POPULATION DIVERSITY AND STRUCTURE OF PHAKOPSORA PACHYRHIZI IN THE SOUTHEASTERN UNITED STATES

by

JOHN CLARY

(Under the Direction of Shavannor M. Smith)

ABSTRACT

Phakopsora pachyrhizi causes Soybean Rust (SBR) on soybeans. SBR is a devastating foliar diseases of soybeans, causing yield losses of up to 50%. *P. pachyrhizi* could overcome common fungicides and adapt to the environment of the continental US. To better understand the behavior of *P. pachyrhizi* and determine effective management practices, this thesis investigates population structure and diversity of *P. pachyrhizi* in the Southeastern US. Genotype by sequencing was used to identify single nucleotide polymorphisms in 49 field isolates of *P. pachyrhizi* collected from four states over ten years (2008-2017). Two clusters of isolates were identified using K-means clustering, discriminant analysis of principal components, and a neighbor-joining tree. These clusters suggested two introductions of *P. pachyrhizi* into the US and two *P. pachyrhizi* genotypes are present in the Southeastern US. *P. pachyrhizi* isolates

demonstrated moderate levels of genetic diversity, with no indications of major adaptations to the Southeastern US.

INDEX WORDS: Plant pathology, plant pathogen interactions, population genetics, population structure, *Phakopsora pachyrhizi*, Asian Soybean Rust, genetic diversity, SBR, Genotype by Sequencing

POPULATION DIVERSITY AND STRUCTURE OF PHAKOPSORA PACHYRHIZI IN THE SOUTHEASTERN UNITED STATES

by

JOHN CLARY

BS, Middle Georgia State University, 2018

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2021

POPULATION DIVERSITY AND STRUCTURE OF PHAKOPSORA PACHYRHIZI IN THE SOUTHEASTERN UNITED STATES

by

JOHN CLARY

Major Professor:	Shavannor Smith
Committee:	James Buck
	James Leebens-Mack
	Chang-Hyun Khang

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2021

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

Introduction 1

Literature Review 3

REFERENCES 17

2 POPULATION DIVERSITY AND STRUCTURE OF PHAKOPSORA PACHYRHIZI IN THE SOUTHEASTERN UNITED STATES 29

Abstract 30

Introduction 30

Materials and Methods 33

Results and Discussion 38

Conclusion 44

References 46

3 SUMMARY 61

REFERENCES 64

APPENDICES

A Quality and quantity of 49 *P. pachyrhizi* southeastern field isolates used for genotype-by-sequencing 65

B Sequencing mean quality scores by position for 49 *P. pachyrhizi* Southeastern field isolates 68

C Per sequence quality scores on y- axis and quality scores on x- axis for the 49 *P. pachyrhizi* Southeastern field isolates 69

D Average allelic heterozygosity proportions in each locality for the 49 *P. pachyrhizi* Southeastern isolates 70

E Heterozygosity averages for each locality and nei's genetic distance over the 40 *P. pachyrhizi* samples in the larger cluster 71

F Heterozygosity averages for each locality and nei's genetic distance over the 9 *P. pachyrhizi* isolates in the smaller cluster 72

G PCoA of 49 *P. pachyrhizi* southeastern isolates using Bray-Curtis dissimilarity matrix 73

LIST OF TABLES

Page

Table 2.1: *Phakopsora pachyrhizi* Isolate Collection.....51

LIST OF FIGURES

Page

Figure 2.1: Schematic representation of Fast-GBS v2 pipeline.....53

Figure 2.2: Genotype accumulation curve for *P. pachyrhizi*.....54

Figure 2.3: Heterozygosity averages for each locality and nei's genetic distance over all *P. pachyrhizi* isolates.....55

Figure 2.4: Genetic distances of *P. pachyrhizi* isolates in the Southeastern United States.....56

Figure 2.5: K-means clustering and DAPC of SBR in the Southeastern United States.....58

Figure 2.5: PCoA of 49 Southeastern United States isolates of soybean rust.....60

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

Soybeans are a rich resource for human and animal feed. As a result, soybeans are important for global food security, and are planted over large acreages throughout the world. Several plant pathogens affect soybean production. One of the most problematic is soybean rust (SBR), caused by the fungus *P. pachyrhizi*, which can cause yield losses of up to 50% in untreated soybean fields (Hartman, et al., 2005). SBR is managed primarily by unsustainable and expensive fungicide inputs throughout each season, causing environmental damage and increasing pathogen fungicide tolerance (Langenback, 2016). Therefore, more sustainable methods are needed to manage this disease, such as improving soybean cultivar resistance, which would increase production efficiency and end-use quality of soybeans.

