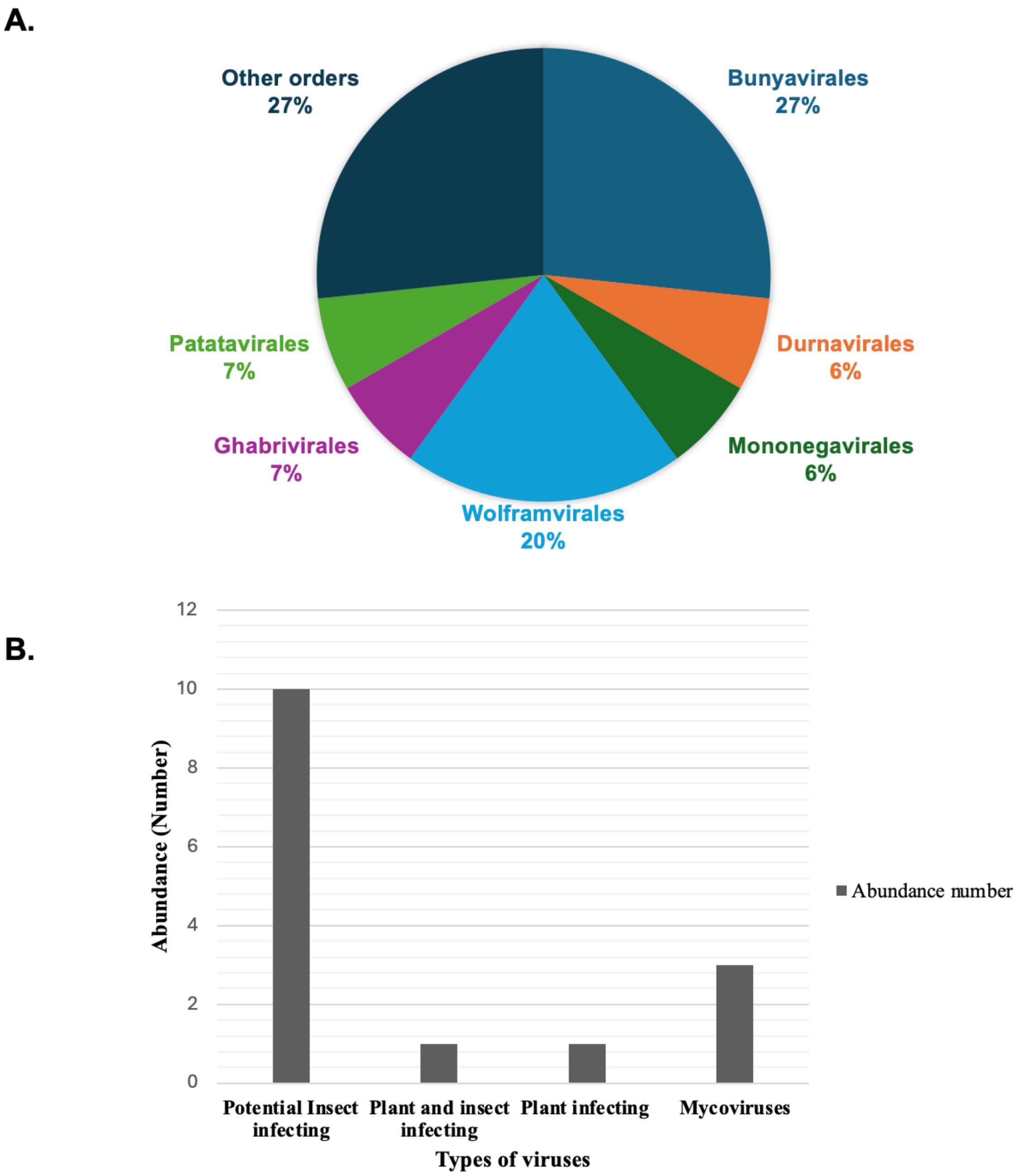


Figure 2. Virus statistics; From top to bottom A. Virus abundance (%) at order level discovered from RNA-seq analysis of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama 2023 B. Graph of abundance of types of viruses discovered from RNA-seq analysis of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama in 2023 as potential insect infecting, potential plant and insect infecting, plant infecting and mycoviruses

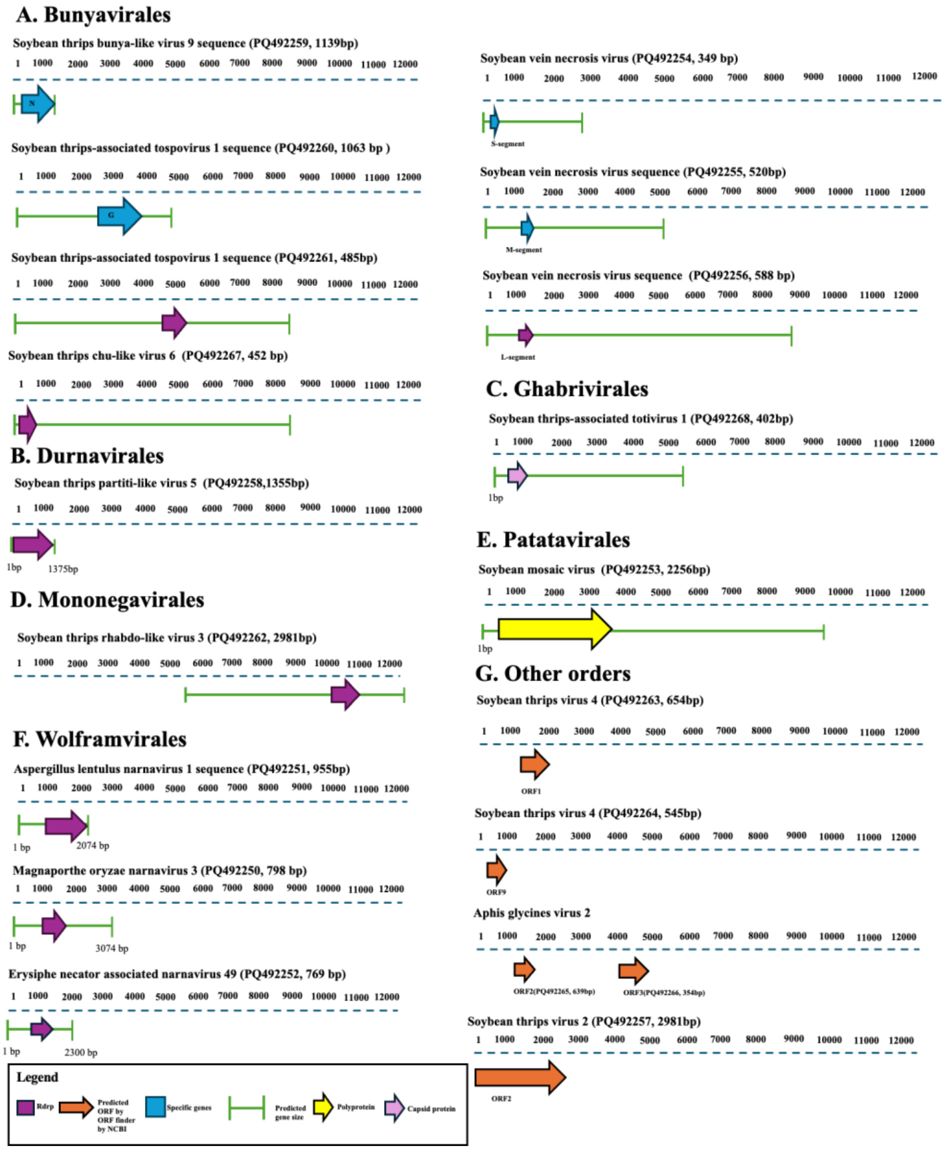


Figure 3. Graphical representation viruses from RNA-seq analysis of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama in 2023. The arrow represents different genes of viruses belonging to (A) Bunyavirales; (B) Durnavirales; (C) Ghabrivirales; (D) Mononegavirales; (E) Patatavirales; (F) Wolframvirales; (G) Other orders; ORF: Open reading frames; RdRP: RNA-dependent RNA polymerase; bp: base pair; NCBI: National Center for Biotechnology Information

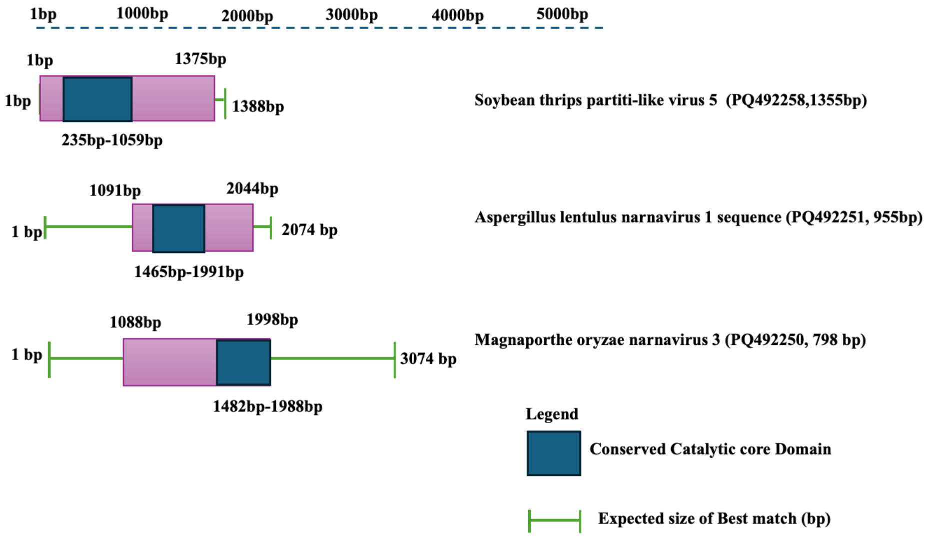


Figure 4. Viruses sharing the Conserved catalytic core domain predicted by NCBI Conserved Catalytic Domain through RNA-seq analysis of soybean thrips (*Neohydatothrips variabilis* Beach) collected in Auburn, Alabama in 2023; soybean thrips partiti-like virus 5 (PQ492258,1355bp; Aspergillus lentulus narnavirus 1 sequence (PQ492251, 955bp); Magnaporthe oryzae narnavirus 3 (PQ492250, 798 bp); bp: base pair; NCBI: National Center for Biotechnology Information

Chapter 3

Discovery of twenty-five different viruses in Tobacco Thrips (*Frankliniella fusca* Hinds) from Peanut and Clover in Headland, Alabama in 2022 and 2023, through RNA-seq analysis

Aasma Sharma[1], Claire Cook[1], Scott H Graham[1], Alana Jacobson[1], Kathleen M. Martin[1]

[1] Auburn University

Abstract

Tobacco thrips (*Frankliniella fusca* Hinds*)* are major agricultural pests and known vectors of plant viruses, particularly in peanut and clover. This study employed RNA sequencing (RNA-seq) to investigate the virome of *F. fusca* collected from Headland, Alabama in peanut and clover. We collected and sequenced 80-85 tobacco thrips from each peanut and clover in year 2022 and 2023. From RNA-seq analysis ten viruses were discovered from clover in 2023, three viruses were discovered from clover in 2022, four viruses were discovered from peanut 2023, four viruses were discovered from peanut in 2022. Most of the viruses discovered belong to the order *Bunyavirales* (33% in *F. fusca* sampled from peanut in 2022, 67% in *F. fusca* sampled from peanut in 2023, 67% in *F. fusca* sampled from clover in 2022, 41% in *F. fusca* sampled from clover in 2023). *F. fusca* collected from peanuts in 2022 and 2023 shared one virus, from clover in 2022 and 2023 is one virus, from clover and peanuts in 2022 is one virus, and from clover and peanuts in 2023 is four viruses. *Orthotospo tomatomaculae* was the common virus in year in both crops. In total *F. fusca* collected in clover has more diversity of viruses in 2023 compared to peanut. This study suggests that *F. fusca* that feeds on clover plant might have higher virus diversity compared to its crop host peanut.

Keywords: Tobacco thrips, peanuts, clover, RNA-seq, virome

Introduction

Tobacco thrips (*Frankliniella fusca* Hinds; Thysanoptera; Thripidae) are an important phytophagous pest and key insect vector of different plant viral diseases. It is distributed across the world and feeds on 57 plant species in 11 families including peanuts and clover. The economic damage is caused by the feeding damage and transmission of major plant viruses in the genus *Orthotospovirus* (Oliver & Whitfield, 2016). Thrips also transmits diverse plant viruses of different genera including *Tospovirus, Ilarvirus, Carmovirus*, *Sobemovirus* and *Machlomovirus* genera (Jones, 2005b).

In peanut in USA, both adults and larvae are efficient transmitter of *Tomato spotted wilt orthotospovirus* (TSWV), a virus that caused an estimated U.S.$140 million in annual peanut production losses from 1996 to 2006 (LaTora et al., 2022b). White clover (*Trifolium repens*) is an important non-host of *F. fusca* (Xie et al., 2012) that might be a reservoir of different plant viruses including viruses that infect peanuts as well. For example: Peanut stunt virus (Bromoviridae) is found to be infecting peanut beside clover in Georgia and Florida (Baker et al., 1999), Bean yellow mosaic virus (Potyvirus) is found to be infecting peanut beside clover (Hill et. al, 1986). In Georgia TSWV that infects peanut has also been reported in white clover (Barkley et al., 2009). Therefore, the exploring the virus diversity in a non-host crop might have significance in the identification of viruses that might infect the crop and might help to prevent the potential crop loss in future.

The importance of exploring viral diversity in insect vectors is not limited only to the discovery of plant-infecting viruses. There might be a wide diversity of viruses including insect-specific viruses or other unknown viruses. The current study is directed towards understanding the relation of plant viruses and tobacco thrips, however, there is limited research and databases on the viruses that are present in the thrips species beyond the plant viruses. This gap can be filled by the virus metagenomics or meta transcriptomics which uses High throughput sequencing (HTS) to identify, characterize the existing virus in the database and discover new viruses present in selected organism (Thekke-Veetil et al., 2020c). There are more than 800 RNA viruses belonging to 39 families, most of them are from *Iflaviridae, Dicistroviridae, Partitiviridae, and Rhabdoviridae* that are stored in the NCBI database from different agriculturally important insect orders like Hymenoptera, Hemiptera, Thysanoptera, Lepidoptera, Diptera, Coleoptera, and Orthoptera (Qi et al., 2023). Previously in 2018, researchers did the RNA-seq analysis in the pooled sample of soybean thrips (*N. variabilis*) in Midwest USA and discovered 181 viral sequences including the arthopod infecting and mycoviruses (Thekke-Veetil et al., 2020c).

To address the goal of this study, the RNA-seq was done in *F. fusca* collected in two neighboring fields of peanuts and clover in 2022 and 2023 and the viral abundance was compared between crops in both years. Similarly, to identify any similarity in the virus diversity between thrips species, the virus diversity of *F. fusca* was compared to a previous study on the identification of the virus diversity of *N. variabilis* collected in Auburn, Alabama in 2023.

Materials and Methods

Thrips collection

Thrips were collected by sampling from different host plants during May of 2022 and 2023 from the Wiregrass Research and Extension Center in Headland, AL. A collection of 100 thrips was attempted from each host plant (peanut, cotton, and clover) via aspiration. The thrips were kept in glass vials containing 90% ethanol in a -80°C freezer until ready for processing.

RNA extraction and sample preparation

Eighty to eighty-five insects from peanut and clover in 2022 and 2023 were moved from their original container into a sterile 1.7 mL tube containing 250 uL of TriZol and ground thoroughly. An additional 750 uL of TriZol was added while grinding and then incubated for 3 minutes using a rotating mixer. Samples were spun down at 12,000 x g at 4^{o}C for 15 minutes. TriZol was collected while carefully avoiding the insect debris and placed in a new sterile 1.7 mL tube. 200 uL of chloroform was added to the sample and shaken vigorously for 15 seconds before incubating at room temperature for 5 minutes. The samples were then spun down at 12,000 x g at 4^{o}C for 10 minutes. The aqueous phase (top layer) was carefully collected and placed in a new sterile 1.7 mL tube with 500 uL of isopropanol. The samples were then shaken gently by hand and incubated at room temperature for 10 minutes. After incubation, the samples were spun down at 12,000 x g at 4^{o}C for 10 minutes. The supernatant was removed, taking caution not to disturb the small white RNA pellet at the bottom. The pellet was washed with 500 uL of 80% ethanol and spun down at 7,500 x g at 4^{o}C for 5 minutes. All ethanol was removed, and after two minutes at room temperature, the pellet was redissolved in 10 μL of molecular-grade water. Sample concentrations were checked on the Nanodrop and stored in a -20 freezer.

Sequencing information

Total RNA samples were submitted to the North Carolina State Genomic Sciences Laboratory (Raleigh, NC, USA) for Illumina RNA library construction and sequencing. Before library construction, RNA integrity, purity, and concentration were assessed using an Agilent 4200 Tapestation using RNA Tape (Agilent Technologies, USA). Purification of messenger RNA (mRNA) was performed using the oligo-dT beads provided in the NEBNExt Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, USA). Complementary DNA (cDNA) libraries for Illumina sequencing were constructed using the NEB Next Ultra II Directional RNA Library

Prep Kit (NEB) and NEBNext Mulitplex Oligos for Illumina (NEB) using the manufacturer-specified protocol. Then, the mRNA was chemically fragmented and primed with random oligos for first strand cDNA synthesis. Second strand cDNA synthesis was then carried out with dUTPs to preserve strand orientation information. The double-stranded cDNA was then purified, end repaired and "a-tailed" for adaptor ligation. Following ligation, the samples were selected for a final library size (adapters included) of 400-550 bp using sequential AMPure XP bead isolation (Beckman Coulter, USA). Library enrichment was performed, and specific indexes for each sample were added during the protocol-specified PCR amplification. The amplified library fragments were purified and checked for quality and final concentration using an Agilent 4200 Tapestation (DNA Tape, Agilent Technologies, USA). The final quantified libraries were pooled in equimolar amounts for clustering and sequencing on an Illumina NovaSeq 6000 DNA sequencer utilizing an S4 150x2 PE sequencing reagent kit (Illumina, USA). The software package Real Time Analysis (RTA) was used to generate raw bcl, or base call files, which were then de-multiplexed by sample into fastq files for data submission.

Bioinformatic Analysis

The extracted r-RNA was sent for sequencing, and the quality of the sequences was checked using FastQC (Version 0.12.1) (Andrews, S., n.d.). A Phred33 Score was used to measure the base quality of sequences. Any sequences that had less than 28 Phred score was eliminated to ensure >99.9% of the base call accuracy. The poor-quality sequences were trimmed to 1-150 bp to eliminate the poor-quality base pairs using Trim sequences tools (Gordon, 2010).

The mapping of reads was done using Bowtie2 (Langmead et al., 2009b; Langmead & Salzberg, 2012) against the full reference genome of tobacco thrips (GCA029379305.1). The unaligned reads were assembled using Trinity (Grabherr et al., 2011b). These contigs were blasted using

NCBI BLAST+ blastn (Cock et al., 2015b)(Camacho et al., 2009b). The virus annotation was sorted based on the key words “virus”, “viridae”, “phage” and the longest contig (>300 bp) with high amino acid identity to the best match was considered Then, the NCBI ORF Finder was used to predict the start of viral genes that have not been characterized previously. For the viruses whose polymerases (RNA-dependent, RNA polymerase, RdRp) were not discovered before, the conserved domain was predicted using NCBI Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml).

Cross-Validation

The contigs from both analyses were compared to identify the common viruses, This was done by making viral sequence from *F. fusca* collected in clover in 2022, in clover in 2023, in peanut 2022, and in peanut 2023 (Table1) the database by using the command “makeblastdb” followed by blasting contigs of viral sequence from *F. fusca* collected in clover in 2022, in clover in 2023, in peanut 2022, in peanut 2022 (Table 6) using the command “blastn” against NCBI-NT 2023 universal database (Camacho et al., 2009c). Similarly, the combined viral contigs from peanut and clover from both years were used as a database against which the viruses discovered from *N. variabilis* in 2023 were blasted against using “makeblastdb” and “blastn” command respectively (Camacho et al., 2009c).

Phylogenetic tree

To understand the close relatives of the *Polerovirus*, White clover mottle virus isolate fusca1 cds (5733 bp, PV083563) the phylogenetic tree was generated using Molecular Evolutionary Genetics Analysis (MEGA 12) software and Muscle Multiple Sequence alignment (Edgar, 2004). Using MEGA, the Neighbor joining tree was created. The top 50 hits of White clover mottle virus isolate fusca1 (5733 bp, PV083563) from the NCBI were pulled, and the phylogenetic tree was constructed. The accession number used for the construction of phylogenetic trees are PV083563.1, NC_031747.1, MZ218220.1, PP946269.1, JF507725.1, AY956384.1, AY956384.1, NC_008249.1, NC_008249.1, OM514388.1, MK050791.1, NC_055495.1 MH464874.1, ON555767.1, LC269033.1, LC269034.1, LC269040.1, LC269035.1, LC269038.1, LC269032.1, LC269037.1, GU930838.1, GU930837.1, AY956385.1, EU541256.1, EU541259.1, EU541261.1, NC_076825.1, EU541258.1, EU541260.1, EU541262.1, EU541269.1, EU541264.1, EU541270.1, OL999602.1, EU541266.1, EU541255.1, EU541265.1, EU541257.1, EU541263.1, LC709261.1, LC709261.1, KX599154.1, KY364846.1, NC_034246.1, MT739416.1, MT739413.1, MT739418.1, MT739414.1, MT739415.1, MT739417.1, NC_028793.2.

Results

Metagenomic analysis statistics

Although a similar concentration of RNA from the same population of the insect was sent for a similar type of sequencing different number of reads were received. For peanut 2022 (P22) ~60 Million reads, for peanut 2023 (P23) ~44 Million reads, for clover 2022 (C22) ~38 Million Reads, and for clover 2023 (C23) ~27 Million reads were received (Table 1). For P22, P23, C22, and C23 respectively approx. 23.02% of total reads, 15.99 % of the reads, 34.36% of the reads, and 12.87% of the reads did not match to the reference genome tobacco thrips, *Frankliniella fusca* Hinds (GCA029379305.1), and these reads were assembled to generate average 810.24 bp, 423.31 bp, 650.56 bp and 581.03 bp contig length respectively (Table 1).

For the *F. fusca* from P22, P23, C22, and C23 respectively resulted in the 16110 viral contigs, 18559 viral contigs, 16428 viral contigs, and 505 viral contigs (Table 1). The clover 2023 had the lowest read from sequencing, and lower contigs that annotate to the viruses, however, contigs belong to a high number of viruses, compared to other datasets. This suggests that the sequence size, % data filtered out through mapping, the number of contigs generated, and contig length are not related to the diversity of viruses. Therefore, the prevalence of high diversity could be found irrespective of these statistical numbers.

Virus profiling

Viruses in Bunyavirales order are abundant in *F. fusca* sampled from peanut and clover in 2022 and 2023

From the RNA-seq analysis of *F. fusca* sampled from peanut 2022 collected from peanut 2022 in Wiregrass, Alabama, the study discovered eleven different contigs belonging to total six different types of viruses belonging to three different orders like Picornavirales (1), Petitvirales (1), Potential picornavirales (2), Bunyavirales (2) (Table 2, Figure 1, Figure 6). The study discovered three viruses from *F. fusca* collected from peanut in 2023 belonging to the order Bunyavirales (5), Piccovirales (1), Picornavirales (2), Sobelivirales (2), and unclassified orders (2) (Table 3, Figure 2, Figure 7). From *F. fusca* sampled from clover 2022, three viruses belonging to the two different orders, Bunyavirales (2 viruses), and Durnavirales (1 virus) were discovered (Table 4, Figure 3, Figure 8). From *F. fusca* sampled from clover in 2023, twelve viruses belonging to the order Bunyavirales (5), Piccovirales (1), Picornavirales (1), Sobelivirales (1), and unclassified orders (4) (Table 5, Figure 4, Figure 9).