To guide decisions regarding breeding for SBR resistance, it is important to understand the pathogen's population structure and genetic diversity. There has been some research analyzing the genetic diversity of *P. pachyrhizi* (Twizeyimana et al., 2011, Freire *et al*., 2008, 2012; Zhang *et al*., 2012) but not recently and no large-scale population genetics studies have been conducted for the southeastern region of the United States. These studies have also focused on individual genes or a small portion of the genome, such as the internal transcribed spacer (ITS) or ADP ribosylation factor regions, without accounting for the diversity of the entire genome or enough to assess genotypes and make conclusions on population genetics (Goellner *et*

al., 2010; Akamatsu *et al.,* 2013). With the availability of three annotated draft genomes of *P. pachyrhizi* (references), new genome wide analyses are imperative to a more comprehensive understanding of the pathogen. The major objective of this project is to better understand the population structure of *P. pachyrhizi,* as well as its evolution since its introduction into the US, therefore providing insight into efficient breeding strategies needed for broad-spectrum and durable resistance. This research could also aid in understanding selection pressures imposed on *P. pachyrhizi* in the US, the pathogen's divergence from its parent population, and insight into new introductions of pathogen isolates into the US. Additionally, data on the specific genotypes of SBR within and among different locations in the southeastern US would indicate where specific R-genes may be most durable. Analyzing evolutionary processes and population genomics of *P. pachyrhizi* is also valuable in identifying effector candidates (Moller and Stukenbrock, 2017). It is possible to measure selective pressure by components of the plant immune system on effector genes in the genome, and studies of this nature have been done extensively (Raffaele and Kamoun, 2012; Stukenbrock, 2013; Persoons *et al.*, 2014; Upadhyaya *et al.*, 2014). These studies are made possible due to the relatively low cost of sequencing a large volume of isolates, meaning it is likely that population genomics and eventually GWAS will be able to unveil novel determinants of pathogenicity.

Project Objectives and hypotheses

Objective 1: *Characterize the genetic diversity of Phakopsora pachyrhizi in the southeastern United States using a metagenomics approach.*

Hypothesis – *P. pachyrhizi isolates from the southeastern US are divergent and rich in genetic polymorphisms that can be used to identify population structure and evolutionary parameters.*

Objective 2: *Determine the population structure of Phakopsora pachyrhizi in the southeastern US.*

***Hypothesis** – P. pachyrhizi isolates from the southeastern US have a distinct geographic population structure.*

LITERATURE REVIEW

Soybean Characteristics

Soybean [*Glycine max* (L.) Merr.] is an annual plant belonging to the leguminosae family. With breeding efforts, soybeans have become a global food commodity and are well adapted to a variety of climates ranging from tropical to temperate (Dashiell *et al.*, 1987). A total of 340 million metric tons of soybean are produced globally, with the United States producing over one third of the global soybean supply at 120 million metric tons (ASA, 2018). The majority of soybean production in the US is concentrated along the Mississippi river and west of the Ohio River. About 31% of farmland in the US is dedicated to soybean production making soybean the second most produced crop in the US followed by maize at 35% (USDA NASS, 2016).

The global use of soybean can be attributed to its protein and oil rich seeds, which provide important food and oil sources for humans, industry, and livestock. Soybeans account for 61% of the global oil-seed production as of 2016 (USDA FAS, 2016). The plant is an increasingly utilized source of protein, unsaturated fatty acids, minerals such as calcium and phosphorus, and vitamins including A, B, and D (Erickson, 2015). Soybean seeds contain around 40% protein, 20% oil, and 20-26% carbohydrates (Wolf *et al.*, 1982). From a soil health and production perspective, soybeans are used extensively as a rotation crop due to their ability to fix nitrogen. This ability is a result of soybean's symbiotic relationship with a nitrogen fixing root nodulating bacterium, *Rhizobium japonicum* (Delves *et al.*, 1986). Nitrogen fixation benefits the

soybean plant, reduces production costs, and makes the soybean a viable rotational crop to use with high nitrogen consuming crops such as maize (Erickson, 2015).