Tomato spotted wilt virus (*Orthotospovirus tomatomaculae*) was common in *F. fusca* from peanut and clover in 2022 and 2023

From the study, the expectation was that there would be shared viruses in the different crops and different years. In *F. fusca* collected from peanut in 2023 and 2022, only *Orthotospovirus tomatomaculae* (PV068193-PV083557, PV068193-PV083558, PV068192-PV083561, PV068192-PV083562) was common (Table 6, Figure 11). Similarly, from clover in 2023 and 2022, *Orthotospovirus tomatomaculae* (PV068188-PV078567) was also common (Table 6, Figure 11). If we compare the crop wide, *F. fusca* collected from clover and peanut in 2022 has two common viruses (Table 6, Figure 11). Similarly, *F. fusca* collected from clover and peanut in 2023 has four different common viruses *Orthotospovirus tomatomaculae* (PV068193-PV068187, PV068192-PV068189, PV068191-PV068188), soybean thrips bunya-like virus (PV068194-PV097222), soybean thrips bunya-like virus 2 (PV068190-PV068190), soybean thrips denso-like virus 3 (PV083565-PV083564).

High abundance of viruses was discovered in *F. fusca* in clover in 2022 and 2023

Similarly, the study expected that the thrips collected from clover would harbor more viruses as it is the alternate host of multiple viruses. Five viruses were commonly discovered between clover in 2022 and 2023, and peanut 2022 and 2023 from *F. fusca* with ten and four unique viruses respectively. In total, clover field harbored 10 viruses whereas peanut harbored 15 viruses (Table 6, Figure 11).

No common viruses to the virus profile of Soybean thrips (Chapter 2)

The study expected that there would be some common viruses between two important thrips species that were sampled from Alabama. However, the comparison result of the viral contig of *F. fusca* (sampled from clover and peanut in 2022 and 2023) did not suggest the presence of any common viruses between soybean thrips, *N. variabilis*, and tobacco thrips, *F. fusca*. Interestingly, the study found that some viruses have similarity to the viruses that were previously discovered from Midwest-USA in 2018 from soybean thrips, *N. variabilis* (Thekke-Veetil et al., 2020c). The viruses whose best match were viruses discovered from Midwest USA study in 2018 in soybean thrips, *N. variabilis* are soybean thrips iflavirus 3 isolate fusca20 polyprotein gene (PV083552) discovered from *F. fusca* collected in peanut in 2022, and soybean thrips bunya-like virus 2 isolate fusca15 (PV068190) discovered from *F. fusca* collected in peanut in 2023, soybean thrips partiti-like virus 8 isolate fusca9 (PV068195) discovered from *F. fusca* collected in clover in 2022, soybean thrips bunya-like virus 2 isolate fusca (PV068190), soybean thrips denso-like virus 3 isolate fusca5 (PV083564), soybean thrips picorna-like virus 8 isolate fusca2 (PV097223), soybean thrips-associated tenui-like virus 2 fusca 3.2 (PV097225), soybean thrips-associated dsRNA virus 2 fusca3.1 (PV097224), soybean thrips-associated dsRNA virus 2 fusca4.1 (PV068185), and soybean thrips-associated dsRNA virus 2 fusca4.2 (PV068186) discovered from *F. fusca* collected in clover in 2023.

Longest contig of White clover mottle virus was found (5733bp, PV083563)

In USA, no other long contig of this virus was found before. The phylogenetic study of White clover mottle virus isolate fusca1 (5733 bp, PV083563) suggests that it was closely related to the White clover mottle virus isolate EFWCMoV genomic sequence (3366 bp, PP946269.1) that was discovered in the leaf hopper (*Empoasca fabae*) in Illinois, USA in 2020 through insect metagenome (Thekke-Veetil et al., 2024) (Table 5). We expected that it would have some similarity to the Cotton leaf roll dwarf virus, as it is also known to infect clover plant. No such finding was observed (Figure 10).

Three Novel viruses were characterized

Fusca virus 1 (2170bp, PV097221)

From the *F. fusca* sampled from clover in 2023, the polyprotein gene of Fusca virus 1 (2170 bp, PV097221) was discovered with 75.3% amino acid identity and an alignment range 6584 bp to 8719 bp to the best match, Victoria bee virus 2 isolate VN2-6 polyprotein gene (8951 bp, MG995724.1). The best match only has a complete gene polyprotein (Best match range 514 bp to 8703 bp). The best match was first identified in the *Apis mellifera* in Victoria, Australia, in April 2014. This is a novel virus, as the amino acid identity to the best match is less than 80%. We have the polyprotein gene, we have 24% of the genome, and we have a missing 74% gene. Based on the best match, the classification of this virus is viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Picornavirales. This virus has not been previously reported from *F. fusca* collected in Wiregrass, Alabama, in clover.

Blattodean nairo-related virus isolate fusca22 (880 bp, PV083556)

From the *F. fusca* sampled from peanut in 2022, Blattodean nairo-related virus isolate fusca22 (880 bp, PV083556) was discovered with 72.931% amino acid identity and an alignment range 976 bp to 4304 bp to the best match Blattodean nairo-related virus OKIAV321 segment M (6063 bp, NC078345.1). The best match was previously found from *Paratemnopteryx couloniana* from Germany in 2011. The best match has 2 genes glycoprotein (328 bp to 3990 bp) and (4373 bp to 5239 bp) (hypothetical protein). It has not been reported in *F. fusca* previously.

Blattodean nairo-related virus isolate fusca 10 segment M (448bp, PV068196)

From the *F. fusca* from clover in 2022, the hypothetical protein 448 bp of Blattodean nairo-related virus isolate fusca 10 segment M **(**PV068196) was found with an alignment range of 4733 bp to 5180 bp and 75.223% amino acid identity to the best match

Blattodean nairo-related virus OKIAV321 segment M genomic sequence (6063 bp, NC078345.1). The best match has two genes hypothetical protein (Best match range: 4373 bp to 5239 bp) and glycoprotein (Best match range: 328 bp to 3990 bp). We have a portion of the hypothetical protein (Best Match range: 4373 bp to 5239 bp). We did not discover glycoprotein gene. The best match was isolated from the host *Paratemnopteryx couloniana* in Germany in 2011. These were the lab culture with samples originating from Germany from the private breeder. Based on the best match the classification for this virus could be: Viruses; Riboviria; Orthornavirae; Negarnaviricota; Polyploviricotina; Ellioviricetes; Bunyavirales; Nairoviridae; Ocetevirus; Ocetevirus paratemnopterygis. This virus is novel and has not been reported in the *F. fusca* before in Alabama, USA, or other thrips spps.

Discussion

Clover is widely used in cover crop or intercrop and is an important alternate host to many viruses (McLaughlin, 1992). This crop is believed to harbor many viruses. So, it is expected that we would discover more viruses from *F. fusca* collected in clover compared to peanut. There is a limited comparative study on virus diversity in *F. fusca* collected from peanut and clover, this study highlights a comprehensive study of virome study in both crop system for two years. For example, studies have found that *F. fusca* from eastern Canada has been found to transmit seven isolates of TSWV (Paliwal, 1974) compared to the natural host like peanut. This suggests that when exposed to a plant host with more viruses, *F. fusca* can retain more viruses. In agroecosystems, wild plants growing alongside crops can act as reservoirs for viruses, which may spread to cultivated plants, leading to epidemics or the emergence of new viruses (Elena et al., 2014; Hasiów-Jaroszewska et al., 2021; Ma et al., 2020; Power & Mitchell, 2004; Roossinck & García-Arenal, 2015).

Because the sampling site of the *F. fusca* in both crops were in proximity, we would expect some common viruses. However, since *F. fusca* has been sampled from two different cropping system, the presence of some unique viruses is also expected. Only one virus is common between *F. fusca* collected from clover in 2023 and 2022, and one virus is common between *F. fusca* collected from peanut in 2023 and 2022. Two viruses are common between *F. fusca* collected from clover in 2022 and peanut in 2022. Four viruses are common between *F. fusca* collected from clover in 2023 and peanut in 2023. *Orthotospovirus tomatomaculae* is a common virus throughout the cropping year and cropping system. Studies have found TSWV in both peanut and clover (Paliwal, 1974; Srinivasan et al., 2017)**.**

However, upon comparison of viruses present in both crops in both years resulted in high abundance of the viruses in *F. fusca* collected from clover compared to peanut. This suggests even for the same insect species, sampling from different cropping systems especially weed hosts might yield in high diversity of the viruses. Therefore, sampling strategy might also play a key factor in determining the high diversity of viruses in an insect metagenome study. From our study, the year of sampling has less significance as more viruses were discovered from *F. fusca* from clover in 2023 compared to clover in 2022 whereas opposite results were obtained in peanut crops.

The expectation is to get a majority of viruses similar to the previous virome study in *N. variabilis* in the Midwestern USA in 2018 (Thekke-Veetil et al., 2020c) because these two insects might feed on the same crop host at some point of their life time (Keough et al., 2018). The ten different viruses had the best match to the viruses that were previously discovered in *N. variabilis* in the Midwestern USA in 2018 (Thekke-Veetil et al., 2020c). The viruses are soybean thrips iflavirus 3 isolate fusca20 polyprotein gene (PV083552) discovered from *F. fusca* collected in peanut in 2022 and soybean thrips bunya-like virus 2 isolate fusca15 (PV068190) discovered from

F. fusca collected in peanut in 2023, soybean thrips partiti-like virus 8 isolate fusca9 (PV068195) discovered from *F. fusca* collected in clover in 2022, soybean thrips bunya-like virus 2 isolate fusca (PV068190), soybean thrips denso-like virus 3 isolate fusca5 (PV083564), soybean thrips picorna-like virus 8 isolate fusca2 (PV097223), soybean thrips-associated tenui-like virus 2 fusca 3.2 (PV097225), soybean thrips-associated dsRNA virus 2 fusca3.1 (PV097224), soybean thrips-associated dsRNA virus 2 fusca4.1 (PV068185), soybean thrips-associated dsRNA virus 2 fusca4.2 (PV068186) discovered from *F. fusca* collected in clover in 2023. This suggests these viruses might have the potential to infect both soybean thrips and tobacco thrips. A study on mosquito virome revealed that specific insect-specific viruses (ISVs) are common across various mosquito species, indicating the presence of a shared viral reservoir within mosquito populations (Gómez et al., 2023). TSWV (*Orthrotospovirus* in Bunyavirales order) is transmitted by different species of thrips like *Thrips tabaci*, *T. setosus*, *F. schultzei*, *F. intonsa*, *F. gemina*, and *F. cephalica* (Riley et al., 2011). So, it is quite possible that the two thrips species soybean thrips and tobacco thrips might share the same pool of viruses while they might have come in contact with each other through feeding (Keough et al., 2018). However, interestingly, the study could not find the common viruses between the *N. variabilis* collected from Auburn, Alabama in 2023 and *F. fusca* collected from Headland, Alabama. This suggests that maybe the common viruses is relevant in these two thrips species at large geographical area.

The study identified three novel viruses. The virus sequence is considered novel based on the % amino acid identity to best match less than 80%. Fusca virus 1 (2170 bp, PV097221) is a novel virus sequence found in *F. fusca* with only 75% amino acid identity match to the Victoria bee virus 2 isolate VN2-6 (8951 bp, MG995724.1) that was reported in Australia in 2014. This virus sequence does not have a closest match that has been discovered in *F. fusca* or any other thrips species previously. Although this virus has a best match to the virus that has been discovered in Honeybee previously, this virus sequence might be only found in *F. fusca* because of <80% amino acid identity. To fully understand the significance of this virus in *F. fusca*, more research on the effect of viruses on the phenotype of the thrips is required.

The other two novel viruses: i. Blattodean nairo-related virus isolate fusca22 hypothetical protein gene, partial cds (880 bp, PV083556) was discovered from *F. fusca* collected from peanut in 2022 with the 72. 931% Amino acid identity to the best match (NC078345.1); ii. Blattodean nairo-related virus isolate fusca 10 segment M **(**PV068196, 448 bp) was found in *F. fusca* sampled from clover in 2022 with 75.223% amino acid identity to the best match (NC078345.1) (Käfer et al., 2019). The best match was isolated from *Paratemnopteryx couloniana* from Germany in 2011. Since both the viral contig have the same annotation, these might be same viruses, however, circulating in the thrips sampled from adjacent clover and peanut field. Although the previous report was on Cockroach species, since the identity to the best match is <80% these might be the novel thrips viruses, closely related to the cockroach viruses.

White clover mottle virus isolate fusca1 (5733 bp, PV083563) was discovered from the study. It is a positive-sense single-stranded RNA virus (+ssRNA) that infects the white clover. There is no literature on the White clover mottle virus isolate. It is believed to be transmitted by aphids as it is a polerovirus however it is surprising to discover this virus in *F. fusca*. This is a new virus that might have potentially might threatened the pasture industry in Alabama in future. Previously, White clover mottle virus isolate EFWCMoV genomic sequence (3366 bp, PP946269.1) was previously discovered in potato hopper (*Empoasca fabae*) in Illinois, USA in 2020 through insect metagenome (Käfer et al., 2019). It had an incomplete P1P2 protein (<1 bp to 1102 bp), readthrough protein (1299 bp to >3366 bp), major capsid protein 1299 bp to 1901 bp), and movement protein (1324 bp to 1905 bp). This was discovered during the High-throughput sequencing analysis of total RNA of potato leaf hopper collected from four Midwestern states in Illinois, Indiana, Iowa, and Minnesota. It has not been previously reported in clover or in the tobacco thrips feeding on the clover in the USA**.** Since tobacco thrips are not a primary vector of this virus, they might have been retained in the insects by feeding on the infected plant. Previously, another relative of this virus has been found through the gut metagenome of the Horse feces in China (MZ218220.1, 4928 bp) (Y.-M. Chen et al., 2022). It is uncommon for the polerovirus to be in thrips species because poleroviruses are phloem-limited and thrips feeds on the the epidermal cells. There needs to be more studies on the economic significance of polerovirus in thrips as there are many poleroviruses that are devastating to economically important agricultural crops.

Even though this is a virus that infects the clover plant, which is a weed host, it might have the potential to be spread to the food crop. Study has suggested that Red clover vein mosaic virus has been widespread in legume (Fletcher et al., 2016). Similarly, in Tasmania the *Polerovirus* genus, in particular Turnip yellows virus (TuYV), has been found to infect pea (another legume) when sampling pea and weeds species around it during 2018-2020 (Umar et al., 2022). This polerovirus P0 and P3 had the genetically diversity to existing TuYV genome, which researchers believe to be because of recombination (Umar et al., 2022). This suggests this *polerovirus* that is currently found to infect clover might also infect the peanut crop threatening the agriculture in Alabama if it recombines with the existing polerovirus Cotton leaf roll dwarf virus in future, which is another important polerovirus found in Alabama.

Bunyavirales is still the predominant virus discovered from RNA-seq analysis: two viruses Blattodean nairo-related virus isolate fusca22 (PV083556), *Orthotospovirus tomatomaculae* (PV083558, PV083559, PV083560, PV083561, PV083562, PV083557)) from *F. fusca* collected from peanut in 2022, two viruses (*orthotospovirus tomatomaculae* (PV068191, PV068192, PV068193, PV068194), soybean thrips bunya-like virus 2 isolate fusca15 (PV068190) from *F. fusca* collected from peanut in 2023, two viruses (*Orthotospovirus tomatomaculae* (PV078565, PV078566, PV078567), Blattodean nairo-related virus isolate fusca 10 segment M (PV068196) from *F. fusca* collected from clover in 2022 and five viruses (soybean thrips bunya-like virus 6 (PV097222), soybean thrips bunya-like virus 2 isolate fusca (PV068190), soybean thrips-associated tenui-like virus 2 fusca 3.2 (PV097225), soybean thrips-associated dsRNA virus 2 fusca3.1 (PV097224), *orthotospovirus tomatomaculae* (PV068187, PV068189, PV068188) from *F. fusca* collected from clover in 2022. Bunyavirales are a group of viruses that can infect both insect and plant species. Unlike many segmented negative-sense RNA viruses that bud at the plasma membrane, most Bunyavirales bud at Golgi membranes, where their envelope glycoproteins (Gn and Gc), undergo processing, trafficking, and interaction with ribonucleoprotein genome segments resulting in reassortment. The reassorted viruses have high pathogenic potential if a cell is co-infected with different bunyaviruses (Barker et al., 2023). Also, the virus might have the potential to be potential crop threat as they can infect both insect and plants. These viruses must be studied in detail as they could be potential crop threat in future.

RNA-dependent RNA polymerase (RdRp) is essential for genome replication and the subsequent transcription process (Venkataraman et al., 2018). The catalytic motifs are the conserved region of RdRp that helps in synthesis of RNA from RNA template in viruses (Jia & Gong, 2019). During the study, only viral RdRp was discovered from the RNA-seq analysis for the two viruses: Soybean thrips partiti-like virus 8 isolate fusca9 (974 bp, PV068195) discovered from *F. fusca* clover 2022 and soybean thrips-associated dsRNA virus 2 fusca4.1 (2876 bp, PV068185) discovered from *F. fusca* in clover in 2023. These two viruses had the conserved catalytic domain. This suggests the RdRp of new viruses might be easy to detect during RNA-seq analysis. Hence, RNA-seq analysis of *F. fusca* collected from weed host and crop host in multiple years resulted in interesting findings and insights on virus discovery by RNA-seq.

Conclusion

Sampling insects from different crops provides greater insights on the diversity of the viruses *F. fusca* acquires in its habitat. Most viruses were discovered from the *F. fusca* from the clover in 2023. The study filters viruses based on the best matches in the NCBI databases, but insect viruses are underrepresented, and novel viruses in *F. fusca* may not be included. This could lead to missing some viruses, so complementary techniques like proteomics should be considered. Biocomputational issues, such as short reads and quality loss during analysis, could cause detection bias, which may be mitigated by using techniques like long-read sequencing, Hi-seq, virus enrichment methods, and incorporating additional approaches to identify DNA viruses potentially integrated into insect genomes.

Reference

Baker, C., Blount, A., & Quesenberry, K. (1999). *Peanut Stunt Virus Infecting Perennial Peanuts in Florida and Georgia*. https://ccmedia.fdacs.gov/content/download/11401/file/pp395.pdf?utm_source=chatgpt.com

Barker, J., daSilva, L. L. P., & Crump, C. M. (2023). Mechanisms of bunyavirus morphogenesis and egress. *Journal of General Virology*, *104*(4). https://doi.org/10.1099/jgv.0.001845

Barkley, N. A., Pinnow, D. L., Wang, M. L., & Pederson, G. A. (2009). First Report of *Tomato spotted wilt virus* Infecting African Clover (*Trifolium tembense*) in Georgia. *Plant Disease*, *93*(2), 202–202. https://doi.org/10.1094/PDIS-93-2-0202C

Batson, J., Dudas, G., Haas-Stapleton, E., Kistler, A. L., Li, L. M., Logan, P., Ratnasiri, K., & Retallack, H. (2021). Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. *eLife*, *10*, e68353. https://doi.org/10.7554/eLife.68353

Beach, A. M. (1897). Contributions to a knowledge of the Thripidae of Iowa. *The Proceedings of the Iowa Academy of Science.*, *3*, 214–227. https://doi.org/10.5962/bhl.part.14877

Brait, N., Hackl, T., Morel, C., Exbrayat, A., Gutierrez, S., & Lequime, S. (2024). A tale of caution: How endogenous viral elements affect virus discovery in transcriptomic data. *Virus Evolution*, *10*(1), vead088. https://doi.org/10.1093/ve/vead088

Brault, V., Mutterer, J., Scheidecker, D., Simonis, M. T., Herrbach, E., Richards, K., & Ziegler-Graff, V. (2000). Effects of Point Mutations in the Readthrough Domain of the Beet

Western Yellows Virus Minor Capsid Protein on Virus Accumulation In Planta and on Transmission by Aphids. *Journal of Virology*, *74*(3), 1140–1148. https://doi.org/10.1128/JVI.74.3.1140-1148.2000

Bruyère, A., Brault, V., Ziegler-Graff, V., Simonis, M.-T., Van Den Heuvel, J. F. J. M., Richards, K., Guilley, H., Jonard, G., & Herrbach, E. (1997). Effects of Mutations in the Beet Western Yellows Virus Readthrough Protein on Its Expression and Packaging and on Virus Accumulation, Symptoms, and Aphid Transmission. *Virology*, *230*(2), 323–334. https://doi.org/10.1006/viro.1997.8476

Burand, J. P., Kim, W., Afonso, C. L., Tulman, E. R., Kutish, G. F., Lu, Z., & Rock, D. L. (2012). Analysis of the genome of the sexually transmitted insect virus Helicoverpa zea nudivirus 2. *Viruses*, *4*(1), 28–61. https://doi.org/10.3390/v4010028

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009a). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009b). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009c). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Card, S. D., Pearson, M. N., & Clover, G. R. G. (2007). Plant pathogens transmitted by pollen. *Australasian Plant Pathology*, *36*(5), 455. https://doi.org/10.1071/AP07050

Chen, A. Y. S., Walker, G. P., Carter, D., & Ng, J. C. K. (2011). A virus capsid component mediates virion retention and transmission by its insect vector. *Proceedings of the National Academy of Sciences*, *108*(40), 16777–16782. https://doi.org/10.1073/pnas.1109384108

Chen, Y., Evans, J., & Feldlaufer, M. (2006). Horizontal and vertical transmission of viruses in the honey bee, Apis mellifera. *Journal of Invertebrate Pathology*, *92*(3), 152–159. https://doi.org/10.1016/j.jip.2006.03.010

Chen, Y.-M., Sadiq, S., Tian, J.-H., Chen, X., Lin, X.-D., Shen, J.-J., Chen, H., Hao, Z.-Y., Wille, M., Zhou, Z.-C., Wu, J., Li, F., Wang, H.-W., Yang, W.-D., Xu, Q.-Y., Wang, W., Gao, W.-H., Holmes, E. C., & Zhang, Y.-Z. (2022). RNA viromes from terrestrial sites across China expand environmental viral diversity. *Nature Microbiology*, *7*(8), 1312–1323. https://doi.org/10.1038/s41564-022-01180-2

Cobbin, J. C., Charon, J., Harvey, E., Holmes, E. C., & Mahar, J. E. (2021). Current challenges to virus discovery by meta-transcriptomics. *Current Opinion in Virology*, *51*, 48–55. https://doi.org/10.1016/j.coviro.2021.09.007

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015a). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015b). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Common types of viruses of floriculture crops and their modes of transmission. (2014, April 3). MSU Extension.

https://www.canr.msu.edu/news/common_types_of_viruses_of_floriculture_crops_and_their_modes_of_transmissi

Conner, K., Sikora, E. J., Zhang, L., & Burmester, C. (2013). First Report of *Soybean vein necrosis-associated virus* Affecting Soybeans in Alabama. *Plant Health Progress*, *14*(1), 49. https://doi.org/10.1094/PHP-2013-0729-03-BR

Culbreath, A. K., Todd, J. W., & Brown, S. L. (2003). E PIDEMIOLOGY AND M ANAGEMENT OF T OMATO S POTTED W ILT IN P EANUT. *Annual Review of Phytopathology*, *41*(1), 53–75. https://doi.org/10.1146/annurev.phyto.41.052002.095522

Dietzgen, R., Mann, K., & Johnson, K. (2016). Plant Virus–Insect Vector Interactions: Current and Potential Future Research Directions. *Viruses*, *8*(11), 303. https://doi.org/10.3390/v8110303

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Elena, S. F., Fraile, A., & García-Arenal, F. (2014). Evolution and Emergence of Plant Viruses. In *Advances in Virus Research* (Vol. 88, pp. 161–191). Elsevier. https://doi.org/10.1016/B978-0-12-800098-4.00003-9

Feng, Y., Krueger, E. N., Liu, S., Dorman, K., Bonning, B. C., & Miller, W. A. (2017). Discovery of Known and Novel Viral Genomes in Soybean Aphid by Deep Sequencing. *Phytobiomes Journal*, *1*(1), 36–45. https://doi.org/10.1094/PBIOMES-11-16-0013-R

Fletcher, J., Tang, J., Blouin, A., Ward, L., MacDiarmid, R., & Ziebell, H. (2016). *Red clover vein mosaic virus*—A Novel Virus to New Zealand that is Widespread in Legumes. *Plant Disease*, *100*(5), 890–895. https://doi.org/10.1094/PDIS-04-15-0465-RE

Gómez, M., Martínez, D., Páez-Triana, L., Luna, N., De Las Salas, J. L., Hernández, C., Flórez, A. Z., Muñoz, M., & Ramírez, J. D. (2023). Characterizing viral species in mosquitoes (Culicidae) in the Colombian Orinoco: Insights from a preliminary metagenomic study. *Scientific Reports*, *13*(1), 22081. https://doi.org/10.1038/s41598-023-49232-9

Gordon, A. (2010). *FASTQ/A short reads pre-processing tools.* [Computer software]. http://hannonlab.cshl.edu/fastx_toolkit/

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011a). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011b). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Hajimorad, M. R., Domier, L. L., Tolin, S. A., Whitham, S. A., & Saghai Maroof, M. A. (2018). *Soybean mosaic virus*: A successful potyvirus with a wide distribution but restricted natural

host range. *Molecular Plant Pathology*, *19*(7), 1563–1579. https://doi.org/10.1111/mpp.12644

Harrewijn, P., Tjallingii, W. F., & Mollema, C. (1996). Electrical recording of plant penetration by western flower thrips. *Entomologia Experimentalis et Applicata*, *79*(3), 345–353. https://doi.org/10.1111/j.1570-7458.1996.tb00842.x

Hasiów-Jaroszewska, B., Boezen, D., & Zwart, M. P. (2021). Metagenomic Studies of Viruses in Weeds and Wild Plants: A Powerful Approach to Characterise Variable Virus Communities. *Viruses*, *13*(10), 1939. https://doi.org/10.3390/v13101939

History and Economic Impact. (2025, March 14). https://tswv.caes.uga.edu/usda-ramp-project/history-and-economic-impact.html?utm_source=chatgpt.com

Hogenhout, S. A., Ammar, E.-D., Whitfield, A. E., & Redinbaugh, M. G. (2008). Insect Vector Interactions with Persistently Transmitted Viruses. *Annual Review of Phytopathology*, *46*(1), 327–359. https://doi.org/10.1146/annurev.phyto.022508.092135

Hunter, W. B., & Ullman, D. E. (1989). Analysis of mouthpart movements during feeding of Frankliniella occidentalis (pergande) and F. schultzei trybom (Thysanoptera: Thripidae). *International Journal of Insect Morphology and Embryology*, *18*(2–3), 161–171. https://doi.org/10.1016/0020-7322(89)90024-X

Hýsková, V., Bělonožníková, K., Chmelík, J., Hoffmeisterová, H., Čeřovská, N., Moravec, T., & Ryšlavá, H. (2024). Potyviral Helper-Component Protease: Multifaced Functions and Interactions with Host Proteins. *Plants*, *13*(9), 1236. https://doi.org/10.3390/plants13091236

Jia, H., & Gong, P. (2019). A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. *Frontiers in Microbiology*, *10*, 1945. https://doi.org/10.3389/fmicb.2019.01945

Jones, D. R. (2005). *Plant viruses transmitted by thrips*. *113*, 119–157. https://doi.org/10.1007/s10658-005-2334-1

Junglen, S., & Drosten, C. (2013). Virus discovery and recent insights into virus diversity in arthropods. *Current Opinion in Microbiology*, *16*(4), 507–513. https://doi.org/10.1016/j.mib.2013.06.005

Käfer, S., Paraskevopoulou, S., Zirkel, F., Wieseke, N., Donath, A., Petersen, M., Jones, T. C., Liu, S., Zhou, X., Middendorf, M., Junglen, S., Misof, B., & Drosten, C. (2019). Re-assessing the diversity of negative strand RNA viruses in insects. *PLOS Pathogens*, *15*(12), e1008224. https://doi.org/10.1371/journal.ppat.1008224

Keough, S., Danielson, J., Marshall, J. M., Lagos-Kutz, D., Voegtlin, D. J., Srinivasan, R., & Nachappa, P. (2018). Factors Affecting Population Dynamics of Thrips Vectors of Soybean vein necrosis virus. *Environmental Entomology*, *47*(3), 734–740. https://doi.org/10.1093/ee/nvy021

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009a). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009b). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Lanzi, G., de Miranda, J. R., Boniotti, M. B., Cameron, C. E., Lavazza, A., Capucci, L., Camazine, S. M., & Rossi, C. (2006). Molecular and biological characterization of deformed wing virus of honeybees (Apis mellifera L.). *Journal of Virology*, *80*(10), 4998–5009. https://doi.org/10.1128/JVI.80.10.4998-5009.2006

LaTora, A. G., Lai, P.-C., Chen, Y.-J., Gautam, S., Abney, M. R., & Srinivasan, R. (2022a). *Frankliniella fusca* (Thysanoptera: Thripidae), The Vector of *Tomato Spotted Wilt Orthotospovirus* Infecting Peanut in the Southeastern United States. *Journal of Integrated Pest Management*, *13*(1), 3. https://doi.org/10.1093/jipm/pmab045

LaTora, A. G., Lai, P.-C., Chen, Y.-J., Gautam, S., Abney, M. R., & Srinivasan, R. (2022b). *Frankliniella fusca* (Thysanoptera: Thripidae), The Vector of *Tomato Spotted Wilt Orthotospovirus* Infecting Peanut in the Southeastern United States. *Journal of Integrated Pest Management*, *13*(1), 3. https://doi.org/10.1093/jipm/pmab045

Lawrence, P., Heung, M., Nave, J., Henkel, C., & Escudero-Pérez, B. (2023). The natural virome and pandemic potential: Disease X. *Current Opinion in Virology*, *63*, 101377. https://doi.org/10.1016/j.coviro.2023.101377

Liu, S., Vijayendr, D., & Bonning, B. C. (2011). *Next Generation Sequencing Technologies for Insect Virus Discovery*. *3*, 1849–1869. https://doi.org/10.3390/v3101849

Looi, L. M., & Chua, K. B. (2007). Lessons from the Nipah virus outbreak in Malaysia. *Malaysian Journal of Pathology*, *29*(2), 63–67.

Ma, Y., Marais, A., Lefebvre, M., Faure, C., & Candresse, T. (2020). Metagenomic analysis of virome cross-talk between cultivated Solanum lycopersicum and wild Solanum nigrum. *Virology*, *540*, 38–44. https://doi.org/10.1016/j.virol.2019.11.009

Marklewitz, M., Zirkel, F., Kurth, A., Drosten, C., & Junglen, S. (2015). Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. *Proceedings of the National Academy of Sciences*, *112*(24), 7536–7541. https://doi.org/10.1073/pnas.1502036112

Marklewitz, M., Zirkel, F., Rwego, I. B., Heidemann, H., Trippner, P., Kurth, A., Kallies, R., Briese, T., Lipkin, W. I., Drosten, C., Gillespie, T. R., & Junglen, S. (2013). Discovery of a Unique Novel Clade of Mosquito-Associated Bunyaviruses. *Journal of Virology*, *87*(23), 12850–12865. https://doi.org/10.1128/JVI.01862-13

McLaughlin, M. R. (1992). Virus Diseases and Stand Decline in a White Clover Pasture. *Plant Disease*, *76*(2), 158. https://doi.org/10.1094/PD-76-0158

Mokili, J. L., Rohwer, F., & Dutilh, B. E. (2012). Metagenomics and future perspectives in virus discovery. *Current Opinion in Virology*, *2*(1), 63–77. https://doi.org/10.1016/j.coviro.2011.12.004

Nault, B. A., & Speese Iii, J. (2002). Major insect pests and economics of fresh-market tomato in eastern Virginia. *Crop Protection*, *21*(5), 359–366. https://doi.org/10.1016/S0261-2194(01)00103-X

Ng, T. F. F., Duffy, S., Polston, J. E., Bixby, E., Vallad, G. E., & Breitbart, M. (2011). Exploring the Diversity of Plant DNA Viruses and Their Satellites Using Vector-Enabled

Metagenomics on Whiteflies. *PLoS ONE*, *6*(4), e19050. https://doi.org/10.1371/journal.pone.0019050

Omkar. (2021). *Polyphagous Pests of Crops*. Springer Singapore Pte. Limited.

Paliwal, Y. C. (1974). Some properties and thrip transmission of tomato spotted wilt virus in Canada. *Canadian Journal of Botany*, *52*(6), 1177–1182. https://doi.org/10.1139/b74-151

Pan, S. (2024, September 25). Insects as Vectors – Features, Types, Examples. *Insects as Vectors – Features, Types, Examples*. https://biologynotesonline.com/insects-as-vectors-features-types-examples/

Peter, K. A., Liang, D., Palukaitis, P., & Gray, S. M. (2008). Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. *Journal of General Virology*, *89*(8), 2037–2045. https://doi.org/10.1099/vir.0.83625-0

Pirone, T. P., & Megahed, E.-S. (1966). Aphid transmissibility of some purified viruses and viral RNA's. *Virology*, *30*(4), 631–637. https://doi.org/10.1016/0042-6822(66)90168-1

Power, A. G., & Mitchell, C. E. (2004). Pathogen Spillover in Disease Epidemics. *The American Naturalist*, *164*(S5), S79–S89. https://doi.org/10.1086/424610

Qi, Y.-H., Ye, Z.-X., Zhang, C.-X., Chen, J.-P., & Li, J.-M. (2023). Diversity of RNA viruses in agricultural insects. *Computational and Structural Biotechnology Journal*, *21*, 4312–4321. https://doi.org/10.1016/j.csbj.2023.08.036

Riley, D. G., Joseph, S. V., Srinivasan, R., & Diffie, S. (2011). Thrips Vectors of Tospoviruses. *Journal of Integrated Pest Management*, *2*(1), I1–I10. https://doi.org/10.1603/IPM10020

Roossinck, M. J., & García-Arenal, F. (2015). Ecosystem simplification, biodiversity loss and plant virus emergence. *Current Opinion in Virology*, *10*, 56–62. https://doi.org/10.1016/j.coviro.2015.01.005

Rosario, K., Capobianco, H., Ng, T. F. F., Breitbart, M., & Polston, J. E. (2014). RNA Viral Metagenome of Whiteflies Leads to the Discovery and Characterization of a Whitefly-Transmitted Carlavirus in North America. *PLoS ONE*, *9*(1), e86748. https://doi.org/10.1371/journal.pone.0086748

Rosario, K., Seah, Y., Marr, C., Varsani, A., Kraberger, S., Stainton, D., Moriones, E., Polston, J., Duffy, S., & Breitbart, M. (2015). Vector-Enabled Metagenomic (VEM) Surveys Using Whiteflies (Aleyrodidae) Reveal Novel Begomovirus Species in the New and OldWorlds. *Viruses*, *7*(10), 5553–5570. https://doi.org/10.3390/v7102895

Rozo-Lopez, P., Brewer, W., Käfer, S., Martin, M. M., & Parker, B. J. (2023). Untangling an insect's virome from its endogenous viral elements. *BMC Genomics*, *24*(1), 636. https://doi.org/10.1186/s12864-023-09737-z

Soybean Thrips. (n.d.). Purdue University - College of Agriculture. Retrieved March 14, 2025, from https://ag.purdue.edu/department/entm/extension/field-crops-ipm/soybean/soybean-thrips.html

Soybean Vein Necrosis Virus—Soybean Disease—Soybean Research & Information Network—SRIN. (n.d.). Soybean Research & Information Network. Retrieved March 29, 2025, from https://soybeanresearchinfo.com/soybean-disease/soybean-vein-necrosis-virus/

Srinivasan, R., Abney, M. R., Culbreath, A. K., Kemerait, R. C., Tubbs, R. S., Monfort, W. S., & Pappu, H. R. (2017). Three decades of managing Tomato spotted wilt virus in peanut in

southeastern United States. *Virus Research*, *241*, 203–212. https://doi.org/10.1016/j.virusres.2017.05.016

Stafford, C. A., Walker, G. P., & Ullman, D. E. (2012). Hitching a ride: Vector feeding and virus transmission. *Communicative & Integrative Biology*, *5*(1), 43–49. https://doi.org/10.4161/cib.18640

Stewart, L. R., Medina, V., Tian, T., Turina, M., Falk, B. W., & Ng, J. C. K. (2010). A Mutation in the *Lettuce Infectious Yellows Virus* Minor Coat Protein Disrupts Whitefly Transmission but Not *In Planta* Systemic Movement. *Journal of Virology*, *84*(23), 12165–12173. https://doi.org/10.1128/JVI.01192-10

The PLOS Pathogens Staff. (2014). Correction: Transovarial Transmission of a Plant Virus Is Mediated by Vitellogenin of Its Insect Vector. *PLoS Pathogens*, *10*(4), e1004141. https://doi.org/10.1371/journal.ppat.1004141

Thekke-Veetil, T., Lagos-Kutz, D., Domier, L. L., McCoppin, N. K., Hartman, G. L., & Clough, S. J. (2024). Exploring Virus Diversity in the Potato leafhopper (Empoasca fabae), an Economically Important Agricultural Pest. *Viruses*, *16*(8), 1305. https://doi.org/10.3390/v16081305

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020a). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020b). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly

Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thekke-Veetil, T., Lagos-Kutz, D., McCoppin, N. K., Hartman, G. L., Ju, H.-K., Lim, H.-S., & Domier, Leslie. L. (2020c). Soybean Thrips (Thysanoptera: Thripidae) Harbor Highly Diverse Populations of Arthropod, Fungal and Plant Viruses. *Viruses*, *12*(12), 1376. https://doi.org/10.3390/v12121376

Thrips Management Guidelines—UC IPM. (n.d.). Retrieved April 6, 2025, from https://ipm.ucanr.edu/PMG/PESTNOTES/pn7429.html?utm_source=chatgpt.com

Tobacco thrips. (2025, February 27). *Tobacco Thrips*. https://www.biobee.com/pests/tobacco-thrips/

Umar, M., Tegg, R. S., Farooq, T., Thangavel, T., & Wilson, C. R. (2022). Abundance of Poleroviruses within Tasmanian Pea Crops and Surrounding Weeds, and the Genetic Diversity of TuYV Isolates Found. *Viruses, 14*(8), 1690. https://doi.org/10.3390/v14081690

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., Guzman, H., Wood, T. G., Popov, V., Gorchakov, R., González, A. V., Haddow, A. D., Watts, D. M., Da Rosa, A. P. A. T., Weaver, S. C., Lipkin, W. I., & Tesh, R. B. (2013). Negevirus: A Proposed New Taxon of Insect-Specific Viruses with Wide Geographic Distribution. *Journal of Virology*, *87*(5), 2475–2488. https://doi.org/10.1128/JVI.00776-12

Vasilakis, N., Guzman, H., Firth, C., Forrester, N. L., Widen, S. G., Wood, T. G., Rossi, S. L., Ghedin, E., Popov, V., Blasdell, K. R., Walker, P. J., & Tesh, R. B. (2014). Mesoniviruses

are mosquito-specific viruses with extensive geographic distribution and host range. *Virology Journal*, *11*(1), 97. https://doi.org/10.1186/1743-422X-11-97

Venkataraman, S., Prasad, B., & Selvarajan, R. (2018). RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. *Viruses*, *10*(2), 76. https://doi.org/10.3390/v10020076

Wamonje, F. O., Michuki, G. N., Braidwood, L. A., Njuguna, J. N., Musembi Mutuku, J., Djikeng, A., Harvey, J. J. W., & Carr, J. P. (2017). Viral metagenomics of aphids present in bean and maize plots on mixed-use farms in Kenya reveals the presence of three dicistroviruses including a novel Big Sioux River virus-like dicistrovirus. *Virology Journal*, *14*(1), 188. https://doi.org/10.1186/s12985-017-0854-x

Wang, Z., Gerstein, M., & Snyder, M. (2009). *RNA-Seq: A revolutionary tool for transcriptomics*. *10*, 57–63.

Wendimu, G. Y., & Gurmu, A. K. (2024). Insect Vectors of Plant Viruses: Host Interactions, Their Effects, and Future Opportunities. *2024*, *2024*(1). https://doi.org/10.1155/aia/6006985

Whitfield, A. E., Falk, B. W., & Rotenberg, D. (2015a). Insect vector-mediated transmission of plant viruses. *Virology*, *479–480*, 278–289. https://doi.org/10.1016/j.virol.2015.03.026

Whitfield, A. E., Falk, B. W., & Rotenberg, D. (2015b). *Insect vector-mediated transmission of plant viruses*. *479*, 278–289.

Wu, H., Pang, R., Cheng, T., Xue, L., Zeng, H., Lei, T., Chen, M., Wu, S., Ding, Y., Zhang, J., Shi, M., & Wu, Q. (2020). Abundant and Diverse RNA Viruses in Insects Revealed by RNA-Seq Analysis: Ecological and Evolutionary Implications. *mSystems*, *5*(4), e00039-20. https://doi.org/10.1128/mSystems.00039-20

Xie, Y.-H., Li, Z.-Y., Dong, K., & Zhang, H.-R. (2012). Changes in the species composition of thrips on *Trifolium repens* (Fabales). *Acta Phytopathologica et Entomologica Hungarica*, *47*(1), 61–67. https://doi.org/10.1556/APhyt.47.2012.1.7

Zhou, J., Kantartzi, S. K., Wen, R.-H., Newman, M., Hajimorad, M. R., Rupe, J. C., & Tzanetakis, I. E. (2011). Molecular characterization of a new tospovirus infecting soybean. *Virus Genes*, *43*(2), 289–295. https://doi.org/10.1007/s11262-011-0621-9

Zhou, J., & Tzanetakis, I. E. (2013). Epidemiology of Soybean vein necrosis-associated virus. *Phytopathology®*, *103*(9), 966–971. https://doi.org/10.1094/PHYTO-12-12-0322-R

Tables

Table 1. Sequences statistics of RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut and clover in 2022 and 2023

	Peanut 2022	Peanut 2023	Clover 2022	Clover 2023
Number of total reads	120400496	896906028	54207354	773737392
Total number of forward and reverse reads	60200248	448453014	27103677	386868696
Length range of the read before trimming (bp)	251	151	251	151
Number of bases before trimming (Gbp)	15.1	67.7	6.8	58.4
GC% for forward reads before trimming	58	52	57	52
GC% for reverse reads before trimming	57	52	55	52
Length range of the read after trimming (bp)	150	NA	150	NA
Number of bases after trimming (Gbp)	9	NA	4.3	NA
GC% after trimming	56	NA	54	NA
Sequences matched to the reference genome %	76.98	84.01	65.64	87.13
Sequences unmatched to the reference genome %	23.02	15.99	34.36	12.87
Number of trinity transcripts	48110	170138	39363	162833

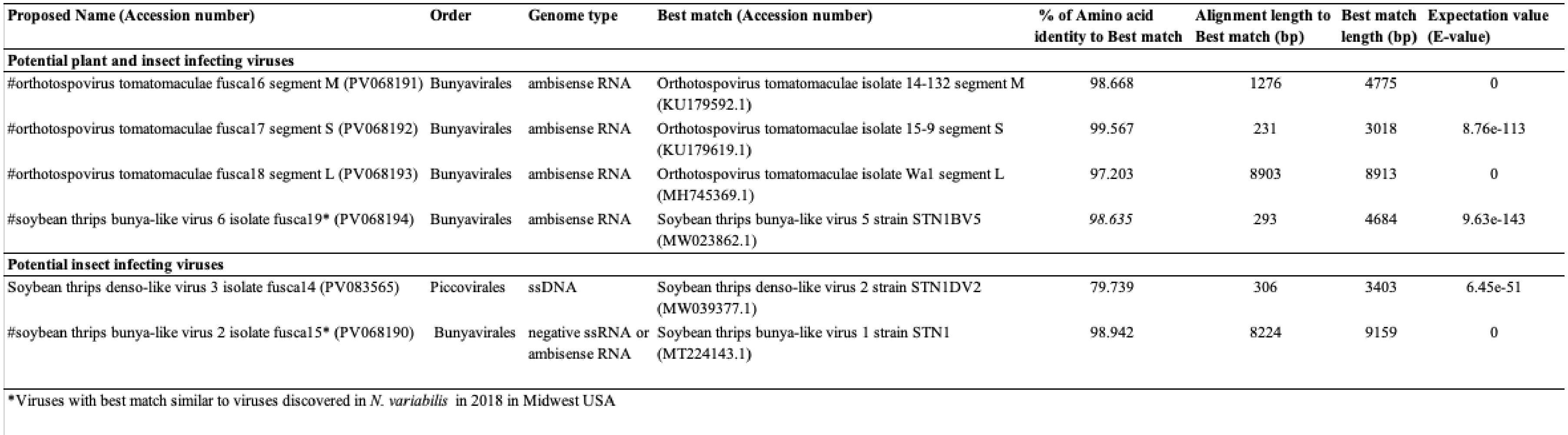
GC% of Trinity transcripts	46.12	44.62	44.75	45.47
Average contig length (bp)	810.24	423.31	689.05	581.03
Average assembled base for the contig	38980730	72021269	27123036	94611092
Contig N10	3051	1786	2205	3516
Contig N20	2168	1059	1565	2048
Contig N30	1681	698	1206	1167
Contig N40	1320	522	956	676
Contig N50	1045	425	773	479
Number of viral contigs	16110	18559	505	16428

Table 2. Diversity of viruses in Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut in 2022 through RNA-seq analysis

Proposed name (Accession number)	**Orders**	**Genome type**	**Best match (Accession number)**	**% of Amino acid identity to Best match**	**Best match length (bp)**	**Alignment length to Best match (bp)**	**Expectation value (E-value)**
Potential plant and insect infecting viruses							
Orthotospovirus tomatomaculae isolate fusca24 RNA-dependent RNA polymerase gene (PV083558)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate LN-HJL segment L (MT241883.1)	97.015	8914	536	0
Orthotospovirus tomatomaculae isolate fusca25 (PV083559)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 15-122 segment M (KU179620.1)	98.743	4772	557	0
Orthotospovirus tomatomaculae isolate fusca26(PV083560)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 15-9 segment S (KU179619.1)	99.153	3018	1062	0
Orthotospovirus tomatomaculae isolate fusca27 (PV083561)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 15-122 segment S (KU179621.1)	99.495	3021	792	0
Orthotospovirus tomatomaculae isolate fusca28 (PV083562)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 14-132 segment S (KU179593.1)	99.085	3013	765	0
Orthotospovirus tomatomaculae isolate fusca23 RNA-dependent RNA polymerase gene (PV083557)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate TifNVHOL-19-Tif RNA-dependent RNA polymerase (RdRp) (MW519468.1)	99.457	1416	552	0
Potential insect infecting viruses							
Soybean thrips iflavirus 3 isolate fusca20 polyprotein gene (PV083552)*	Picornavirales	positive ssRNA	Soybean thrips iflavirus 2 strain STN1 (MT195547.1)	71.744	10286	1766	1.29E-123
Soybean thrips iflavirus 1 hypothetical protein 1 gene (PV083553)	NA	NA	NA				
Soybean thrips iflavirus 1 hypothetical protein 2 gene (PV083554)	NA	NA	NA				
Sinsheimervirus sp. isolate fusca21(PV083555)	Petitvirales	ssDNA	Sinsheimervirus phiX174 strain (MH378442.1)	98.549	5386	758	0
Blattodean nairo-related virus isolate fusca22 (PV083556)	Bunyavirales	negative ssRNA or ambisense RNA	Blattodean nairo-related virus OKIAV321 segment M (NC_078345.1)	72.931	6063	894	2.41e-70

*Viruses with best match similar to viruses discovered in *N. variabilis* in 2018 in Midwest USA

Table 3. Diversity of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut in 2023 through RNA-seq analysis



Proposed Name (Accession number)	**Order**	**Genome type**	**Best match (Accession number)**	**% of Amino acid identity to Best match**	**Alignment length to Best match (bp)**	**Best match length (bp)**	**Expectation value (E-value)**
Potential plant and insect infecting viruses							
#orthotospovirus tomatomaculae fusca16 segment M (PV068191)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 14-132 segment M (KU179592.1)	98.668	1276	4775	0
#orthotospovirus tomatomaculae fusca17 segment S (PV068192)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 15-9 segment S (KU179619.1)	99.567	231	3018	8.76e-113
#orthotospovirus tomatomaculae fusca18 segment L (PV068193)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate Wa1 segment L (MH745369.1)	97.203	8903	8913	0
#soybean thrips bunya-like virus 6 isolate fusca19* (PV068194)	Bunyavirales	ambisense RNA	Soybean thrips bunya-like virus 5 strain STN1BV5 (MW023862.1)	*98.635*	293	4684	9.63e-143
Potential insect infecting viruses							
Soybean thrips denso-like virus 3 isolate fusca14 (PV083565)	Piccovirales	ssDNA	Soybean thrips denso-like virus 2 strain STN1DV2 (MW039377.1)	79.739	306	3403	6.45e-51
#soybean thrips bunya-like virus 2 isolate fusca15* (PV068190)	Bunyavirales	negative ssRNA or ambisense RNA	Soybean thrips bunya-like virus 1 strain STN1 (MT224143.1)	98.942	8224	9159	0