Pucciniomycotina and Soybean Rust

Rust fungi, in the subfamily *Pucciniomycotina*, are the predominant cause of the most devastating foliar diseases to agronomic crops such as cereal grains and soybeans (Aime *et al.*, 2017). They represent one of the largest fungal orders, with over 8000 species described (Aime *et al.*, 2014). Rust fungi are obligate biotrophs that typically exhibit highly specific host interactions, typically resulting in coevolution with their host plants (Aime *et al.*, 2017). Rust pathogens are found on a wide variety of host plants including ferns, monocots, gymnosperms, angiosperms, and everything in-between. They are considered among the most serious threats to agricultural crops such as wheat (*Triticum aestivum*), maize (*Zea mays*), coffee (*Coffea arabica*), and soybean (*Glycine max*) (Dean *et al.*, 2012; Fischer *et al.*, 2012). On a global scale, production costs associated with rust diseases cost an estimated $4-5 billion USD annually (Figueroa *et al.*, 2017).

Two morphologically similar fungi, *Phakopsora pachyrhizi* and *Phakopsora meibomiae*, cause rust disease on soybeans (Bonde, 1988; Bromfield, 1984). While both fungi are considered pathogens, *P. meibomiae*, known as New World soybean rust because of its initial identification in Central and South America, does not threaten yields (Bonde *et al.*, 2006; Ono *et al.*, 1992). *P. pachyrhizi*, known as Asian Soybean Rust (SBR) due to its eastern Asian origins, is associated with more aggressive virulence characteristics across a wide range of soybean accessions (Goellner *et al.*, 2010). Morphological characteristics and species-specific DNA primers can be used to differentiate between the two species (Fredrick *et al.*, 2002; Ono *et al.*, 1992).

Phakopsora pachyrhizi Syd. is an obligate biotrophic fungus, meaning it can only complete its lifecycle on a living host (Bromfield, 1984). This dependence on a living host means that culturing the fungus on artificial media is not possible. Unlike most rust pathogens that are limited to one to two host species, *P. pachyrhizi* has a wide host range with over 80 species in the legume subfamily *Faboideae* susceptible to infection (Goellner *et al.*, 2010). Of these 80 plus hosts, soybeans are the most agronomically important (Slaminko *et al.*, 2008; Bromfield, 1984). *Glycine soja,* a wild soybean progenitor, *Pachyrhizus erosus,* Jicama, and *Pueraria lobata,* kudzu, are also notable hosts for *P. pachyrhizi* (Ono *et al.*, 1992). SBR causes losses between 10% and 80% in untreated fields (Bromfield, 1984). The variability in loss is primarily due to environmental fluctuations that influence *P. pachyrhizi's* survivability as well as host susceptibility. The pathogen is also hindered by a high UV index and temperatures above 30°C (Li *et al.*, 2010). *P. pachyrhizi* pustules primarily form on the abaxial side of soybean leaflets and produce large numbers of urediniospores on susceptible hosts (Bromfield, 1984). The symptoms indicative of infection on susceptible soybean hosts include small chlorotic spots on into uredia (Koch *et al.*, 1983). Under high disease pressure defoliation can occur early and, in combination with lesions on remaining leaves, reduce photosynthesis and yield (Kumundini *et al.*, 2008).

Rust Life Cycles, Parasexuality, and SBR Infection

The life cycle of most rusts is complex, involving up to five spore types and two hosts to complete their life cycle. A typical rust life cycle includes sexual reproduction via formation of teliospores that germinate into basidia after karyogamy of a dikaryon. This is followed by division and reduction to four haploid basidiospores, that spread to a secondary host in

heteroecious rusts, or reinfect the primary host in the case of autoecious rusts. In macrocyclic-heteroecious rusts, the haploid basidiospores infect the aecial host, producing pycnia with receptive hyphae. Fertilization can occur between two mating types during this process as pycniospores are formed inside exuded droplets on the receptive hyphae and are spread around by insects. Plasmogamy occurs and forms dikaryotic aecia on the abaxial side of the leaf. Aeciospores are created and are windblown back to the telial host. Aeciospore infection produces uredia and urediniospores, which are the repeating vegetative cycle of the rust. Uredia differentiate into telia where karyogamy occurs, giving rise to diploid teliospores that are typically an overwintering structure. An example of a typical rust life cycle is detailed in Lorrain *et al.* (2019).