*Viruses with best match similar to viruses discovered in *N. variabilis* in 2018 in Midwest USA

Table 4. Diversity of viruses of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in clover in 2022 through RNA-seq analysis

Proposed Name (Accession number)	**Order**	**Genotype**	**Best match (Accession number)**	**Best match length (bp)**	**% of Amino acid identity to Best match**	**Alignment length to Best match (bp)**	**Expectation value (E-value)**
Potential plant and insect infecting viruses							
Orthotospovirus tomatomaculae isolate fusca11 segment S (PV078565)	Bunyavirales	ambisense	Orthotospovirus tomatomaculae isolate TUFRunner297-19-Tif (MW519234.1)	777	99.495	396	0
Orthotospovirus tomatomaculae isolate fusca12 segment L (PV078566)	Bunyavirales	ambisense	Orthotospovirus tomatomaculae isolate PepCal_12 segment L (MK348941.1)	8909	97.772	404	0
Orthotospovirus tomatomaculae isolate fusca13 segment M (PV078567)	Bunyavirales	ambisense	Orthotospovirus tomatomaculae isolate 14-132 segment M (KU179592.1)	4775	99.291	423	0
Potential insect infecting viruses							
Soybean thrips partiti-like virus 8 isolate fusca9 (PV068195)	Durnavirales	dsRNA	Soybean thrips partiti-like virus 7 isolate STN1PV7 (MT648427.1)	1429	88.706	974	0
Blattodean nairo-related virus isolate fusca 10 segment M (PV068196)	Bunyavirales	negative ssRNA or ambisense	Blattodean nairo-related virus OKIAV321 segment M (NC_078345.1)	6063	75.223	448	3.44e-50

Table 5. Diversity of viruses of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama, in clover in 2023 through RNA-seq analysis

Proposed Name (Accession number)	Order	Genome type	Best match (Accession number)	Best match length (bp)	% of Amino acid identity to Best Match	Alignment length to best match (bp)	Expectation value (E-value)
Potential plant infecting viruses							
White clover mottle virus isolate fusca1 (PV083563)	Sobelivirales	positive ssRNA	White clover mottle virus genomic RNA (NC_031747.1)	6205	98.273	5734	0
Cotton leafroll dwarf virus isolate wiregrass Bankit2933961	Sobelivirales	positive ssRNA	Cotton leafroll dwarf virus isolate SC_Barnwell5 (OQ300129.1)	4265	98.773	652	
Potential plant and insect infecting viruses							
orthotospovirus tomatomaculae isolate fusca5, L segment (PV068187)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate Wa1 segment L (MH745369.1)	8913	97.249	7198	0
orthotospovirus tomatomaculae isolate fusca6, S segment (PV068189)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 15-122 segment S (KU179621.1)	3021	98.842	1727	0
orthotospovirus tomatomaculae isolate fusca7, M segment (PV068188)	Bunyavirales	ambisense RNA	Orthotospovirus tomatomaculae isolate 14-198 segment M (KU179600.1)	4771	98.869	4774	0
Potential insect infecting viruses							
Fusca virus 1 (PV097221)	Picornavirales	positive ssRNA	Victoria bee virus 2 isolate VN2-6 (MG995724.1)	8951	75.3	2170	0
Soybean thrips bunya-like virus 6* (PV097222)	Bunyavirales	negative ssRNA or ambisense	Soybean thrips bunya-like virus 5 strain STN1BV5 (MW023862.1)	4684	99.08	4673	0
Soybean thrips bunya-like virus 2 isolate fusca* (PV068190)	Bunyavirales	negative ssRNA or ambisense	Soybean thrips bunya-like virus 1 strain STN1 (MT224143.1)	9159	98.689	9152	0
Soybean thrips denso-like virus 3 isolate fusca5* (PV083564)	Piccovirales	ssDNA	Soybean thrips denso-like virus 2 strain STN1DV2 (MW039377.1)	3403	99.295	1560	0
Soybean thrips picorna-like virus 8 isolate fusca2* (PV097223)	Picornavirales	positive ssRNA	Soybean thrips picorna-like virus 7 strain STN1 (MT293128.1)	8960	98.645	3764	0
Soybean thrips-associated tenui-like virus 2 fusca 3.2* (PV097225)	Bunyavirales	negative ssRNA or ambisense	Soybean thrips-associated tenui-like virus 1 strain STN1TV1 (MW033650.1)	1815	99.198	1122	0
Soybean thrips-associated dsRNA virus 2 fusca3.1* (PV097224)	Bunyavirales	negative ssRNA or ambisense	NA			668	
Soybean thrips-associated dsRNA virus 2 fusca4.1* (PV068185)	NA	NA	Soybean thrips-associated dsRNA virus-1 strain STN1 (MT293146.1)	3695	99.444	2876	0
Soybean thrips-associated dsRNA virus 2 fusca4.2* (PV068186)	NA	NA	Soybean thrips-associated dsRNA virus 4 strain STN1DRV4 (MW033658.1)	3037	99.86	738	0

*Viruses with best match similar to viruses discovered in *N. variabilis i*n 2018 in Midwest USA

Table 6. Comparison of all the viruses discovered from both clover and peanut crop in 2022 and 2023

Database viruses	Database viruses length (bp)	Virus that were blasted against the database viruses	Length(bp)	%identity of nt	Alignment length to the database	Alignment %	Potential virus	Order	E-value
Peanut field 2023		**Peanut field 2022**							
orthotospovirus tomatomaculae fusca18 segment L (PV068193)	8903	orthotospovirus tomatomaculae fusca23 segment L (PV083557)	552	98.732	552	100	Plant and insect	Bunyavirales	0
orthotospovirus tomatomaculae fusca18 segment L (PV068193)	8903	orthotospovirus tomatomaculae fusca24 segment L (PV083558)	536	97.015	536	100	Plant and insect	Bunyavirales	0
orthotospovirus tomatomaculae fusca17 segment S (PV068192)	231	orthotospovirus tomatomaculae fusca27 segment S (PV083561)	792	97.297	37	4.67171717	Plant and insect	Bunyavirales	3.37E-12
orthotospovirus tomatomaculae fusca17 segment S (PV068192)	231	orthotospovirus tomatomaculae fusca28 segment S (PV083562)	765	100	180	23.5294118	Plant and insect	Bunyavirales	5.75E-94
Clover field in 2023		**Clover field in 2022**							
orthotospovirus tomatomaculae isolate fusca7, M segment (PV068188)	4774	orthotospovirus tomatomaculae isolate fusca13 segment M (PV078567)	423	99.054	423	100	Plant and insect	Bunyavirales	0
Clover field 2022		**Peanut field 2022**							
Blattodean nairo-related virus isolate fusc	448	blattodean nairo-related virus fusca22	894	100	448	50.1118568	Insect	Bunyavirales	0
orthotospovirus tomatomaculae isolate fusca11 segment S PV078565	396	orthotospovirus tomatomaculae fusca26 segment S (PV083560)	1062	98.737	396	37.2881356	Plant and insect	Bunyavirales	0
Peanut field in 2023		**Clover field in 2023**							
soybean thrips bunya-like virus 6 isolate fusca19 PV068194	293	soybean thrips bunya-like virus 6 (PV097222)	4673	98.976	293	6.27006206	Insect	Bunyavirales	5.49E-151
soybean thrips bunya-like virus 2 isolate fusca15 PV068190	8224	soybean thrips bunya-like virus 2 isolate fusca (PV068190)	9152	97.846	8219	89.805507	Insect	Bunyavirales	0
Soybean thrips denso-like virus 3 isolate fusca14 putative protein 1 and putative protein 2 PV083565	306	soybean thrips denso-like virus 3 isolate fusca (PV083564)	1560	80.488	287	18.3974359	Insect	Piccovirales	1.61E-57
orthotospovirus tomatomaculae fusca18 segment L PV068193	8903	9.orthotospovirus tomatomaculae isolate fusca5 , L segment (PV068187)	7198	98.069	7198	100	Plant and insect	Bunyavirales	0
orthotospovirus tomatomaculae fusca17 segment S PV068192	231	orthotospovirus tomatomaculae isolate fusca6, S segment (PV068189)	1727	99.134	231	13.3757962	Plant and insect	Bunyavirales	4.55E-118
orthotospovirus tomatomaculae fusca16 segment M PV068191	1276	orthotospovirus tomatomaculae isolate fusca7, M segment (PV068188)	4774	97.962	1276	26.7281106	Plant and insect	Bunyavirales	0
Peanut field		**Clover field**							
soybean thrips denso-like virus 3 isolate fusca (PV083564)	1560	soybean thrips denso-like virus 3 isolate fusca14 PV083565	306	80.488	287	93.7908497	Insect	Piccovirales	1.14E-57
soybean thrips bunya-like virus 2 isolate fusca (PV068190)	9152	soybean thrips bunya-like virus 2 isolate fusca15 PV068190	8224	97.846	8218	99.9270428	Insect	Bunyavirales	0
orthotospovirus tomatomaculae isolate fusca7, M segment PV068188	4774	orthotospovirus tomatomaculae fusca16 segment M PV068191	1276	97.962	1276	100	Plant and insect	Bunyavirales	0
Orthotospovirus tomatomaculae isolate 15-122 segment S, complete sequence	1727	orthotospovirus tomatomaculae fusca17 segment S PV068192	231	99.134	231	100	Plant and insect	Bunyavirales	2.13E-118
11 orthotospovirus tomatomaculae isolate fusca7, M segment (PV068188)	4774	orthotospovirus tomatomaculae fusca18 segment L (PV068193)	8903	98.069	7198	80.849152	Plant and insect	Bunyavirales	0
orthotospovirus tomatomaculae isolate fusca12 segment L PV078566	404	orthotospovirus tomatomaculae fusca18 segment L (PV068193)	8903	98.01	402	4.51533191	Plant and insect	Bunyavirales	0
soybean thrips bunya-like virus 6 PV097222	4673	soybean thrips bunya-like virus 6 isolate fusca19 PV068194	293	98.976	293	100	Plant and insect	Bunyavirales	1.07E-151
Blattodean nairo-related virus isolate fusc	448	blattodean nairo-related virus fusca22 PV083556	894	100	467	52.2371365	Insect	Bunyavirales	0
9.orthotospovirus tomatomaculae isolate fusca5 , L segment (PV068187)	7198	orthotospovirus tomatomaculae fusca23 segment L PV083557	552	99.19	247	44.7463768	Plant and insect	Bunyavirales	1.75E-127
9.orthotospovirus tomatomaculae isolate fusca5 , L segment (PV068187)	7198	orthotospovirus tomatomaculae fusca24 segment L (PV083558)	536	97.761	536	100	Plant and insect	Bunyavirales	0
9.orthotospovirus tomatomaculae isolate fusca5 , L segment (PV068187)	7198	orthotospovirus tomatomaculae fusca23 segment L (PV083557)	552	99.19	247	44.7463768	Plant and insect	Bunyavirales	1.75E-127
11 orthotospovirus tomatomaculae isolate fusca7, M segment (PV068188)	4774	orthotospovirus tomatomaculae fusca25 segment M (PV083559)	557	98.205	557	100	Plant and insect	Bunyavirales	0
Orthotospovirus tomatomaculae isolate fusca11 segment S (PV078565)	396	orthotospovirus tomatomaculae fusca26 segment S (PV083560)	1062	98.737	396	37.2881356	Plant and insect	Bunyavirales	0
10 orthotospovirus tomatomaculae isolate fusca6, S segment (PV068189)	1727	orthotospovirus tomatomaculae fusca27 segment S (PV083561)	792	98.987	790	99.7474747	Plant and insect	Bunyavirales	0
10 orthotospovirus tomatomaculae isolate fusca6, S segment (PV068189)	1727	orthotospovirus tomatomaculae fusca28 segment S (PV083562)	765	99.085	765	100	Plant and insect	Bunyavirales	0

Figures

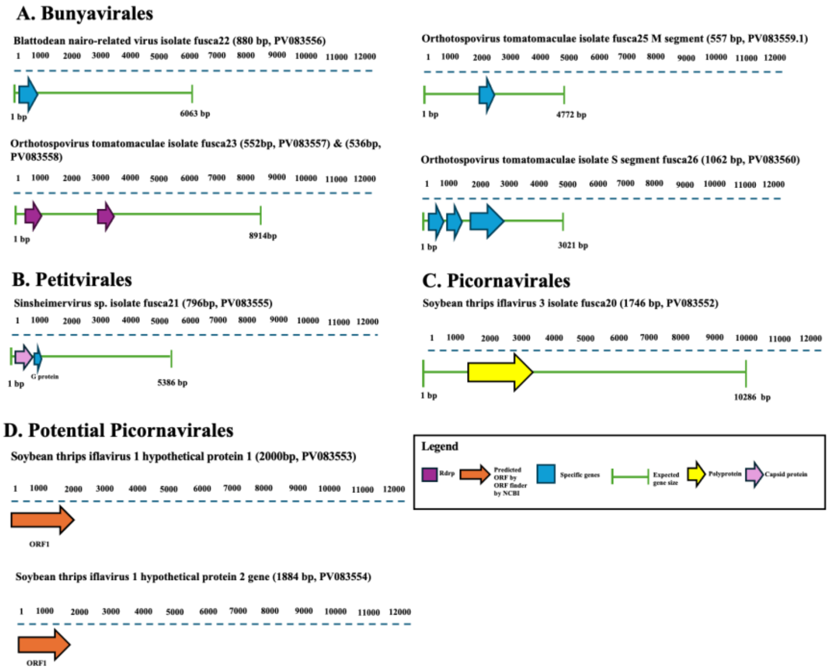


Figure 1. Graphical representation viruses found from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut in 2022. The arrow represents different genes of viruses belonging to (A) Bunyavirales; (B) Petitvirales; (C) Picornavirales ; (D) Potential Picornavirales; ORF: Open reading frames; RdRP: RNA-dependent RNA polymerase; NCBI: National Center for Biotechnology Information; bp: base pair

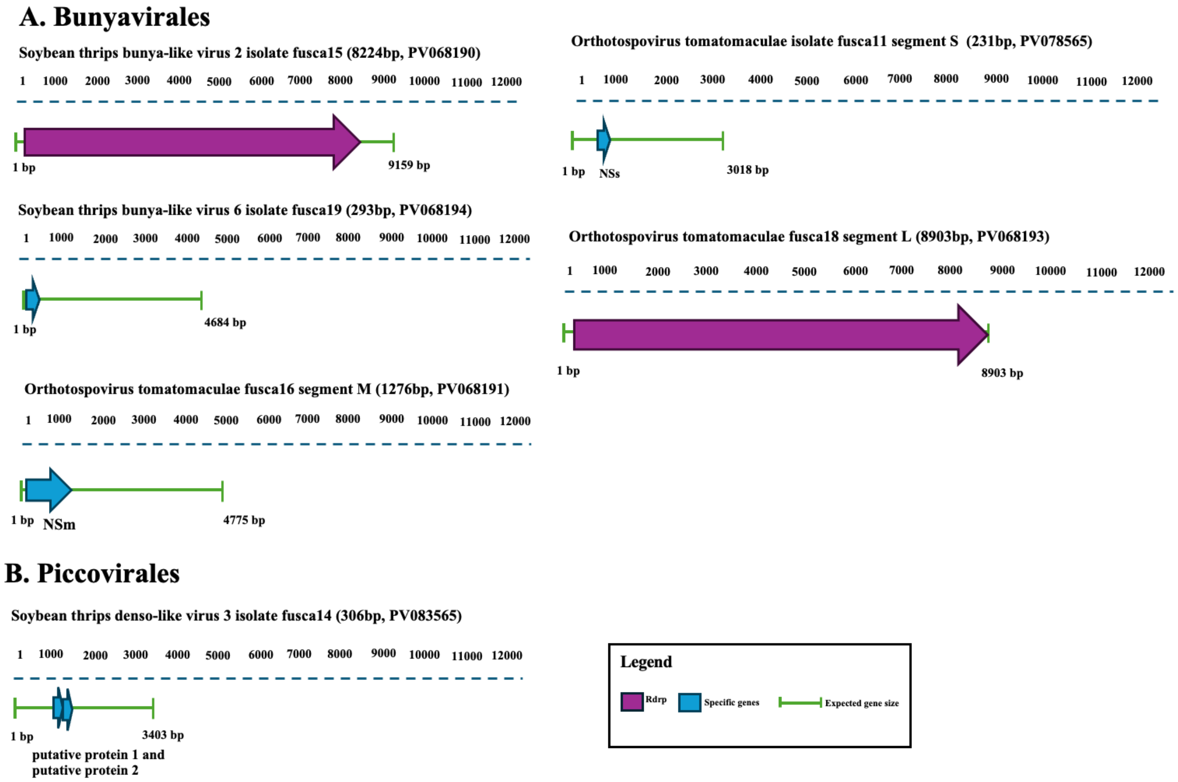


Figure 2. Graphical representation viruses from Hiseq of Tobacco thrips *Frankliniella fusca* (Hinds) collected in Wiregrass, Alabama in Peanut field in 2023. The arrow represents different genes of viruses belonging to (A) Bunyavirales (B) Piccovirales; bp: base pair; Nsm:non –structural; NSs: non-structural silencing suppressor protein movement protein; RdRP: RNA-dependent RNA polymerase; NCBI: National Center for Biotechnology Information; bp:base pair

Figure 2. Graphical representation viruses found from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut in 2023. The arrow represents different genes of viruses belonging to (A) Bunyavirales (B) Piccovirales; bp: base pair; Nsm:non –structural; NSs: non-structural silencing suppressor protein movement protein; RdRP: RNA-dependent RNA polymerase; NCBI: National Center for Biotechnology Information; bp: base pai

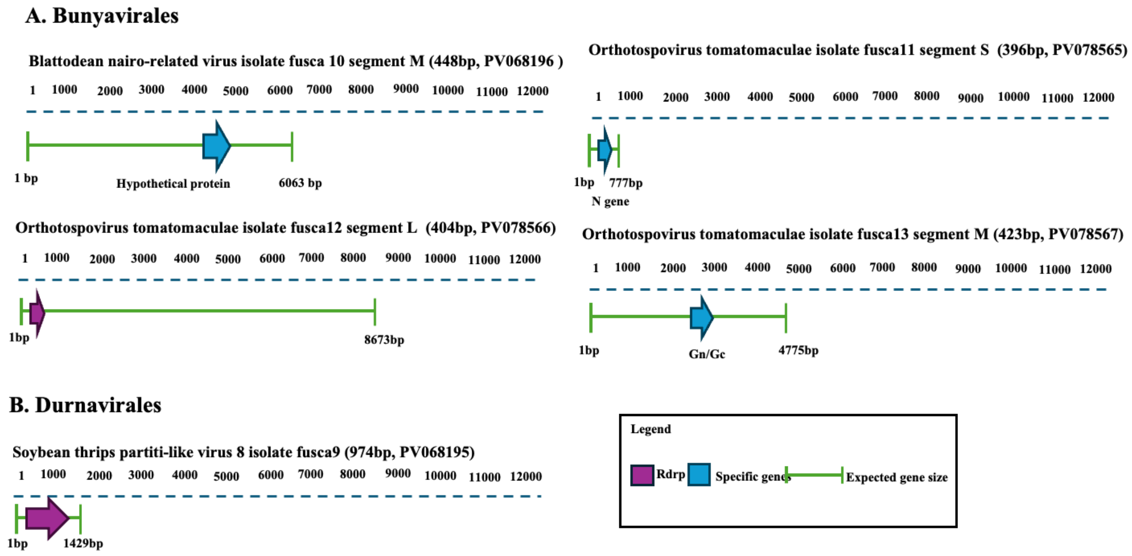


Figure 3. Graphical representation viruses found from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in clover in 2022. The arrow represents different genes of viruses belonging to (A) Bunyavirales, (B) Durnavirales; bp: Base pair; N: Nucleocapsid gene; RdRP: RNA-dependent RNA polymerase; NCBI: National Center for Biotechnology Information; bp: base pair

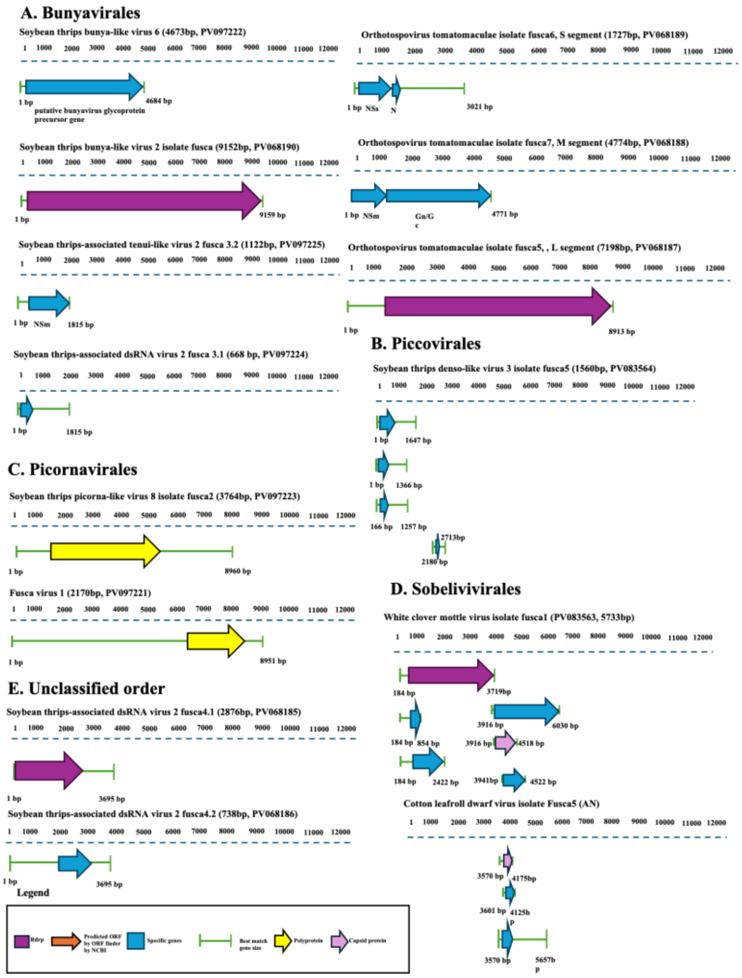


Figure 4. Graphical representation viruses from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in clover in 2023 Viruses belonging to (A) Bunyavirales (B) Petitvirales; (C) Picornavirales ; (D) Potential Picornavirales; bp: base pair; Nsm:non –structural; NSs: non-structural silencing suppressor protein movement protein;Gn/Gc: Glycoprotein precursor protein; ORF: Open reading frames; RdRP: RNA-dependent RNA polymerase; NCBI: National Center for Biotechnology Information; bp: base pair

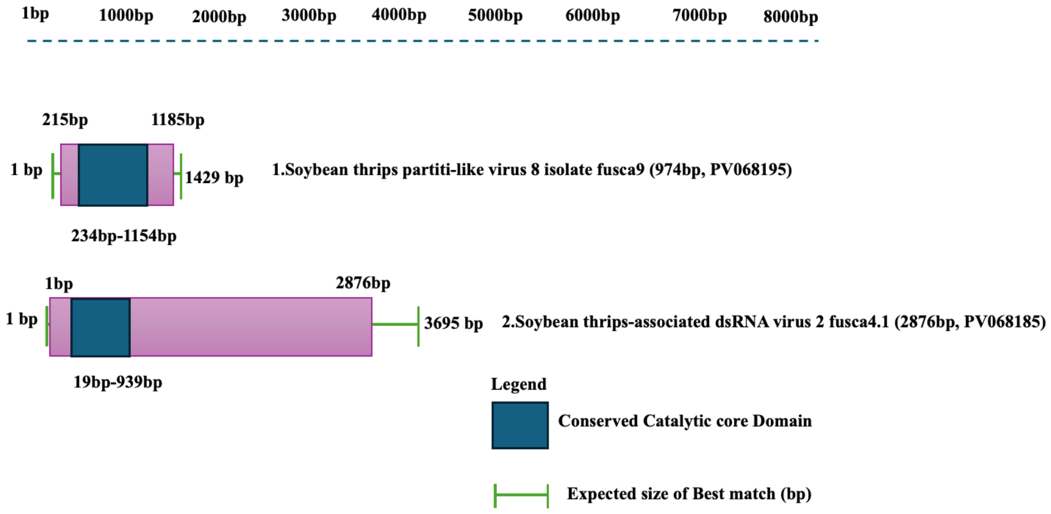


Figure 5. Viruses sharing the Conserved catalytic core domain predicted by NCBI Conserved Catalytic Domain through RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama: Soybean thrips partiti-like virus 8 isolate fusca9 (974bp, PV068195); Virus collected in *F. fusca* in clover in 2023: Soybean thrips-associated dsRNA virus 2 fusca4.1 (2876bp, PV068185); bp: base pair; NCBI: National Center for Biotechnology

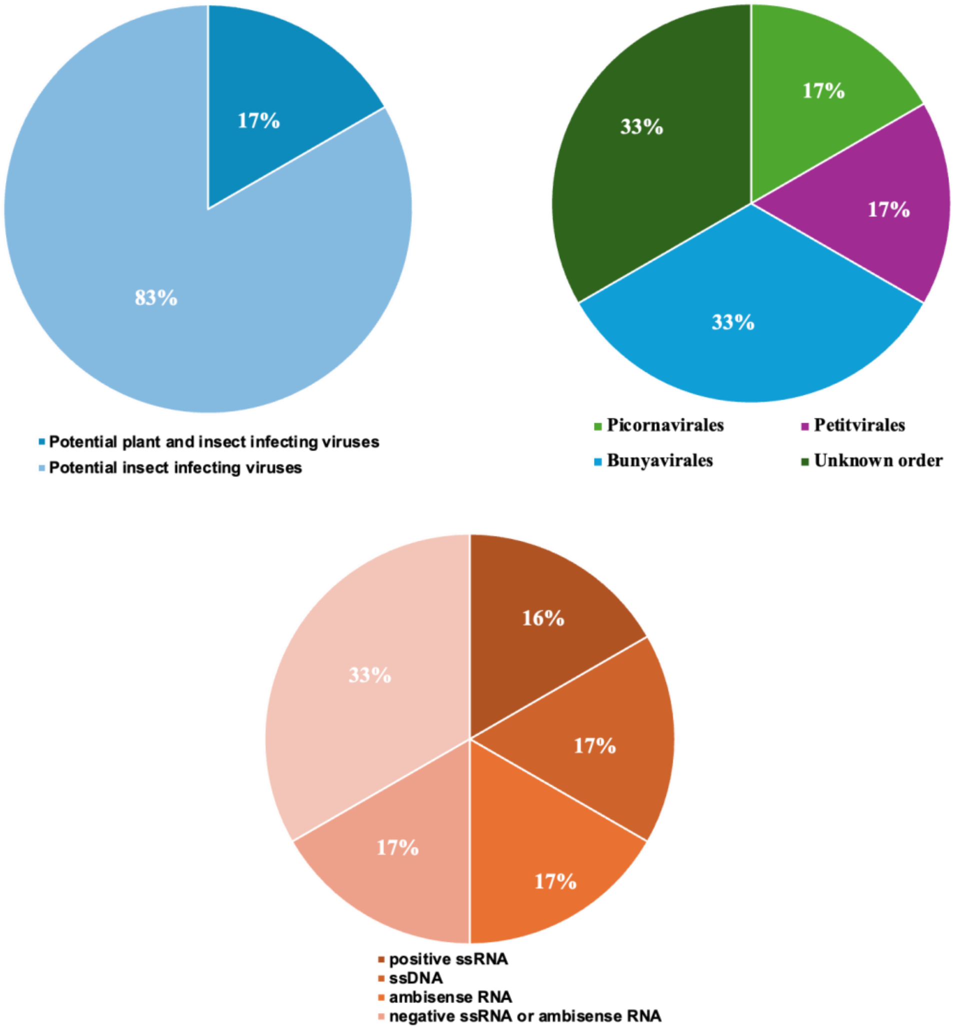


Figure 6. Abundance of viruses A. Potential plant infecting and potential insect infecting viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Wiregrass, Alabama in peanut field in 2022 B. Order of viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut field in 2022 C. Genome type of virus discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut field in 2022

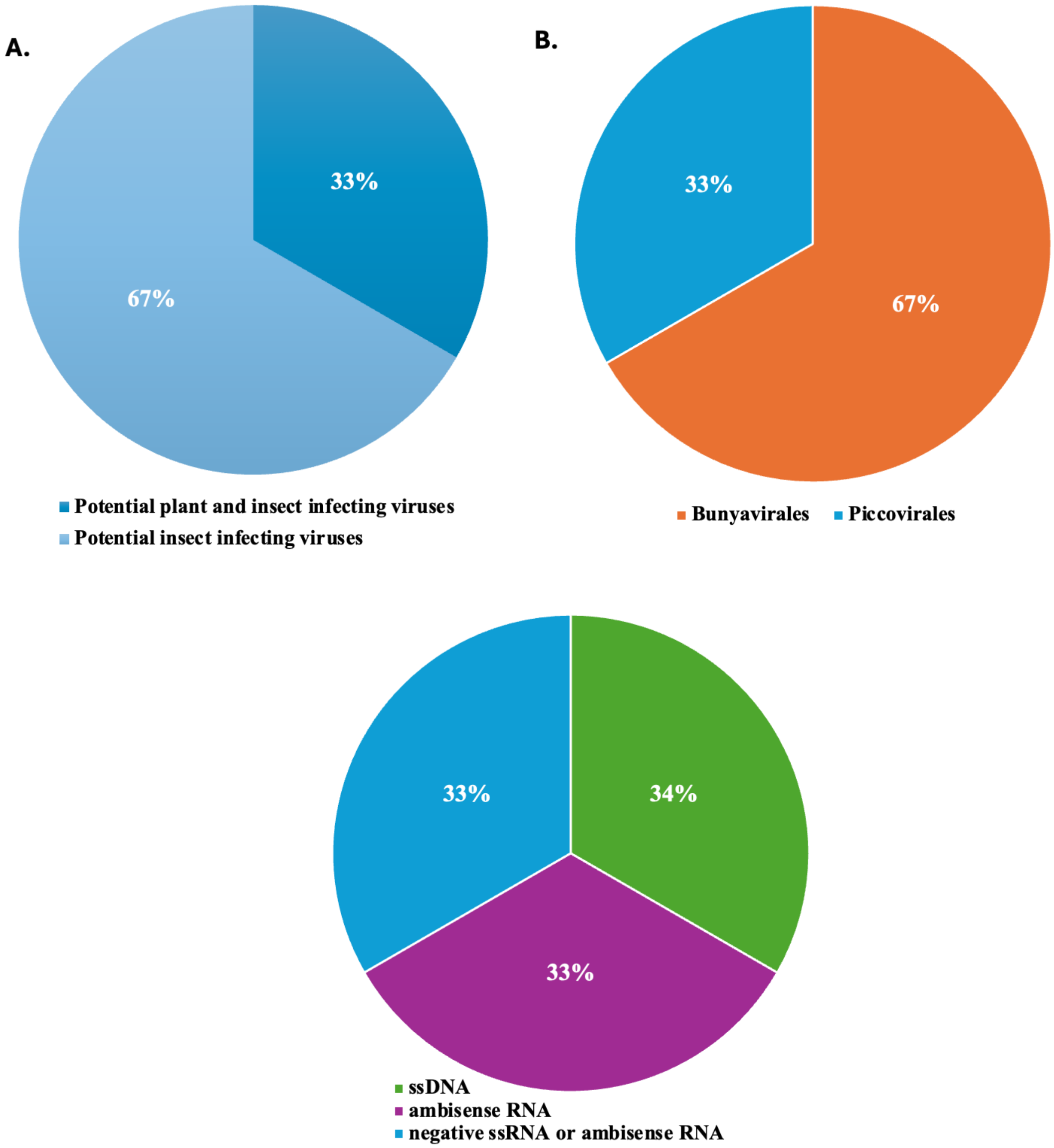


Figure 7. Abundance of viruses A. Potential plant infecting and potential insect infecting viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Wiregrass, Alabama in peanut in 2023 B. Order of viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut in 2023 C. Genome type of virus discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in peanut in 2023

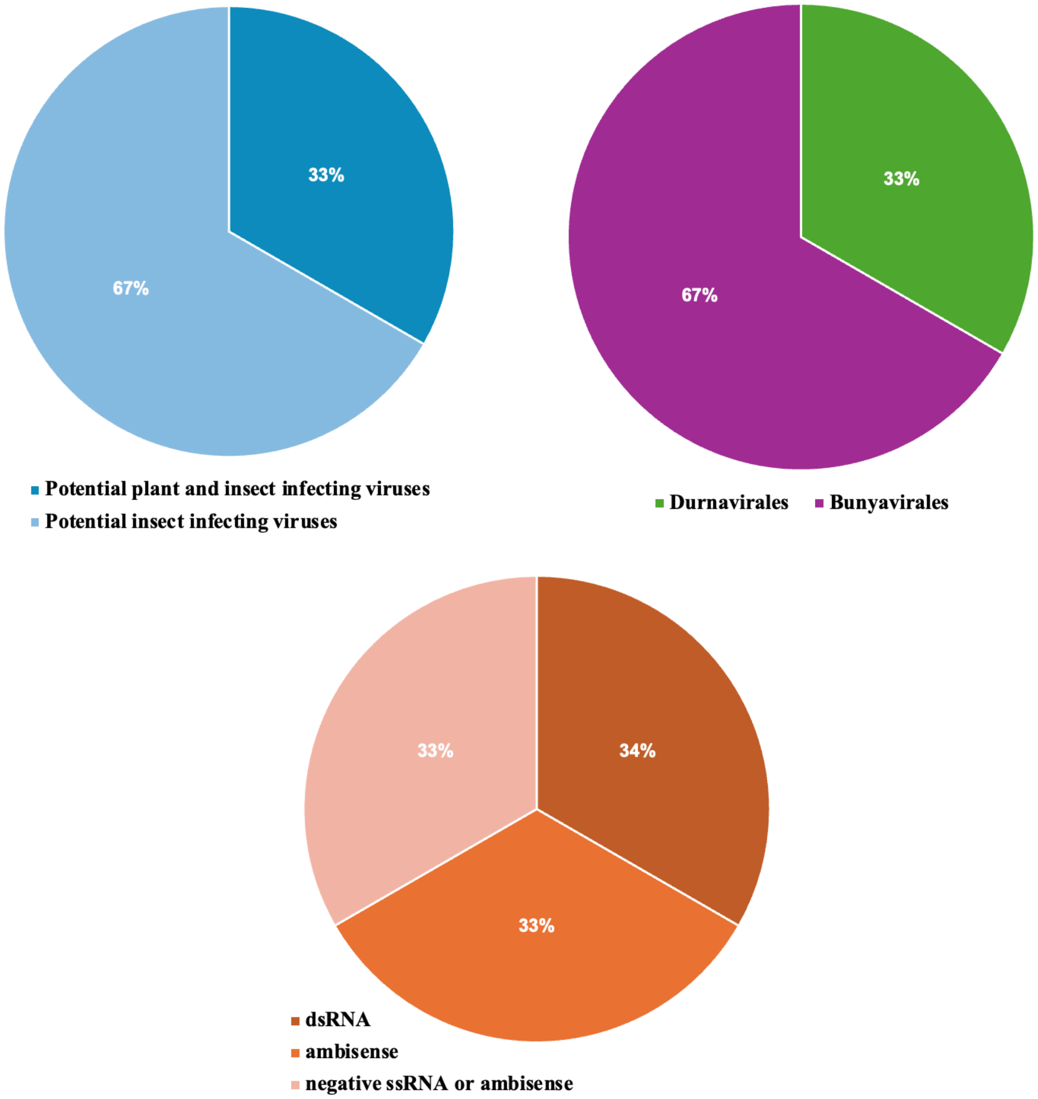


Figure 8. Abundance of viruses A. Potential plant infecting and potential insect infecting viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Wiregrass, Alabama in Clover in 2022 B. Order of viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in clover in 2022 C. Genome type of virus discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in clover in 2022

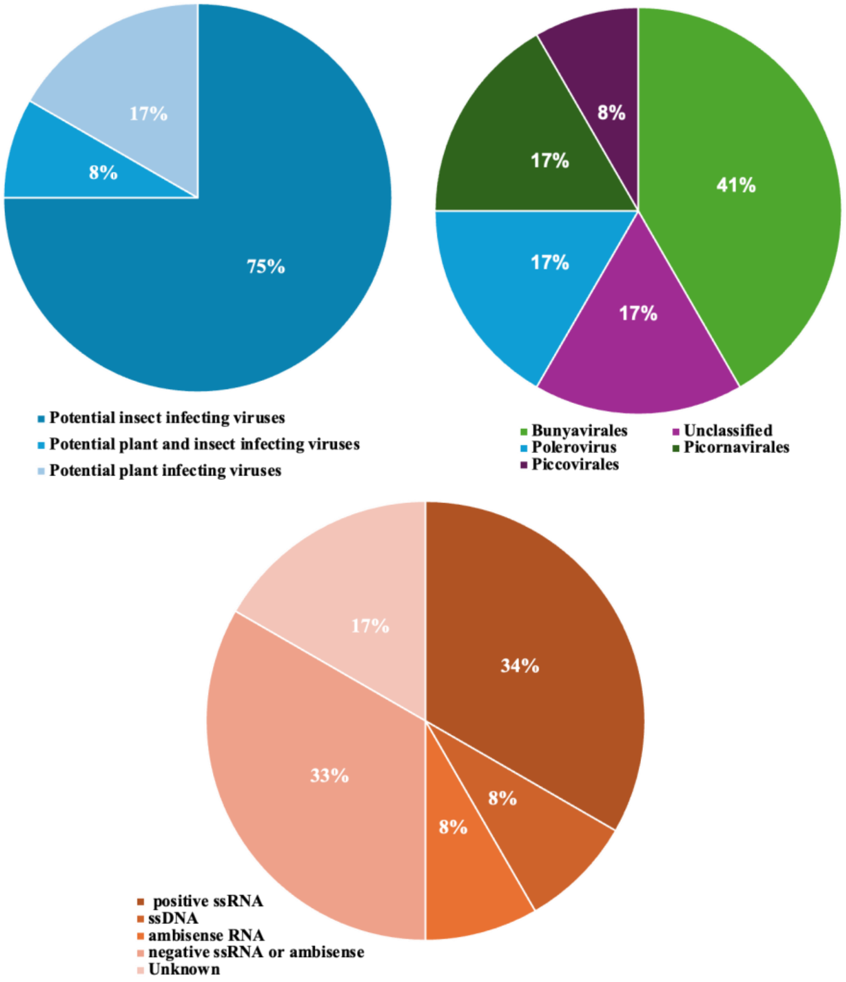


Figure 9. Abundance of viruses A. Potential plant infecting and potential insect infecting viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Wiregrass, Alabama in Clover in 2023 B. Order of viruses discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama, in clover in 2023 C. Genome type of virus discovered from RNA-seq analysis of Tobacco thrips (*Frankliniella fusca* Hinds) collected in Headland, Alabama in clover in 2023

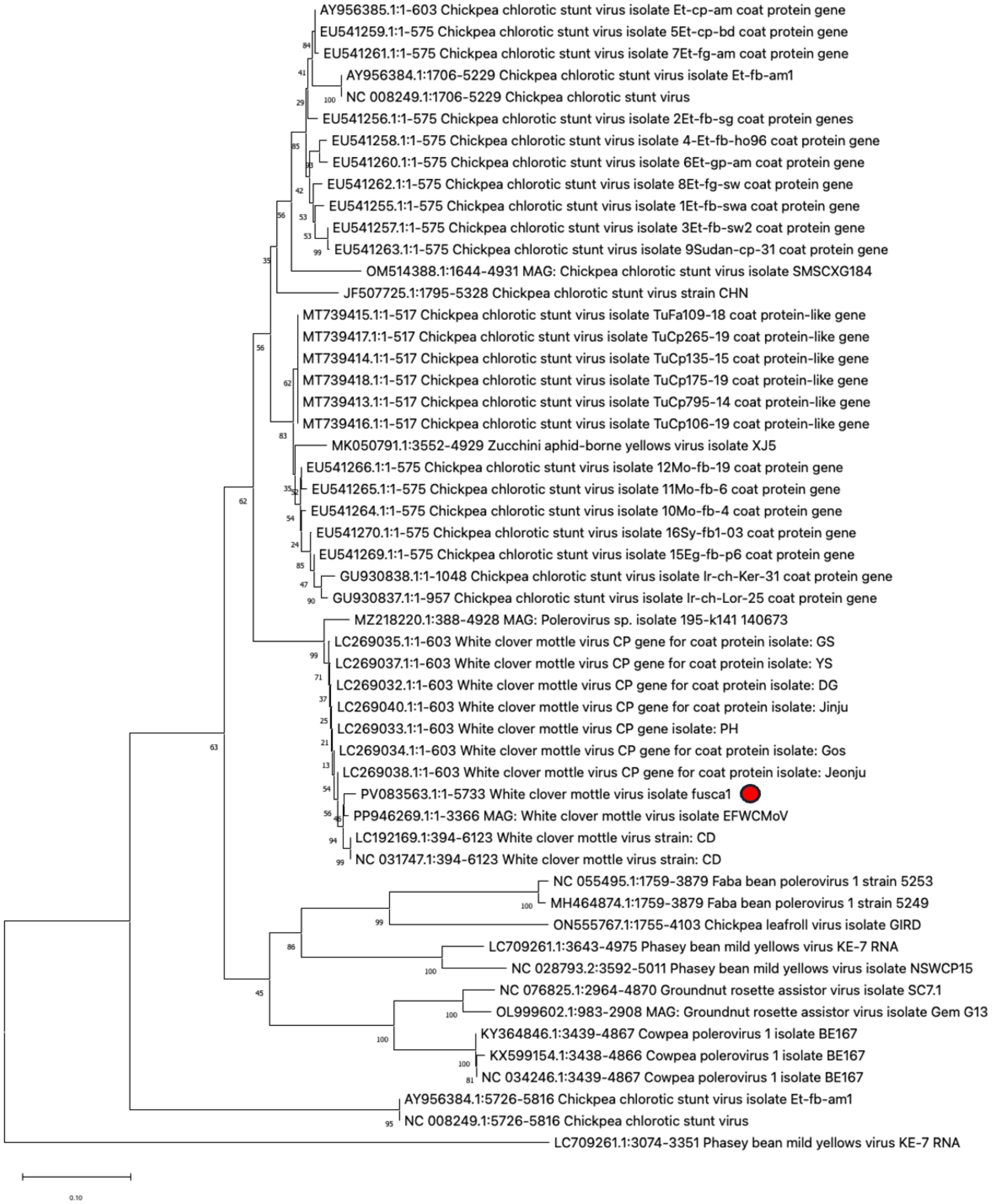


Figure 10. Phylogenetic tree of first fifty blast result available in NCBI of White clover mottle virus (PV083563.1) discovered in Tobacco thrips (*F. fusca*) collected from clover in 2023 from Headland, Alabama. The alignment was done using MUSCLE. The values at the nodes denote the percentage of bootstrap support (500 replicates); Red dot is the virus that was discovered through the study

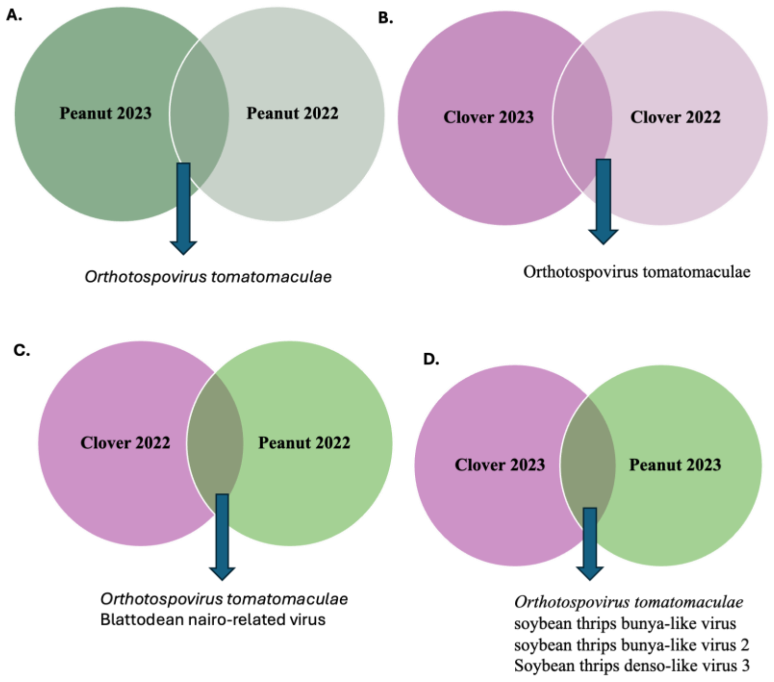


Figure 11. Common viruses discovered in *F. fusca* in peanut and clover in 2022 and 2023 A. Common viruses discovered from RNA-seq analysis of Tobacco thrips, *Frankliniella fusca* Hinds collected in Headland, Alabama between peanut field 2023 and peanut field 2022 B. Common viruses discovered from RNA-seq analysis of Tobacco thrips, *Frankliniella fusca* Hinds collected in Headland, Alabama between clover field 2023 and clover field 2022 B. Common viruses discovered from RNA-seq analysis of Tobacco thrips, *Frankliniella fusca* Hinds collected in Headland, Alabama between clover field 2022 and peanut field 2022 C. Common viruses discovered from RNA-seq analysis of Tobacco thrips, *Frankliniella fusca* Hinds collected in Headland, Alabama between clover field 2022 and peanut field 2022 D. Common viruses discovered from RNA-seq analysis of Tobacco thrips, *Frankliniella fusca* Hinds collected in Headland, Alabama between clover field 2023 and peanut field 2023

Supplemental Chapter 1

Virome characterization of *Culex erraticus* reveals fifteen different viruses using dual Bioinformatic Approaches

Aasma Sharma[1], Kyle Oswalt[1], Natalie Wong[1], John F. Beckmann[1], Kathleen Martin[1]*

[1] Department of Entomology and Plant Pathology, Auburn University, Auburn, Alabama

*Corresponding author

Abstract

Mosquitoes are the cosmopolitan important insect vector that is a human threat as well as reservoir of various viruses. *Culex erraticus* is a prevalent mosquito species in Auburn, Alabama. As *Culex erraticus* harbors medically important viruses, studies are often directed exclusively to human pathogens, without consideration as to what else may be present. To address this, an RNA-seq analysis was conducted to understand viral diversity. To ensure a complete virome, we used two different approaches for bioinformatic analyses to eliminate bias from one. In first method the previously characterized genome of *Culex quinquesfasciatus* (GCF_015732765.1) was used as a reference genome with a mapping cutoff followed by the annotation of assembled contigs with a universal blast and traditional megablast against the NCBI Nucleotide database (01 sept 2023) in galaxy platform. In a second method, the contigs were annotated with a specific blast against NCBI viral genomic database (viral.1.1.genomic.fna) without mapping to a reference. In addition, the cross-validation of the contigs generated by two methods was conducted to identify the common and unique viral contigs. RNA-seq analysis suggested total of fifteen different viruses. Both methods identified three common viruses like Merida virus isolate Cx.erraticus2 (PQ963471), Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472), and Zheijang mosquito virus isolate

Cx.erraticus17 (PQ963484). The first method also identified total ten unique viruses and the second method generated two unique viruses. This suggests that having different approaches to bioinformatic analysis complement each other to gain better representation of virus the in Virome diversity study.