Phakopsora pachyrhizi's lifecycle is unique in that the common sexual stages have not been fully characterized. It is not known whether *P. pachyrhizi* is heteroecious or autoecious, as germination of spore types other than urediniospores, has not been observed in nature. In Asia, there is record of teliospore formation, but no germination into a basidium (Bromfield, 1984). Artificial germination of teliospores into a basidium, followed by basidiospore formation has been observed in a lab setting (Saksiriat and Hoppe, 1991). Additionally, an aecial host has not been found (Goellner *et al.,* 2010). The parasexual cycle is characterized by somatic hyphal fusion and genetic recombination in the absence of meiosis and contributes to the genetic diversity in *P. pachyrhizi* despite the pathogen's lack of a known sexual cycle in nature (Goellner, 2010). Somatic hyphae from different genotypes fuse in a process called anastomosis, followed by heterokaryosis, nuclear fusion, recombination, and chromosomal reassortment to complete the parasexual cycle (Park and Wellings, 2012). The parasexual cycle has been supported by evidence of nuclear migration between hyphal and germ tube anastomosis (Vittal *et*

al., 2011). This parasexual cycle has been confirmed by molecular techniques with the Ug99 lineage of the wheat rust pathogen *Puccinia graminis* f. sp. *tritici* (Li *et al.*, 2019)

SBR infection begins with a urediniospore on the leaf surface of a soybean or other host. The urediniospore germinates forming a germ tube, typically on the abaxial side of the leaf, and terminates with a non-ornamented appressorium (Koch and Hoppe, 1988). The appressorium penetrates directly through the epidermal layer of the host plant via pressure build up and a penetration peg. The process of direct penetration on the primary host plant is unique among most obligate fungal pathogens and is possibly the reason for *P. pachyrhizi's* virulence and large host range (Goellner *et al.*, 2010). A septum is laid down to separate the penetration hyphae from primary and secondary hyphae that continues to colonize the intercellular space of the host plant (Koch *et al*., 1983). Hyphae differentiate into haustoria near mesophyll cells, which begin to form between 24-48 hours post contact of host plant and urediniospore. Haustoria are vital structures that exchange nutrients and proteins between the plant and the pathogen (Voegele and Mendgen, 2011; Kemen *et al.*, 2005; Dracatos *et al.*, 2018). Following haustorial formation, mesophyll tissue and intercellular spaces are colonized. Five to eight days after infection, asexual urediniospores are formed within a uredium on short stalks on foliage. When the urediniospores are mature, they are released through the uredium, and dispersed by the wind, traveling long distances to find another susceptible host (Goellner *et al.*, 2010).

Distribution History of *Phakopsora pachyrhizi*

The first reporting of the disease was in 1902 under the name *Uredo sojae* Henn. in Japan, where it was found on *Glycine soja* Siebold and Zucc. (Bromfield, 1984). The fungus was reported under its current name in 1913, where it was found on jicama [*Pachyrhizus erosus* (L.)

Urb.] in Taiwan (Bromfield, 1984). From there, it spread to Australia in 1934 (Bromfield, 1984). It was thought to spread to Central and South American countries in the 1970s and early 1980s, but this was actually *P. meibomiae* (Bonde *et al.*, 2006). *P. pachyrhizi* did eventually spread to Hawaii in 1994, to central African countries from 1997 to 2001, and then to South and Central American countries from 1997 to 2003 (Freire *et al.*, 2008; Killgore *et al.*, 1994; Levy, 2005; Pretorius *et al.*, 2001). The first report in the continental US was in November of 2004 (Schneider *et al.*, 2005). Urediniospores are thought to have been blown in from South America by tropical storms, as well as through clothing and increased human mobility *et al.*, 2005). In 2004, SBR was also found on Kudzu (*Pueraria lobata*) in Florida, where the pathogen is thought to overwinter in the soybean free time of the year (Harmon *et al.*, 2005). *P. pachyrhizi* has not spread significantly in areas of the US other than the deep south, as freezing weather in the winter kills urediniospores (Pivonia, 2004). Yield losses in the US have been reported upwards of 33% in fields with no fungicide applications (Mueller *et al.*, 2008). The Soybean Rust Commentary (http://www.asrusa.net/) scouting data in 2017 showed observation of SBR in 78 counties in Mississippi, 59 counties in Alabama, 15 counties in Georgia, 11 counties in Tennessee, 9 parishes in Louisiana, 4 counties in Florida, 7 counties in South Carolina, 7 counties in Arkansas, and 1 county in Kentucky, for a total of 191 positive counties/parishes. August 2021 data indicates 23 counties in Georgia, Alabama, and Florida were found to have SBR in the field as early as March 2021 (https://soybean.ipmpipe.org/soybeanrust/).[1] These data indicate pockets of infection in line with the idea that the *P. pachyrhizi* spores are windblown into farms each year from overwintering sites in Florida .

[1] Scouting data for SBR in the US is currently being migrated over to an integrated pest management website, data prior to 2019 is currently unavailable, and not all scouting data for 2019 is available on this website (https://soybean.ipmpipe.org/soybeanrust/).