Keywords: *Culex erraticus*, virome, RNA-seq

Introduction

Mosquitoes (Diptera: *Culicidae*) are a cosmopolitan insect group that has more than 3600 insect species, majority of which are the vectors of parasites, viruses, and bacteria (Nebbak et al., 2021). In addition to viruses, they are globally recognized vectors for medically and veterinary significant pathogens contributing to millions of diseases and The *Culex* and *Aedes* genera transmit epidemic viral diseases belonging to *Flaviviridae* and *Togaviridae* (Gómez et al., 2023). Other major groups of viruses of interest that can be transmitted by mosquitos include members of *Bunyavirales* and *Rhabdoviridae,* viruses that are known to cross kingdoms and infect a diverse array of hosts. The eggs of the majority of *culex* and *anopheles'* mosquitoes important in transmitting human pathogens require water, thus agricultural fields, wastewater, sewage, bird baths and anywhere where water can collect such as tires are attractive areas for habitation (Morse et al., 2019).

The sub-tropical climate of the gulf coast regions of United States including Alabama is a complex habitat. In association with the fisheries in our state, there is an optimal environment for birds such as raptors (bald eagles, ospreys, various hawk species), wading birds (multiple heron species), vultures, kingfishers, geese, ibis, cormorants and various diving ducks throughout the year. There are also songbird species in the trees and fields surrounding the ponds. Other animals present include, musk turtles and sliders, and less common, softshell and snapping turtles. Snakes are also present and very common such as water snakes, cottonmouths, copperheads, rattlesnakes and rat snakes. Therefore, there are many different important mosquito species in Alabama for example: *Aedes aegypti, Aedes albopictus, Culex pipien, tarsalis, quinquefasciatus* and *Anopheles quadrimaculatis* (Morse et al., 2019). Also, previously in the Tuskegee National Forest, *Culiseta melanura, Culex restuans, Aedes vexans, Coquillettidia perturbans, Cx. erraticus, Cx.*

peccator, Cx. territans, and Ochlerotatus sticticus were all identified as feeding on a wide variety of birds and reptiles(Estep et al., 2011).

As humans urbanize and prepare land for agriculture they come into contact with diverse mosquito species, reservoirs, and novel diseases. Even in the first world, this process is still ongoing. To investigate novel viruses and microbiomes of mosquitos in our locale, we focused our study on an ecosystem having three key factors: proximity to housing developments and human habitation, proximity to undeveloped land such as forested areas, and the proximity to aquatic fisheries agriculture stations that would support mosquito populations. The overlap of these three factors can foster disease outbreaks. Specifically, in Auburn, a fisheries pond was selected in the north part of the community, in close proximity to the Kreher forest and less than a mile from a large housing development. We sought to investigate the mosquito populations of this unique habitat in the context of fisheries agriculture which is important to the state of Alabama. We collected mosquitoes during early fall 2022 with the objective of assessing the diversity of species locally and also to identify the most abundant mosquito vector. Therein, we hoped to next assess the diversity of viruses within that predominant vector species. One hypothesis was that the predominant vector would carry a variety of viruses which may not have been characterized previously. We did not expect to find dangerous viruses, given the lack of reports of human infections locally. Our investigative interest was more in determining the mosquito virome associated with Alabama aquaculture. Also, as two main groups such as rhabdovirues and bunyaviruses are known to have diverse host ranges, it was of interest to surveil if any were present in the area.

Materials and Methods

Collection sites (Information provided by Kyle Oswalt and Natalie Wong)

Mosquitoes utilized for this study were trapped at the Auburn University, E.W. Shell Fisheries center located in Auburn, AL. USA. The E.W. Shell Fisheries center is a 1600-acre research station consisting of infrastructure and natural habitat. Man-made ponds facilitate studies utilizing diverse species of catfish and trout. The fisheries also provide a habitat for a variety of wildlife including Bald Eagles, Belted Kingfishers, blue herons, wading birds, songbirds, reptiles, and amphibians. The fisheries are located on AL- 147, Auburn, Alabama (Coordinates) south of the Kreher Preserve & Nature Center and north of Auburn campus. This unique location of the fisheries allows it to serve as a novel model representing the interface between a more urban environment to the south and a more natural environment to the north. This location allows for a more biodiverse ecosystem to study zoonosis.

Mosquito collection and identification (Information provided by Kyle Oswalt and Natalie Wong)

Live adult mosquitoes were collected using Centers for Disease Control (CDC) miniature light traps (Model 512, John W. Hock) baited with CO_2 emitted from approximately 0.9 kg of dry ice per trap. .The traps were set overnight from 4 PM until 10 AM the following morning. The collection period extended from late August through the end of October 2022. Traps were strategically placed along the edges of the fisheries and forests or at recently drained ponds within E.W Shell Fisheries (Auburn, AL) to gather a diverse population of mosquitos. Upon retrieval, live adult mosquitoes were transported to the laboratory and euthanized by incubation at -20°C for 10 minutes. Subsequently, the mosquitoes were sexed and identified to species level using a stereomicroscope. Specimens were then grouped by species and sex into pools, with each pool containing a maximum of 20 adult mosquitoes. These pools were submerged in 100% ethanol and stored at -20°C until nucleic acid extraction.

Viral RNA Extraction (Information provided by Micheal Mayfield)

As *Culex erraticus* was the predominant species at this location, the study was focused on this species moving forward. Female *C. erraticus* mosquitoes were removed from ethanol and allowed to dry under a laminar flow hood for 10 minutes. Individuals were then pooled into four different groups ranging from 15 to 17 mosquitoes and placed into 1.5ml microcentrifuge tubes. Samples were homogenized by pestle in liquid nitrogen and RNA was extracted from each group with 1ml of TriZol (Invitrogen, Waltham, MA) according to manufacturer's instructions. Modifications to manufacturer's instructions were implemented to increase the yield and quality of RNA. These steps included repeating the chloroform extraction twice, conducting salt precipitation with 100% Isopropanol and Sodium Acetate (pH 5.2), and 2 additional cold 70% ethanol washes. Total RNA concentration was quantified by nanodrop (Thermo Scientific, Waltham, MA) and ranged from 1600 ng/ul to 1800 ng/ul. The 260/280 ranged from 2.01 to 2.05 and 260/230 ranged from 1.88 to 1.92. RNA samples from the four different collections were pooled to create one sample for sequencing. Their concentrations were normalized, so that each sample contributed equally within the final pooled sample. Samples were then pooled, to increase the number of insects from different timepoints to be included in the virome analysis. For example, an equal amount of RNA was included for each collection timepoint, although not necessarily the same volume.

Ribosomal RNA Depletion and Sequencing (Information provided by Micheal Mayfield)

The total RNA sample was sent to SeqCenter (Pittsburgh, Pennsylvania) and prepared for RNA-seq through their traditional pipeline. The sample was treated with RNAse-free DNase (Invitrogen, Waltham, MA) and the library prepped using an Illumina Stranded Total RNA Prep Ligation with Ribo-Zero Plus kit and 10bp unique dual indices (UDI) (Illumina, San Diego, CA). Sequencing was performed on a NovaSeq X Plus (Illumina, San Diego, CA). This produced paired-end 150 bp reads with the expectation of 100M reads. Demultiplexing, quality control, and adapter trimming was performed with bcl-convert (v4.1.5) through SeqCenter.

Bioinformatic analysis pipelines

Once the sequences were received, the two different analyses pipelines were followed. In Method 1, the sequences were mapped the contigs and then annotated non-mapped contigs to the universal NCBI database. In Method 2 the sequences were annotated to the virus NCBI viral genomic database (viral.1.1.genomic.fna) (*NCBI Virus*, n.d.) without mapping.

In method one, the mapping of reads was done using Bowtie2 (Langmead et al., 2009a), (Langmead et al., 2009b) against the full reference genome of *Culex quinquesfasciatus* (GCF_015732765.1). Assembled contigs were separated based on size of contigs using Filter sequences by length (Galaxy Version 1.2) into 0-400 bp, 400-1000bp and >1000 bp. These contigs were blasted using NCBI BLAST+ blastn (Cock et al., 2015) (Camacho et al., 2009a) against NCBI-NT 2023 universal database. The largest contig with highest amino acid identity (%) of viruses with >400 bp long contigs were considered after final assembly (Batson et al., 2021). The taxonomic report was also generated using kraken2 (Version 2.1.3) (Wood & Salzberg, 2014). For unique contigs longest contig with high Amino acid identity (%) annotation was considered to ensure the sensitivity of virus during detection and reduced need of manual genome extension (X. Deng et al., 2015) (Z. Deng & Delwart, 2021).

In the second method, raw sequencing data were subjected to quality control using FastQC (version 0.74) and trimmed with Trimmomatic (version 0.3.9). To remove low-quality reads and adapters. High-quality reads were then assembled de novo using Trinity (Galaxy version 2.11.0) Citation to generate contigs. These contigs were queried against the NCBI viral genomic database (viral.1.1.genomic.fna) (*NCBI Virus*, n.d.) which includes complete viral genome sequences using BLASTn (BLAST+ 2.16.0) Citation with an E-value cutoff of 1e-5. Significant viral matches were clustered to generate consensus sequences that incorporate IUPAC nucleotide codes to account for polymorphic sites.

Cross-Validation

The contigs from both analyses were cross blast by making contigs from Method 1 as the database by using the command "makeblastdb" and blasting contigs from method 2 against the database using the command "blastn". (Camacho et al., 2009b).

Phylogenetic analysis

To understand the close relative of the Merida virus isolate Cx.erraticus2 (PQ963471) to other complete or nearly complete Merida virus isolate present in the NCBI database the maximum likelihood phylogenetic tree was generated using Muscle Multiple Sequence alignment (Edgar, 2004) and ITOL tree maker (Letunic & Bork, 2024). The accession number of the virus >10,000kb were used to generate the trees: NC_040599.1, MW434769.1, MW434773.1, OM817546.1, MW434771.1, MW434770.1, MW434772.1, MH188000.1, MH310083.1, OL700074.1, MT577803.1, NC040532.1, MF882997.1, OQ725971.1, OQ725976.1, OQ725974.1, OQ725980.1, OQ725972.1, PQ773030.1, MK440621.1, PQ773029.1, OQ725979.1, OQ725979.1, OQ725977.1, OQ725977.1, OQ725978.1, OQ725978.1 and OQ725973.1.

Results

Culex erraticus was the most abundant mosquito (Information provided by Natalie Wong)

Between August through October of 2022, a total of 11 mosquito collections were sampled from multiple locations within E.W. Shell Fisheries in Auburn, Alabama. Mosquitoes were collected from ecotones between forested areas and fisheries, as well as recently drained ponds, to enhance biodiversity surveillance. From these collections, a total of 240 female mosquitoes were collected, sexed, and morphologically identified to genus and species with the use of a stereomicroscope and taxonomic key. The results revealed the presence of 3 unique genera and species at the fisheries, *Culex* (Melanoconion) erraticus, *Psorophora* sp., and *Anopheles crucians*. Of the three mosquito species collected, *Culex erraticus* was the most abundant, comprising 97% of the total specimens, followed by *Psorophora spp*. (2.5%) and *Anopheles crucians* (0.4%). The presence of these species was expected given their widespread distribution in the southeastern region of North America (Darsie and Ward, 2016). Due to its high abundance and generalist feeding behavior, *Culex erraticus* was selected for further study (Table 3, Figure 2).

Bioinformatic statistics from Method 1

The total 158709252 reads (79354626 paired-end forward and reverse reads each) has total 10.5Gbp bases. The length range of the sequences is 35bp to 150 bp long with GC 45%, and Phred score 33 were received (Table 1). The BioProject ID for the raw read sequences is PRJNA1227547(submission id:SAMN46988613). Since there was no presence of the *Culex erraticus* genome in the NCBI database, the mapping was done to the available mosquito genome of *Culex quinquesfasciatus.* In total 16.72% sequences aligned to the reference genome of *Culex quinquesfasciatus* (GCF_015732765.1) and the rest 83.28% of sequences did not align to the reference genome. This nonaligned portion of the reads was moved ahead for the further assembly. With the expectation of removing the significant read overlap, the nonaligned data from mapping the data was assembled using the Trinity tools. The total non-mapped trinity transcripts are

1006209 with GC 41.27% with average contig length 580.25bp, assembled bases 583848302. 128835 contigs were of size <400bp, 326633 contigs were of size 400-1000bp, 63986316bp contigs were >1000bp. From the contig range of >400bp, we discovered total 742 different viral contigs that belonged to 12 different viruses (Table 1). To find the further information on the potential viral contigs from the Method 1 after final blast the kracken2 report was generated against the viral 2020 database that identified the 48 different viral taxonomy (Table 4). However, after the final assembly and filter we discovered 12 different viruses (Table 2).

Bioinformatic statistics from Method 2 (Information provided by Kyle Oswalt)

The de novo assembly performed in Method 2 yielded 1,187,014 contigs, with sequence lengths ranging from 172 bp to 33,372 bp. The total assembly spanned 671,847,487 bp, resulting in an average contig length of 566 bp and an N50 value of 726 bp. Contig size distribution was broad, with the majority falling between 500–3000 bp, indicative of a well-dispersed transcriptomic landscape. Quality filtering was conducted using [FastQC version 0.74], followed by adapter and low-quality base trimming via [Trimmomatic version 0.3.9]. Trinity (Galaxy version 2.11.0) was subsequently employed to perform the de novo assembly. The resulting contigs were queried using BLASTn (BLAST+ 2.16.0) against the NCBI viral genomic database (viral.1.1.genomic.fna), utilizing an E-value threshold of 1e-5 to optimize detection sensitivity. This analysis identified 45 unique contigs (~0.004% of the total) as viral based on significant alignments. Among these, two contigs corresponded to viruses not recovered in Method 1: Spodoptera litura nucleopolyhedrovirus isolate Cx.erraticus5 (PQ963474) and Spodoptera littoralis nucleopolyhedrovirus isolate Cx.erraticus6 (PQ963475). The remaining 1,186,969 contigs failed to align with known viral genomes, suggesting origins from the host transcriptome, environmental contaminants, or potentially novel viral taxa. As with Method 1, no complete viral genome was represented by a single contig, necessitating additional scaffolding. For the Merida virus, 29 individual contigs were integrated using an IUPAC consensus-based approach to account for nucleotide ambiguity across overlapping regions. This process culminated in the reconstruction of a full-length genome sequence totaling 6,994 bp. Collectively, these resultss underscore the fragmented nature of virainnome representation in metatranscriptomic-data and highlight the critical role of contig integration in the accurate reconstruction of viral genomes. Further methodological details regarding contig overlap thresholds and manual validation criteria are provided in the Methods section.

Fifteen different viruses were discovered from the two methods

Six viruses were positive sense ssRNA viruses, three viruses were negative sense ssRNA viruses and eight viruses has uncharacterized genome (Figure 2). Seventy-one percent of viruses were RNA viruses (Figure 3).

From method 1, after mapping 16.23% host reads assembled to the *Culex quinquesfasciatus* (GCF_015732765.1) genome and 83.77% to non- host reads which were processed for further analysis. Upon assembly the total contigs generated were 1,006,209 with average contig length of 580.25 bp. For combined virus the longest contig annotation was chosen. For unique contigs, longest contig with high Amino acid identity (%) annotation was considered to ensure the sensitivity of virus during detection and reduced need of manual genome extension (Deng et al., 2015) (Z. Deng & Delwart, 2021).

The kracken 2 report on the total contigs from method 1 also included two viruses (Spodoptera litura nucleopolyhedrovirus and Zhejiang mosquito virus 1) that were found in method 2. The contig for these viruses could not be filtered out after final annotation. This report also suggests that there are viruses that matches to the Merida virus, Hubei virga-like virus 2, and

Hubei narna-like virus 17 whose contigs might have been filtered out in the final annotation. Still, there are potential of having more virus in our dataset whose contigs have not been recovered by either of the method or might have been cutoff because of the size of contig <400bp after final annotation.

The Method 1 generated ten unique viruses like Hubei virga-like virus 3 isolate Cx.erraticus1 (PQ963470), Culex rhabdovirus isolate Cx.erraticus7 (PQ963476), Ecclesville picorna-like virus isolate Cx.erraticus8 (PQ963477), Elemess virus isolate Cx.erraticus9 (PQ963478), Elemess virus isolate Cx.erraticus10 (PQ963479), Qingnian mosquito virus isolate Cx.erraticus11 (PQ963480), Rinkaby virus isolate Cx.erraticus14 (PQ963485), Serbia narna-like virus 2 isolate Cx.erraticus15 (PQ963482), XiangYun narna-levi-like virus 8 isolate Cx.erraticus16 (PQ963483), Bunyavirales sp. isolate Cx.erraticus12 (PQ869759) and Qingnian mosquito virus isolate Cx.erraticus13 (PQ963481). The method 2 also generated two unique viruses Spodoptera litura nucleopolyhedrovirus isolate Cx.erraticus5 (PQ963474) and Spodoptera littoralis nucleopolyhedrovirus isolate Cx.erraticus6 (PQ963475) (Table2).

From both methods we discovered the nearly complete genome of two virus like Merida virus isolate Cx.erraticus2 (PQ963471), and Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472) (Figure 5). Merida virus isolate Cx.erraticus2 (PQ963471), Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472) and Zheijang mosquito virus isolate Cx.erraticus17 (PQ963484*)* were also found by both methods (Table 2).

For Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472) the method 1 had a contig length of 663bp with 89.577% Amino acid identity with best match to Atrato picorna-like virus 1 strain An 1771-1 genomic sequence, sequence (MN661035.1) and method 2 had a contig length of 9012 bp with 77.16% Amino acid identity to Zhejiang mosquito virus 1 (NC033716.1). Upon cross blast the shorter contig overlapped to the longer contig so the longer contig was considered for the NCBI submission.

For Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472) the Method 1 generated 265 bp long viral contig with an alignment range of 81bp to 345bp to Hubei mosquito virus 4 (4971bp, NC_032231.1) whereas method 2 generated the same virus that is 3734 bp long viral contig with an alignment range 1 to 3643bp and 89.45% Amino acid identity to Hubei mosquito virus 4 (4971bp, NC_032231.1). So, the longer contig 3734 bp was considered for the NCBI submission.

Mononegavirales like viral sequences

There are two viruses pulled as the members of Mononegavirales belonging to family Rhabdoviridae with non-segmented single stranded negative sense RNA genome (Figure 5). The virus Merida virus isolate Cx.erraticus2 (PQ963471) is 11766 bp long & 98.74% Amino acid identity to the Merida virus isolate MERD-Mex07 (11798bp, NC 040599.1) with 99.72% viral coverage with 5 genes for 5 proteins: large protein (L), glycoprotein (G), nucleoprotein (N), phosphoprotein (P), and matrix protein (M). For the first time this virus genome was identified in *Culex quinquefasciatus* in Mexico in 2007 (Charles et al., 2016). There has been no reporting in *Culex erraticus* in USA as well as in Alabama.

The Culex rhabdovirus isolate Cx.erraticus7 (PQ963476) has 3666 bp long contig that has the 98.636 % Amino acid identity to Culex rhabdovirus strain CRV/Kern (MH188000.1). We identified three complete genes: Nucleoprotein gene (116bp to 1552bp) and Phosphoprotein gene (1642bp to 2844bp), Matrix protein gene (2899bp to 3459bp) and a chunk (80bp) of Glycoprotein located 3568bp to 3602bp. This virus was previously isolated from the host pool of *Culex pipiens complex*, *Cx. tarsalis* and *Cx. Erythrothorax* in California, USA in 2016(Sadeghi et al., 2018).

There is no report of its occurrence in Alabama. The closest relative is Merida virus isolate MERD-Mex07, complete genome (NC040599.1) with 98% Amino acid identity.

Wolframvirales like viral sequences

The RdRp of positive ss RNA of Serbia narna-like virus 2 isolate Cx.erraticus15 (PQ963482) identified. It is 2751 bp long with an alignment range 3 bp to 2753bp and 94 % Amino acid identity to Serbia narna-like virus 1 isolate 85061 RNA-dependent RNA polymerase gene, partial cds (3280 bp, MT822183.1) (Figure 5). The full genome of this virus has not been reported yet. It was first time reported in Hematophagous Arthropods from Belgrade, Serbia that had the pool of mosquitoes, ticks and bedbugs(Stanojević et al., 2020a). The closest relative is RNA-dependent RNA polymerase gene of *Narnaviridae sp*. strain YX788 (MW452310.1) with 93.99% Amino acid identity. Any previous report of this virus in USA is unknown and undocumented in any mosquito species including *Culex erraticus.*

Picornavirales like viral sequences

The RdRp of positive ss RNA of Ecclesville picorna-like virus isolate Cx.erraticus8 (PQ963477) which is 662 bp with 90.18% Amino acid identity to Ecclesville picorna-like virus isolate TriniPa-Ecclesville-E04 nonfunctional polyprotein gene, complete sequence (2134bp, MW650743.1) that encodes for nonfunctional polyprotein gene from 851bp to 1512bp (Figure 5). It was first time discovered in Trinidad, West Indies in 2017/2018 from Haemagogus sp (Ali et al., 2021) . The closest relative is Atrato picorna-like virus 1 strain An 1771-1 genomic sequence with 89.58% Amino acid identity, sequence discovered from Anopheles sp. in Colombia in 2016 which was 9579bp long linear RNA . The ORF (220bp to 9486 bp) encoded for polyprotein. Any previous report of this virus in USA is unknown and undocumented in any mosquito species including Culex erraticus. The longert portion of the contigs 3,518 bp and 1,881 bp were unaligned to the best match so we reported the viral fragment as the Bunyavirales sp. isolate Cx.erraticus12 (PQ869759) and Qingnian mosquito virus isolate Cx.erraticus13 (PQ963481) respectively.

Bunyavirales like viral sequence

The RdRp of positive ss RNA of Qingnian mosquito virus isolate Cx.erraticus11 (PQ963480) (Figure 5) belonging to family Peribunyaviridae which 5078 bp long with 70.25% Amino acid identity to Qingnian Mosquito Virus strain YC179 RNA-dependent RNA polymerase gene, complete cds (7103bp, MW452282.1) collected from mosquito (Species unknown) from China in 2018. Its closest relative has not been reported in NCBI. Any previous report of this virus in USA is unknown and undocumented in any mosquito species including *Culex erraticus.*

Riboviria-like viral sequences related to unidentified family

Negevirus like viral sequence

Rinkaby virus isolate Cx.erraticus14 (PQ963485) (Figure 5) is 4264bp long with 29.41% viral coverage (10216bp to 14478bp) that encodes for fragment of hypothetical protein (ORF3: 8845bp to12480bp). It has 92.073% Amino acid identity to *Rinkaby virus isolate* OTU75 hypothetical protein, RNA-dependent RNA polymerase, and hypothetical protein genes, complete cds (14498 bp, MK440650.1) that was identified in *Culex sp*. in Sweden in 2007. The closest relative is *Riboviria* sp. strain YX683 hypothetical protein genes, complete cds (MW452294.1) discovered in mosquito in China in 2018 with 90.83% Amino acid identity. Any previous report of this virus in USA is unknown and undocumented in any mosquito species including *Culex erraticus.*

Other Riboviria like viral sequence similar to virus isolated in USA

There are two linear RNA virus like sequences that resembles the Hubei virga-like virus and Elemess virus that was isolated in *Culex erythrothorax* from USA in 2017 (Batson et al., 2021) (Figure 5).

Hubei virga-like virus 3 isolate Cx.erraticus1 (PQ963470) with 99.97% viral coverage (10867 bp) was discovered that has 99.26% Amino acid identity to Hubei virga-like virus 2 isolate CMS002_053a_PLCR genomic sequence (10870bp, MW435006.1). The genes that were discovered are Polyprotein (67 bp to 10046 bp) and capsid protein (10091bp to 10708 bp). This virus was first time isolated in Placer Valley, California from the mosquito collected in agriculture & residential habitat. The closest relative to this might be Hubei virga-like virus 2 isolate CMS001_024_ALCO genomic (10902 bp, MW434995.1) with 99.25% Amino acid identity collected in female, blood fed, collected in marsh habitat USA: Alameda County, California, USA in 2017(Batson et al., 2021). Any previous report of this virus is unknown and undocumented in *Culex erraticus.*

Two different segments of Elmesis virus (Positive ssRNA virus) were discovered. Elemess virus isolate Cx.erraticus10 (PQ963479) has 89.91% Amino acid identity to Elemess virus isolate CMS002_047g_WVAL major capsid protein gene, complete cds (3700bp, MW434732.1) from 5bp to 3689bp that was isolated from female, blood fed *Culex erythrothorax* collected in marsh habitat from Alameda County, California, USA in 2017 (Batson et al., 2021). The closest relative is Elemess virus isolate CMS001_049_ALCO major capsid protein gene, complete cds (3685bp, MW434734.1) with 89.89% Amino acid identity Alameda County, California, USA in 2017. Any previous report of this virus is unknown and undocumented in *Culex erraticus.*

Elemess virus isolate Cx.erraticus9 (PQ963478) has 89.44% Amino acid identity to Elemess virus isolate CMS001_020_ALCO RNA-dependent RNA polymerase gene, complete cds (3713 bp, MW434751.1) from 12bp to 3712bp that discovered from the RNA of female, blood fed *Culex erythrothorax* collected in marsh / wildlife refuge / park habitat from Alameda County, California, USA in 2017(Batson et al., 2021). The closest relative is Elemess virus isolate CMS001_035_ALCO RNA-dependent RNA polymerase gene, complete cds (3696 bp, MW434752.1) with 89.18% Amino acid identity discovered in Alameda County, California, USA in 2017. Any previous report of this virus is unknown and undocumented in *Culex erraticus.*

Other Riboviria like viral sequence similar to virus isolated in China

There are four linear RNA virus-like sequences that resemble the Zhejiang mosquito virus 1 strain, the Hubei narna-like virus 17 strain mosHB204971, Hubei mosquito virus 4 strain 3mos6213, and XiangYun narna-levi-like virus 8 isolate XY101736 (Figure 5).

Zheijang mosquito virus isolate Cx.erraticus17 (PQ963484) has 82.67% Amino acid identity to Zhejiang mosquito virus 1 strain mosZJ35497 hypothetical protein gene, complete cds (9558bp, NC033716.1) from 592 bp to 9558 bp that was isolated from Mosquitoes(species unidentified) in 2013 from China (Shi et al., 2016a). It encodes for hypothetical protein. The closest relative is Atrato picorna-like virus 1 strain An 1771-1 genomic sequence, sequence (9579bp, MN661035.1) that was isolated from Anopheles sp. from Colombia in 2016 with 88.56% AA identity. Any previous report of this virus in USA is unknown and undocumented in any mosquito species including *Culex erraticus.*

Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472) has 88.63% Amino acid identity to Hubei mosquito virus 4 strain 3mos6213 hypothetical protein 1, hypothetical protein 2, and hypothetical protein 3 genes, complete cds (4971bp, NC032231.1) from 1bp to 3643bp that was isolated from pool of mosquitoes from China in 2013 (Shi et al., 2016a). The closest relative is Hubei mosquito virus 4 strain mosHB232766 hypothetical protein 1, hypothetical protein 2, hypothetical protein 3, and hypothetical protein 4 genes, complete cds (4917 bp, KX883124.1) isolated from pool of mosquitoes from China in 2013.

The RdRp of XiangYun narna-levi-like virus 8 isolate Cx.erraticus16 (PQ963483) was discovered which is 2396 bp long with an alignment range 258bp to 2649bp with 90.44% Amino acid identity to XiangYun narna-levi-like virus 8 isolate XY101736 RNA-dependent RNA polymerase (2396 bp, OL700144.1) from *Culex pipiens* from Yunnan province, China in 2018 (Feng et al., 2022). There has been no reporting of this virus in USA. The closest relative is Serbia narna-like virus 1 isolate Mediterranean_2015 (OP264878.1) with 92.84% Amino acid identity that was isolated for the first time in Mediterranean region (Specific area: Not available) in 2015 from *Culex pipiens*. Any previous report of this virus in USA is unknown and undocumented in any mosquito species including *Culex erraticus.*

Lefavirales like viral sequences

The ORF23 that encode for ribonucleotide reductase of Spodoptera litura nucleopolyhedrovirus isolate Cx.erraticus5 (PQ963474) (Figure 5) is 2166 bp long with an alignment range 22104bp to 24240bp and 71.73% Amino acid identity to Spodoptera litura nucleopolyhedrovirus (137998bp, NC003102.1). It was reported in china from Spodoptera litura nucleopolyhedrovirus. It is not reported previously from Auburn, Alabama in *Culex erraticus*.

The ORF20 that encode for ribonucleotide reductase of Spodoptera littoralis nucleopolyhedrovirus isolate Cx.erraticus6 (PQ963475) (Figure 5) is 2154bp long with an alignment range 22226bp to 24349bp and 71.73% % Amino acid identity to Spodoptera littoralis nucleopolyhedrovirus (137998bp, NC038369.1). It was reported in China from Spodoptera litura nucleopolyhedrovirus. It is not reported previously from Auburn, Alabama in *Culex erraticus*.

Phylogenetic analysis of Merida virus isolate Cx.erraticus2 (PQ963471)

The phylogenetic analysis (Figure 6) of Merida virus isolate Cx.erraticus2 (PQ963471) suggests that the closest match is to the Culex rhabdovirus isolate Cx.erraticus7 (PQ963476) obtained from our own study and closest relative to Merida virus isolate MERD-Mex07 (NC 040599.1) that was isolated from *Culex quinquefasciatus*. The other isolate of meridavirus (MW434769.1) has been isolated from USA from California from *Culex tarsalis.* However, This is the first time for the isolation from Auburn, Alabama from *Culex erraticus.*

Discussion

At the onset of this study, the interest was determining what the diversity of viruses was in a species of mosquitos that could easily be found in an area of Auburn not frequently studied in The collection site was chosen based on the location close to a forested area, however, it contained fisheries ponds which would provide access to alternative hosts such as fish and birds. This site also offered easy access to human populations living in a planned neighborhood. The question asked given this diverse area was two-fold, what would be the most predominant population of mosquitoes and what viruses would they carry in reflection to that. If the population of mosquitoes were feeding on humans more often, the mosquito species would be expected to be one known to be problematic to humans such as *Aedes egyptii* and the expected viruses found would be those that infected humans. However, if the mosquito species found was a species not known to prefer humans, viruses found would be those that would not have humans as part of their host range but instead would be insect viruses or viruses of other species such as birds or reptiles. In this case, a diversity of viruses that are either novel to Alabama or this host would be found mostly due to the lack of study in this kind of environment.

Cx. erraticus is an abundant mosquito collected in spring till summer to early fall in Auburn, Alabama. Despite the high prevalence there is a huge gap in the understanding of the diversity of viruses. The potential insect-specific viruses (ISVs) has not been looked at because of limited infectivity outside their natural hosts(De Coninck et al., 2024). In our study we discovered total 15 different types of viruses belonging to five orders: Mononegavirales, Wolframvirales, Picornavirales, Bunyavirales, Lefavirales and Riboviria realm. This suggests that there is a significant diversity of viruses in the system of *Cx. erraticus.* Majority of viruses, approx. 50% of total viruses, were RNA viruses. This was expected because RNA viruses are abundant in nature because of its fast-paced mutation and ssRNA viruses mutate faster than dsRNA(Sanjuán & Domingo-Calap, 2016). We also found one Novel virus Qingnian mosquito virus isolate Cx.erraticus11 (PQ963480) with 70% Amino acid identity to the best match Qingnian Mosquito Virus strain YC179 RNA-dependent RNA polymerase gene (MW452282.1).

For RNA viruses, RNA dependent RNA polymerase (RdRp) is highly important for the replication of the genome and subsequent transcription (Venkataraman et al., 2018) which is the second step of the virus replication. The study discovered five viral RdRp of five different viruses. The polymerase catalytic motifs are the conserved portion of the RNA viruses (Jia & Gong, 2019). In our study, the three viruses (out of five) Serbia narna-like virus 2 isolate Cx.erraticus15 (PQ963482), XiangYun narna-levi-like virus 8 isolate Cx.erraticus16 (PQ963483) and Elemess virus isolate Cx.erraticus9 (PQ963478) have the conserved catalytic domains. This suggests that viruses with the conserved domain might be easier to discover in RNA-seq analysis. However, there is a chance that other viruses without the conserved domain would have been missed to be discovered, creating the detection bias in the RNA-seq analysis in virome study.

Our study is based on filtering out the viruses with the best match available in the NCBI universal and viral database. There is an under-representation of the insect viruses in the database because they are still being explored. Similarly, novel viruses might be present in *Cx. erraticus* but never been reported in NCBI universal and viral database. Because of this, there is a possibility that we might miss discovering the virus whose best match is not available in such database. For this the parallel techniques like Proteomics should also be considered. There might also be a potential for some biocomputational reasons involved in this bias, like presence of short reads, quality loss during analysis, etc. To address these different methods like long read sequencing, different methods of sequencing (Hi-seq) and use of virus enrichment techniques could be done. Additionally, it could be difficult to discover different DNA viruses from the insect genome, and as some might be integrated in the genome insects. Therefore, more techniques should be incorporated to address this detection bias.

Most of the best match of the viruses that we discovered before has been discovered in different species of the mosquitoes but were not reported in *Cx. erraticus* before. Out of fifteen viruses the best match of only two species of viruses (Hubei virga-like virus 3 isolate Cx.erraticus1 (PQ963470), and Elemess virus isolate Cx.erraticus10 (PQ963479) and Elemess virus isolate Cx.erraticus9 (PQ963478) were discovered in the USA before. Rest of the viruses are reported for the first time in *Cx. erraticus* through our study in USA. Since the best match of the viruses have been discovered previously from China (Shi et al., 2016b), Serbia (Stanojević et al., 2020b), Sweden, and the Mediterranean. This suggests the mosquito virome has broad geographical expansion among the mosquito species.

Previously there has been isolate of Meridavirus discovered from different mosquito species in USA. However, Merida virus isolate Cx.erraticus2 (PQ963471) is closest to the Meridavirus discovered from *Culex quinquefasciatus* in Mexico in 2007 (Charles et al., 2016). This suggests maybe the virus found here mutated or infected mosquito from Mexico found a way to arrive to Auburn, Alabama maybe by Hurricane. Most research have previously suggested the positive impact of hurricane to the insect community (Novais et al., 2018). But, in Auburn in Cx erraticus this is the first report of this virus. Meridavirus has been reported in other mosquito species like *Cx. tarsalis, Culex quinquefasciatus,* etc. This suggests that this virus might also be prevalent in other different species of mosquito species in USA, but the isolate might be different. Similarly, the study discovered longer contig for Merida virus isolate Cx.erraticus2 (PQ963471) i.e. 11766bp out of 11798bp of the full virus and for Hubei virga-like virus 3 isolate Cx.erraticus1 (PQ963470) i.e. 10813bp out of 10870bp. This suggests that these two viruses might be the most prevalent and be in the most active stage in *Cx. erraticus* during the sampling.

The approach of using the dual bioinformatic methods was supportive of the comprehensive identification of viral diversity in *Cx. erraticus*. Both methods had three viruses (Merida virus isolate Cx.erraticus 2 (PQ963471), Hubei mosquito virus 5 isolate Cx.erraticus3 (PQ963472), Zheijang mosquito virus isolate Cx. erraticus 17 (PQ963484)) as a common virus. From the second method, we were able to identify the two viruses: Spodoptera litura nucleopolyhedrovirus isolate Cx.erraticus5 (PQ963474) and Spodoptera littoralis nucleopolyhedrovirus isolate Cx.erraticus6 (PQ963475). The prediction is that these viruses might be incorporated into the genome of the *Cx. erraticus*. Genome integration has been found to be a mandatory step in some groups of viruses like retroviruses during replication (Hindmarsh & Leis, 1999). Also, if these viruses integrate in the germline they can be inherited by the progeny (Blair et al., 2019). The literature suggests that non-retroviruses are also incorporated into the genome of the insects. (Blair et al., 2019). In Hematophagous insects like mosquitoes different viruses like Flavivirus (Crochu et al., 2004)(Roiz et al., 2009) (Sánchez-Vargas et al., 2009)(Rizzo et al., 2014)(Vázquez et al., 2012)(Katzourakis & Gifford, 2010), (Whitfield et al., 2017), Reoviruses, Rhabdovirus (Katzourakis & Gifford, 2010) (Whitfield et al., 2017) Chuviruses, vigraviruses (Whitfield et al., 2017) are found to show the genome integration. Both of these viruses also belong to the Baculoviridae family that have the potential to cause mortality in mosquito larvae (Agboli et al., 2019).So, these viruses might have biocontrol potential and biotechnological implications for the control of insect populations (Williams et al., 2017). A clear understanding of these two viruses requires further investigation. However, the kraken2 (Version 2.1.3) report on the total contigs from method 1 also included two viruses that were found from method 2 such as Spodoptera litura nucleopolyhedrovirus and Zhejiang mosquito virus 1. This suggests that these viruses might be present in the contigs found from method 1 and might have been annotated differently or removed in the final filtering. Additionally, there is the potential of the presence of more viruses which were not found from both methods based on this report. More research is needed to be done to address this.

The E.W. Shell Fisheries center is a 1600-acre research station consisting of infrastructure and natural habitat. Man-made ponds facilitate studies utilizing diverse species of catfish and trout. The fisheries also provide a habitat for a variety of wildlife including Bald Eagles, Belted Kingfishers, blue herons, wading birds, songbirds, reptiles, and amphibians. From our study we expected to find different viruses infecting the mosquito, as well as some bird and fish infecting viruses. Since *Cx. erraticus* does not feed on human blood we did not expect to find any human viruses. We did not find the fish viruses. However, the best match of Ecclesville picorna-like virus isolate Cx.erraticus8 (PQ963477) was discovered from the mosquito *Haemagogus sp* that was sampled in Caroni Bird Sanctuary located in National Park in Trinidad and Tobago (Ali et al., 2021). This virus belongs to the family *Picornavirales*. This family comprises viruses that are non-enveloped, positive-sense and non-segmented RNA(6.7–10.1 kb). This family is important because it contains the viruses that infects many eukaryotes like mammals, fish, arthropods, plants, algae (Zell et al., 2017). This suggests this virus might be potential bird or fish infecting viruses. More research is needed to support this. This result also suggests that sampling strategies is crucial in virome representation. This suggests two things; first this virus might have broad host range including *Cx. erraticus* and it might also be potential bird infecting virus that might have been retained in the body of *Cx. erraticus* while feeding on that birds that comes to feed on the fish in the ecosystem of the sampling site.

Although several viruses were identified from *Cx erraticus*, the knowledge on the ecological and evolutionary roles of these viruses within mosquito populations and their potential to be transmitted to other hosts is limited. However, bioinformatics approach with combined analysis methods contributes significantly to the understanding on diversity of virus associated with mosquito.

Conclusion

The Mi-Seq of the most abundant mosquito *Cx. erraticus* in Auburn, Alabama in 2022 suggests the presence of fifteen different viruses. Merida virus isolate Cx.erraticus2 (PQ963471) is newly discovered in the Auburn, Alabama from this insect species. The finding emphasizes the importance of engaging multiple bioinformatic analysis approach might be a better approach to incorporate the maximum diversity of viruses from complex dataset.

References

Agboli, E., Leggewie, M., Altinli, M., & Schnettler, E. (2019). Mosquito-Specific Viruses—Transmission and Interaction. *Viruses*, *11*(9), 873. https://doi.org/10.3390/v11090873

Ali, R., Jayaraj, J., Mohammed, A., Chinnaraja, C., Carrington, C. V. F., Severson, D. W., & Ramsubhag, A. (2021). Characterization of the virome associated with Haemagogus mosquitoes in Trinidad, West Indies. *Scientific Reports*, *11*(1), 16584. https://doi.org/10.1038/s41598-021-95842-6

Batson, J., Dudas, G., Haas-Stapleton, E., Kistler, A. L., Li, L. M., Logan, P., Ratnasiri, K., & Retallack, H. (2021). Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay. *eLife*, *10*, e68353. https://doi.org/10.7554/eLife.68353

Blair, C. D., Olson, K. E., & Bonizzoni, M. (2019). The Widespread Occurrence and Potential Biological Roles of Endogenous Viral Elements in Insect Genomes. In *Insect Molecular Virology: Advances and Emerging Trends*. Caister Academic Press. https://doi.org/10.21775/9781912530083.02

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009a). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009b). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Charles, J., Firth, A. E., Loroño-Pino, M. A., Garcia-Rejon, J. E., Farfan-Ale, J. A., Lipkin, W. I., Blitvich, B. J., & Briese, T. (2016). Merida virus, a putative novel rhabdovirus discovered in Culex and Ochlerotatus spp. Mosquitoes in the Yucatan Peninsula of Mexico. *Journal of General Virology*, *97*(4), 977–987. https://doi.org/10.1099/jgv.0.000424

Cock, P. J. A., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015). NCBI BLAST+ integrated into Galaxy. *Gigascience*, *4*(1), s13742-015-0080–0087. https://doi.org/10.1186/s13742-015-0080-7

Crochu, S., Cook, S., Attoui, H., Charrel, R. N., De Chesse, R., Belhouchet, M., Lemasson, J.-J., De Micco, P., & De Lamballerie, X. (2004). Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. Mosquitoes. *Journal of General Virology*, *85*(7), 1971–1980. https://doi.org/10.1099/vir.0.79850-0

De Coninck, L., Soto, A., Wang, L., De Wolf, K., Smitz, N., Deblauwe, I., Mbigha Donfack, K. C., Müller, R., Delang, L., & Matthijnssens, J. (2024). Lack of abundant core virome in *Culex* mosquitoes from a temperate climate region despite a mosquito species-specific virome. *mSystems*, *9*(6), e00012-24. https://doi.org/10.1128/msystems.00012-24

Deng, X., Naccache, S. N., Ng, T., Federman, S., Li, L., Chiu, C. Y., & Delwart, E. L. (2015). An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. *Nucleic Acids Research*, *43*(7), e46–e46. https://doi.org/10.1093/nar/gkv002

Deng, Z., & Delwart, E. (2021). ContigExtender: A new approach to improving de novo sequence assembly for viral metagenomics data. *BMC Bioinformatics*, *22*(1), 119. https://doi.org/10.1186/s12859-021-04038-2

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Estep, L. K., McClure, C. J. W., Burkett-Cadena, N. D., Hassan, H. K., Hicks, T. L., Unnasch, T. R., & Hill, G. E. (2011). A multi-year study of mosquito feeding patterns on avian hosts in a southeastern focus of eastern equine encephalitis virus. *The American Journal of Tropical Medicine and Hygiene*, *84*(5), 718–726. https://doi.org/10.4269/ajtmh.2011.10-0586

Feng, Y., Gou, Q., Yang, W., Wu, W., Wang, J., Holmes, E. C., Liang, G., & Shi, M. (2022). A time-series meta-transcriptomic analysis reveals the seasonal, host, and gender structure of mosquito viromes. *Virus Evolution*, *8*(1), veac006. https://doi.org/10.1093/ve/veac006

Gómez, M., Martínez, D., Páez-Triana, L., Luna, N., De Las Salas, J. L., Hernández, C., Flórez, A. Z., Muñoz, M., & Ramírez, J. D. (2023). Characterizing viral species in mosquitoes (Culicidae) in the Colombian Orinoco: Insights from a preliminary metagenomic study. *Scientific Reports*, *13*(1), 22081. https://doi.org/10.1038/s41598-023-49232-9

Hindmarsh, P., & Leis, J. (1999). Retroviral DNA Integration. *Microbiology and Molecular Biology Reviews*, *63*(4), 836–843. https://doi.org/10.1128/MMBR.63.4.836-843.1999

Jia, H., & Gong, P. (2019). A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. *Frontiers in Microbiology*, *10*, 1945. https://doi.org/10.3389/fmicb.2019.01945

Katzourakis, A., & Gifford, R. J. (2010). Endogenous Viral Elements in Animal Genomes. *PLoS Genetics*, *6*(11), e1001191. https://doi.org/10.1371/journal.pgen.1001191

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009a). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009b). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, *52*(W1), W78–W82. https://doi.org/10.1093/nar/gkae268

Morse, W., Izenour, K., McKenzie, B., Lessard, S., & Zohdy, S. (2019). Perceptions and practices of mosquito-borne diseases in Alabama – is concern where it should be? *BMC Public Health*, *19*(1), 987. https://doi.org/10.1186/s12889-019-7308-x

NCBI Virus. (n.d.). Retrieved March 6, 2025, from https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/

Nebbak, A., Monteil-Bouchard, S., Berenger, J.-M., Almeras, L., Parola, P., & Desnues, C. (2021). Virome Diversity among Mosquito Populations in a Sub-Urban Region of Marseille, France. *Viruses*, *13*(5), 768. https://doi.org/10.3390/v13050768

Novais, S., Macedo-Reis, L. E., Cristobal-Peréz, E. J., Sánchez-Montoya, G., Janda, M., Neves, F., & Quesada, M. (2018). Positive effects of the catastrophic Hurricane Patricia on insect communities. *Scientific Reports*, *8*(1), 15042. https://doi.org/10.1038/s41598-018-33210-7

Rizzo, F., Cerutti, F., Ballardini, M., Mosca, A., Vitale, N., Radaelli, M. C., Desiato, R., Prearo, M., Pautasso, A., Casalone, C., Acutis, P., Peletto, S., & Mandola, M. L. (2014). Molecular characterization of flaviviruses from field-collected mosquitoes in northwestern Italy, 2011–2012. *Parasites & Vectors*, *7*(1), 395. https://doi.org/10.1186/1756-3305-7-395

Roiz, D., Vázquez, A., Seco, M., Tenorio, A., & Rizzoli, A. (2009). Detection of novel insect flavivirus sequences integrated in Aedes albopictus (Diptera: Culicidae) in Northern Italy. *Virology Journal*, *6*(1), 93. https://doi.org/10.1186/1743-422X-6-93

Sadeghi, M., Altan, E., Deng, X., Barker, C. M., Fang, Y., Coffey, L. L., & Delwart, E. (2018). Virome of > 12 thousand Culex mosquitoes from throughout California. *Virology, 523,* 74–88. https://doi.org/10.1016/j.virol.2018.07.029

Sánchez-Vargas, I., Scott, J. C., Poole-Smith, B. K., Franz, A. W. E., Barbosa-Solomieu, V., Wilusz, J., Olson, K. E., & Blair, C. D. (2009). Dengue Virus Type 2 Infections of Aedes aegypti Are Modulated by the Mosquito's RNA Interference Pathway. *PLoS Pathogens*, *5*(2), e1000299. https://doi.org/10.1371/journal.ppat.1000299

Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. *Cellular and Molecular Life Sciences*, *73*(23), 4433–4448. https://doi.org/10.1007/s00018-016-2299-6

Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden, J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E. C., & Zhang, Y.-Z. (2016a). Redefining the invertebrate RNA virosphere. *Nature*, *540*(7634), 539–543. https://doi.org/10.1038/nature20167

Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden, J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E. C., & Zhang, Y.-Z. (2016b). Redefining the invertebrate RNA virosphere. *Nature*, *540*(7634), 539–543. https://doi.org/10.1038/nature20167

Stanojević, M., Li, K., Stamenković, G., Ilić, B., Paunović, M., Pešić, B., Maslovara, I. Đ., Šiljić, M., Ćirković, V., & Zhang, Y. (2020a). Depicting the RNA Virome of Hematophagous Arthropods from Belgrade, Serbia. *Viruses*, *12*(9), 975. https://doi.org/10.3390/v12090975

Stanojević, M., Li, K., Stamenković, G., Ilić, B., Paunović, M., Pešić, B., Maslovara, I. Đ., Šiljić, M., Ćirković, V., & Zhang, Y. (2020b). Depicting the RNA Virome of Hematophagous Arthropods from Belgrade, Serbia. *Viruses*, *12*(9), 975. https://doi.org/10.3390/v12090975

Vázquez, A., Sánchez-Seco, M.-P., Palacios, G., Molero, F., Reyes, N., Ruiz, S., Aranda, C., Marqués, E., Escosa, R., Moreno, J., Figuerola, J., & Tenorio, A. (2012). Novel Flaviviruses Detected in Different Species of Mosquitoes in Spain. *Vector-Borne and Zoonotic Diseases*, *12*(3), 223–229. https://doi.org/10.1089/vbz.2011.0687

Venkataraman, S., Prasad, B., & Selvarajan, R. (2018). RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. *Viruses*, *10*(2), 76. https://doi.org/10.3390/v10020076

Whitfield, Z. J., Dolan, P. T., Kunitomi, M., Tassetto, M., Seetin, M. G., Oh, S., Heiner, C., Paxinos, E., & Andino, R. (2017). The Diversity, Structure, and Function of Heritable Adaptive Immunity Sequences in the Aedes aegypti Genome. *Current Biology*, *27*(22), 3511-3519.e7. https://doi.org/10.1016/j.cub.2017.09.067

Williams, T., Virto, C., Murillo, R., & Caballero, P. (2017). Covert Infection of Insects by Baculoviruses. *Frontiers in Microbiology*, *8*, 1337. https://doi.org/10.3389/fmicb.2017.01337

Wood, D. E., & Salzberg, S. L. (2014). Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Zell, R., Delwart, E., Gorbalenya, A. E., Hovi, T., King, A. M. Q., Knowles, N. J., Lindberg, A. M., Pallansch, M. A., Palmenberg, A. C., Reuter, G., Simmonds, P., Skern, T., Stanway, G., Yamashita, T., & ICTV Report Consortium. (2017). ICTV Virus Taxonomy Profile: Picornaviridae. *Journal of General Virology*, *98*(10), 2421–2422. https://doi.org/10.1099/jgv.0.000911

Tables

Table 1. Sequence statistics of the reads generated by Mi-Seq of RNA of *Culex erraticus* collected in E.W. Shell Fisheries between late August to Late October in 2022

	Method1	Method2:
Number of total reads	158709252	
Total number of forward and reverse reads	79354626 each	
Length range of the read	35-150 bp	
GC%	45	
Sequences matched to the reference genome %	16.72	
Sequences unmatched to the reference genome %	83.28	
Number of trinity transcripts	1006209	
GC% of Trinity transcripts	41.27	
Average contig length (bp)	580.25	
Average assembled base for the contig	583848302	
Number of contigs of 0-400bp	128835	
Number of contigs of 400-1000bp	326633	

Number of contigs of >1000bp	63986316
Number of viral contigs	742
Number of virus contig <400bp	955 hits (387 contigs)
Number of virus contig 400bp to 1000bp	590 hits (255 contigs)
Number of virus >1000bp	367hits (100 viral contigs)

Method1: Methods with mapping to the reference genome and blast to the NCBI database

Method2: Methods without mapping to the reference genome and blast to the viral database

Table 2. Diversity of viruses. The table shows the diversity of viruses from RNA-seq analysis of *Culex erraticus* collected in E.W. Shell Fisheries between late August to Late October in 2022; RNA: Ribo-Nucleic Acid by dual bio-informatic methods

	Method 1 (With Mapping with reference genome (GCF_015732765.1)				**Method 2 (Without Mapping)**			
Proposed Name (Assigned Accession Number)	**Best Match (Best Match Accession Number)**	**Best Match length (bp)**	**Alignment length (bp)**	**Amino acid identity (%)**	**Best Match (Best Match Accession Number)**	**Best Match length (bp)**	**Alignment length (bp)**	**Amino acid identity (%)**
Merida virus isolate Cx.erraticus2 (PQ963471)	Merida virus isolate MERD-Mex07 (NC_040599.1)	11798	11766*	98.742	Merida virus isolate MERD-Mex07 (NC040599.1)	11798	11766	
Hubei mosquito virus 5 isolate Cx.erraticus3	Hubei mosquito virus 4 (NC032231.1)	4971	265		Hubei mosquito virus 4 (NC032231.1)	4971	3734*	89.45
Zheijang mosquito virus isolate Cx.erraticus17 (PQ963484)	Atrato picorna-like virus 1 strain An 1771-1 genomic sequence, sequence (MN661035.1)	9579	663	89.577	Zhejiang mosquito virus 1 (NC033716.1)	*9558*	9012*	77.16
Spodoptera litura nucleopolyhedrovirus isolate Cx.erraticus5 (PQ963474)					Spodoptera litura nucleopolyhedrovirus (NC003102.1)	139342	2166*	71.02
Spodoptera littoralis nucleopolyhedrovirus isolate Cx.erraticus6 (PQ963475)					Spodoptera littoralis nucleopolyhedrovirus (NC038369.1)	137998	2154*	71.73
Hubei virga-like virus 3 isolate Cx.erraticus1 (PQ963470)	Hubei virga-like virus 2 isolate CMS002_053a_PLCR genomic sequence	10870	10813*	99.27				
Culex rhabdovirus isolate Cx.erraticus7	Culex rhabdovirus strain CRV/Kern	11824	3666*	98.636				
Ecclesville picorna-like virus isolate Cx.erraticus8 (PQ963477)	Ecclesville picorna-like virus isolate TriniPa-Ecclesville-E04 nonfunctional polyprotein gene	2134	662*	90.181				
Elemess virus isolate Cx.erraticus9 (PQ963478)	Elemess virus isolate CMS001_020_ALCO RNA-dependent RNA polymerase gene	3711	3701*	89.111				
Elemess virus isolate Cx.erraticus10 (PQ963479)	Elemess virus isolate CMS002_047g_WVAL major capsid protein gene (MW434732.1)	3663	3663*	89.708				
Qingnian mosquito virus isolate Cx.erraticus11 (PQ963480)	Qingnian Mosquito Virus strain YC179 RNA-dependent RNA polymerase gene	7103	1583*	70.246				
Rinkaby virus isolate Cx.erraticus14 (PQ963485)	Rinkaby virus isolate OTU75 hypothetical protein, RNA-dependent RNA polymerase, and hypothetical protein genes (MK440650.1)	14498	4264*	88.115				
Serbia narna-like virus 2 isolate Cx.erraticus15 (PQ963482)	Serbia narna-like virus 1 isolate 85061 RNA-dependent RNA polymerase gene	3280	2751*	94.002				
XiangYun narna-levi-like virus 8 isolate Cx.erraticus16 (PQ963483)	XiangYun narna-levi-like virus 8 isolate XY101736 RNA-dependent RNA polymerase	2649	2652*	90.611				
Bunyavirales sp. isolate Cx.erraticus12 (PQ869759)								
Qingnian mosquito virus isolate Cx.erraticus13 (								

Table 3. Mosquito Collection at the Fisheries. This Table depicts the mosquito collections taken from E.W. Shell Fisheries between late August to Late October in 2022 (Information received from Natalie Wong)

Sampling Date	Species of Mosquito collected	Male/Female	Number
8/23/22	*Culex erraticus*	Female	21
8/31/22	*Culex erraticus*	Female	12
9/1/22	*Anopheles Crucians*	Female	1
9/7/22	*Culex erraticus*	Female	8
9/7/22	*Culex erraticus*	Female	5
9/7/22	*Psorophora spp.*	Female	5
9/21/22	*Culex erraticus*	Female	80
9/21/22	*Culex erraticus*	Female	4
10/12/22	*Culex erraticus*	Female	20
10/26/22	*Culex erraticus*	Female	80
10/26/22	*Psorophora spp.*	Female	1
Total			237

Table 4. Kracken2 statistics of the reads generated by Mi-Seq of RNA of *Culex erraticus* collected in E.W. Shell Fisheries between late August to Late October in 2022

Taxonomical id	Kraken2: Classification
1.Escherichia_phage_vB_EcoM_VR20	1
2.Escherichia_phage_ECML-4	1
3.Erwinia_phage_PhiEaH1	3
4.Bacillus_phage_Shbh1	4
5.Gordonia_phage_Yvonnetastic	1
6.unclassified_Wphvirus	1
7.Acanthamoeba_polyphaga_moumouvirus	2
8.Diadromus_pulchellus_ascovirus_4a	12
9.Viruses	2
10.Caudovirales	1
11.Mimiviridae	1
12.Alphabaculovirus	1
13.Pandoravirus	5
14.Slopekvirus	1
15.Bocaparvovirus	4

16.Lymantria_dispar_multiple_nucleopolyhedrovirus	1
17.Spodoptera_litura_nucleopolyhedrovirus	5
18.Leucania_separata_nucleopolyhedrovirus	9
19.Choristoneura_fumiferana_granulovirus	17
20.Hyposoter_fugitivus_ichnovirus	10
21.Glypta_fumiferanae_ichnovirus	1
22.Cotesia_congregata_bracovirus	1
23.Pandoravirus_salinus	2
24.Pandoravirus_inopinatum	7
25.Pandoravirus_macleodensis	7
26.Pandoravirus_quercus	2
27.Phormidium_phage_MIS-PhV1B	1
28.Thermoproteus_tenax_virus_1	1
29.Hubei_mosquito_virus_4	12
30.Hubei_narna-like_virus_17	1
31.Hubei_virga-like_virus_2	4
32.Wenzhou_tombus-like_virus_11	1
33.Zhejiang_mosquito_virus_1	4

34.Espirito_Santo_virus	2
35.Perinet_vesiculovirus	1
36.Merida_virus	103
37.Enterovirus_sp.	1
38.Dioscorea_mosaic_associated_virus	1
39.Fruit_bat_alphaherpesvirus_1	1
40.Cyanophage_S-RIM32	1
41.Bacillus_virus_G	4
42.Escherichia_phage_HK639	1
43.Escherichia_virus_SSL2009a	1
44.Streptococcus_phage_20617	1
45.Chrysochromulina_ericina_virus	1
46.Megavirus_chiliensis	4
47.Mimivirus_terra2	1
48.Marseillevirus_marseillevirus	12
49.Skunkpox_virus	1

Figures

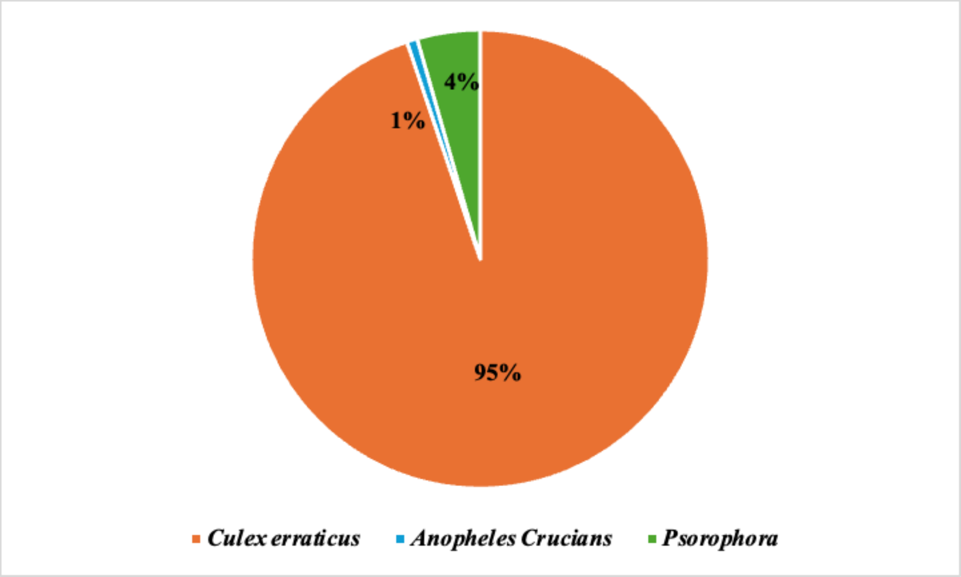


Figure 1. Sampling of Mosquito. The pie-chart shows the abundance of mosquito species during sampling in E.W. Shell Fisheries between late August to Late October in 2022

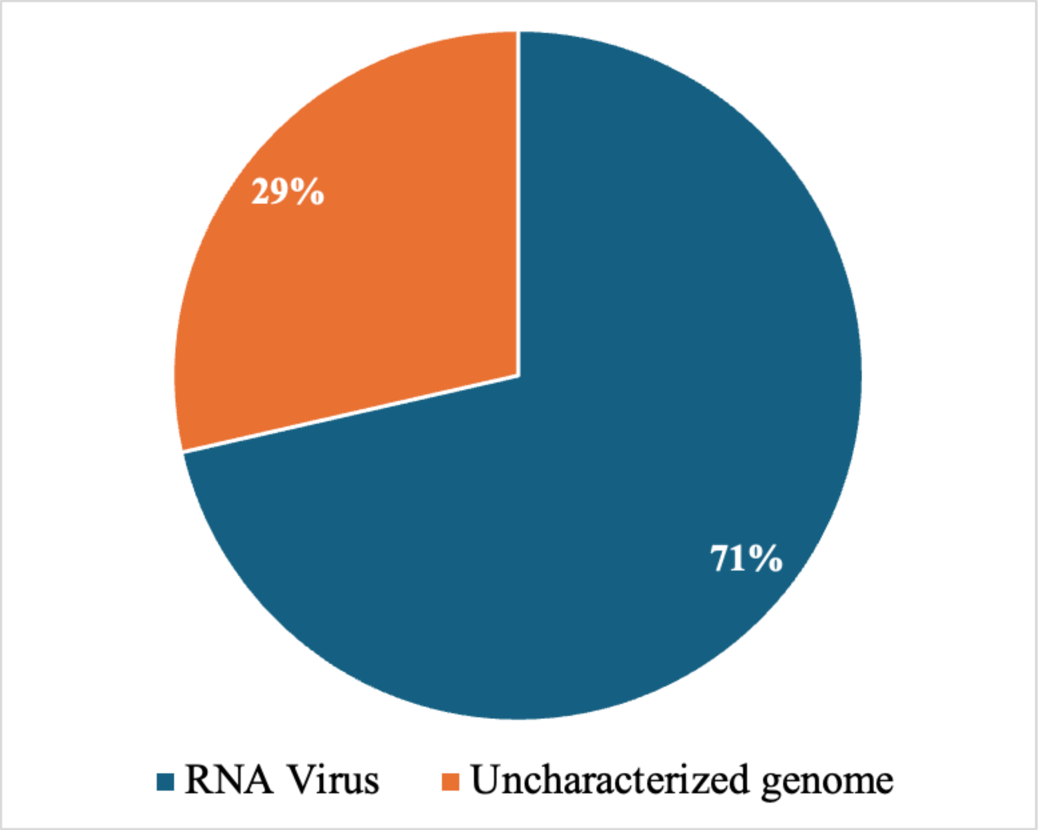


Figure 2. Genome of the viruses. The pie-chart represents the types of genomes (DNA or RNA or uncharacterized) of the viruses discovered by RNA-seq analysis of *Culex erraticus* collected in E.W. Shell Fisheries between late August to Late October in 2022; RNA: Ribo-Nucleic Acid

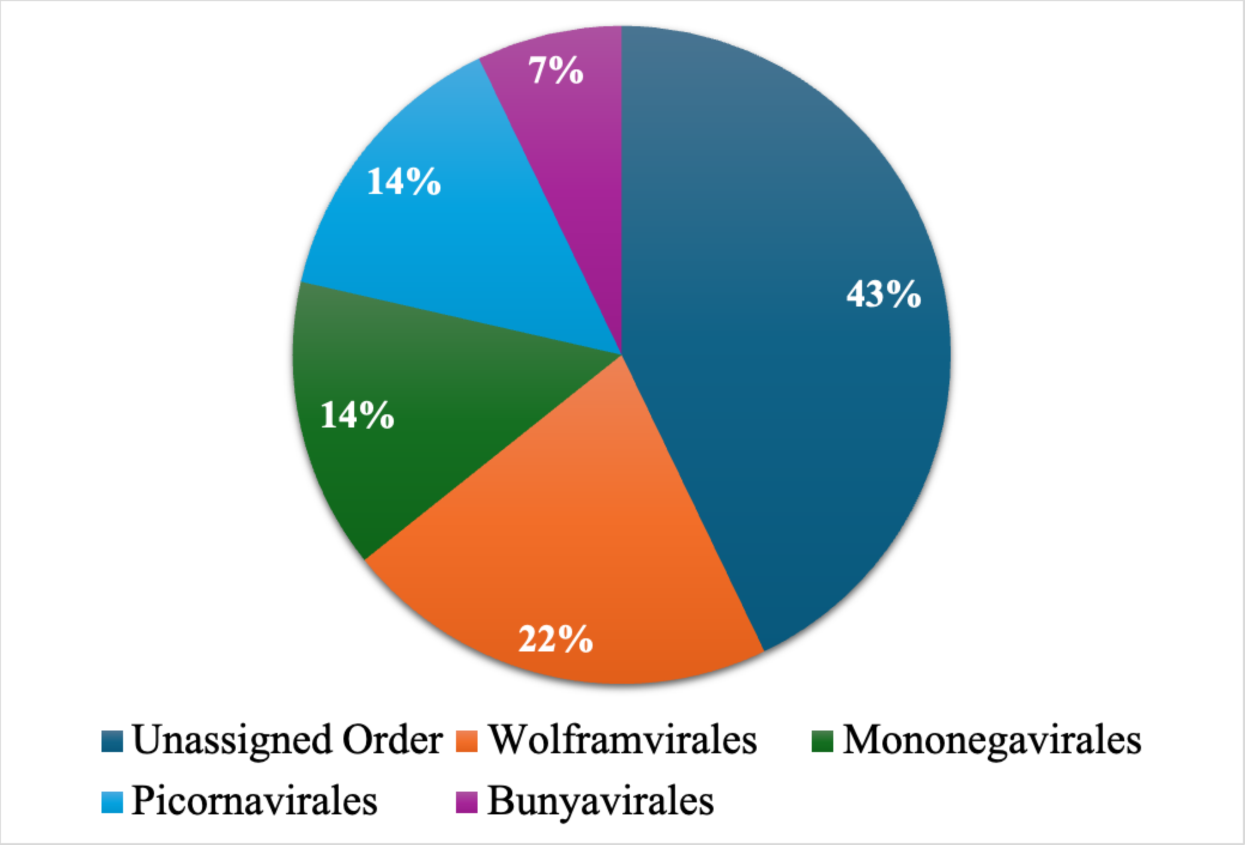


Figure 3. Statistics of virus belonging to different orders. The pie-chart represents the virus diversity at order level that were discovered by RNA-seq analysis of *Culex erraticus* collected in E.W. Shell Fisheries between late August to Late October in 2022

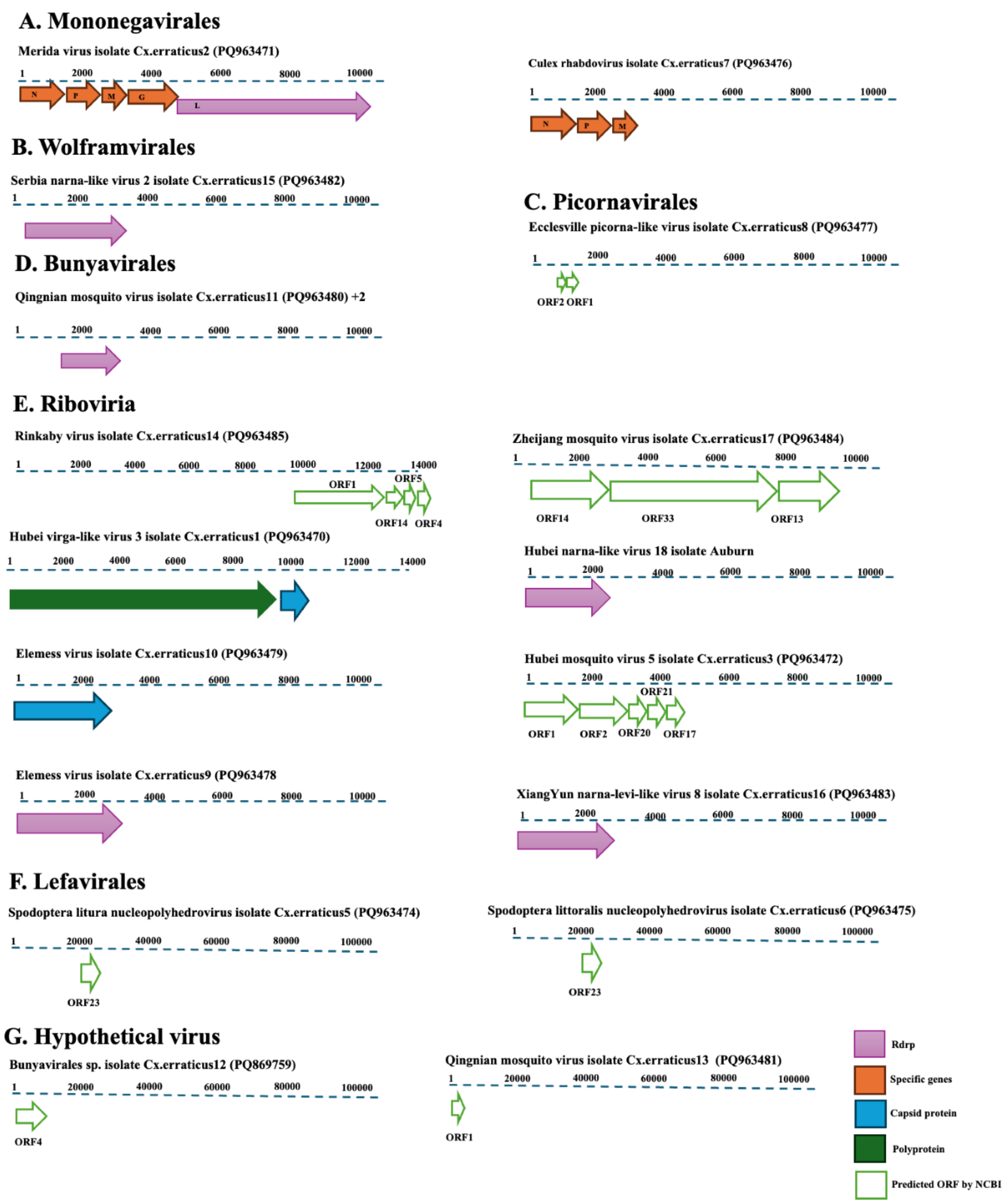


Figure 4. Graphical representation viruses from Mi-Seq of Mosquito(*Culex erraticus*) collected in Auburn, Alabama 2022; Viruses belonging to (A) Mononegavirales (B) Wolframvirales; (C) Picornavirales; (D)Bunyavirales; (E) Patatavirales; (F) Riboviria realm (G). Lefavirales (H);Hypothetical viruses; ORF: Open reading frames; RdRP: RNA-dependent RNA polymerase; NCBI: National Center for Biotechnology Information

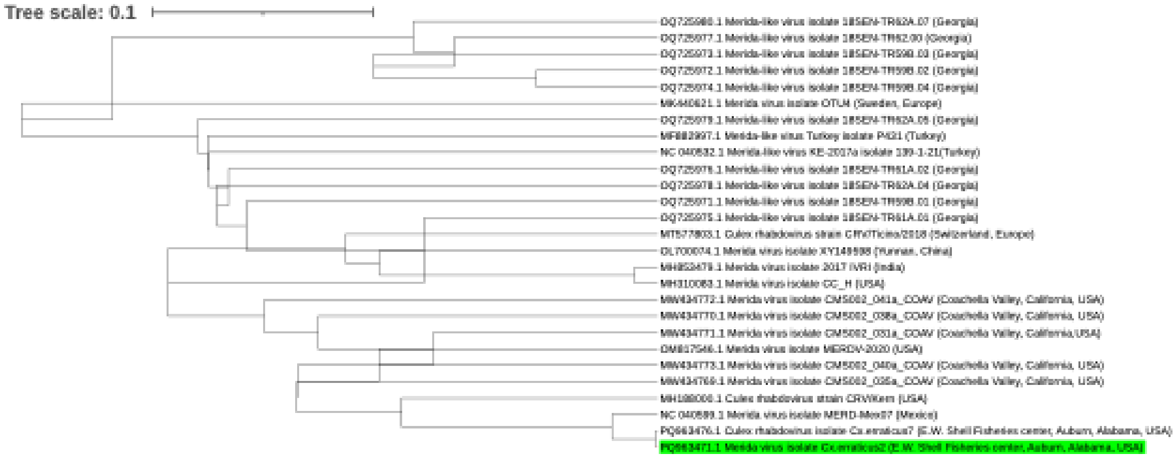


Figure 5. Phylogenetic analysis of Merida virus isolate Cx.erraticus2 (PQ963471) (Highlighted in Green) assembled from Mi-Seq analysis of RNA of *Culex erraticus* collected in E.W. Shell Fisheries between late August to Late October in 2022