Soybean Rust Management

SBR disease management fall into three distinct categories: cultural practices, chemical control, and host resistance through breeding for genetically resistant lines (Yorinori *et al.*, 2005; Sikora *et al.*, 2009; Langenbach *et al.*, 2016). The overall goal of these practices is to provide durable and broad-spectrum resistance against pathogens, and to enhance crop yield. Cultural methods for disease management are modified depending on the geographic regions where the crop is being grown. However, some information about them can be found in Langenbach *et al.* (2016).

Chemical and Genetic Management of SBR

The most effective short-term management strategy for preventing and combating SBR from year to year is through application of fungicides (Dalla Lana et al., 2018). Disease control through fungicide use is costly. In Brazil, three treatments per season costs the growers ~2 billion USD annually (Godoy *et al.*, 2016). Chemicals in the demethylation inhibitors (DMI) and quinone outside inhibitor ***(***QoI) classes, as well as the highly active SDHI class, are effective means for managing SBR (Guicherit *et al.*, 2014). A predominant concern of fungicide usage is in the potential for pathogen populations to develop insensitivity. Fungicide insensitivity in *P. pachyrhizi* has been observed with DMIs, caused by either point mutations in, or overexpression of, the *cyp51* gene (Scherm *et al.*, 2009; Schmitz *et al.*, 2014; Reis *et al.*, 2015). QoI insensitivity has also been observed in *P. pachyrhizi*, as a result of a *cyt b* mutation (F129L) (Klosowski *et al.*, 2015). Since the introduction of SDHI use for SBR in 2013, this class of fungicides has become increasingly used and at least one isolate of *P. pachyrhizi* has been found

to be insensitive (Godoy *et al.*, 2015; Simões *et al.*, 2018). Widespread usage of highly effective fungicides with single modes of action cause selective pressure on a pathogen to develop resistance to the fungicide. This effect is compounded by the polycyclic nature of *P. pachyrhizi* (Bradley, 2007). An additional issue with chemical applications is that the timing of application is linked with its efficacy, meaning that preventative and early applications in the disease cycle are the most effective means of control for SBR (Mueller *et al.*, 2009; Godoy, 2011; Childs *et al.*, 2019). Early detection and forecasting are required for efficient disease management. In the southeastern US, a sentinel plot management system is utilized. Sentinel plots are untreated soybean plots that are planted prior to the main crop and are scouted weekly to assess disease. Disease develops on these more mature soybean plants more quickly than the regular crop, meaning there is a forecasting aspect to the sentinel plots as well as detection of disease presence. The plots were created for SBR detection but have since been utilized for scouting many other diseases (Bob Kemerait, personal communication).

A sustainable practice for disease management is needed to accomplish the goal of broad-spectrum and durable resistance, and therefore provide predictable harvests for this global commodity. A method that could fill this requirement is genetic resistance of soybean varieties. This method of management minimizes risk of ecological damage and sanitation processes required of fungicide usage, decreases production costs, and improves the CO_2 footprint of soybean products (Maltby *et al.*, 2009; Verweij *et al.*, 2009; Wightwick *et al.*, 2010).

Plants are generally resistant to the majority of pathogens, a phenomenon called non-host resistance. Plants have also evolved a complex multi-layered defense system to protect themselves against pathogens that have the ability to cause serious diseases (Dodds and Rathjen, 2010). The first layer of plant defenses is passive resistance. This type of resistance consists of

preformed barriers that are present in the plant prior to infection. Pre-formed plant defense structures include leaf cuticle wax, leaf trichomes and the chemical structure, thickness and crosslinking of epidermal cell walls, that all function to stop the pathogen from entering the plant (Freeman and Beattie, 2008). The next layer of plant defenses is the active plant immune system that can be conceptually divided into two modes of responses, pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) (Dodds and Rathjen, 2010).

PTI is induced when pathogen associated molecular patterns (PAMPs) are recognized by plant pattern recognition receptors (PRRs), causing a signaling cascade in the plant that results in immunity from the pathogen. This PTI defense response has sometimes been overcome by pathogens through secreted proteins called effector molecules that typically act to suppress signaling components in the PTI pathway, resulting in effector triggered susceptibility (ETS). In response, plants have developed resistance genes (R-genes) that produce proteins which recognize specific and complementary effector proteins. This type of interaction is qualitative. There are different classes of R-genes that have multiple domains that allow for recognition of effectors followed by activation of plant defenses (Petit-Houdenot and Fudal, 2017). The largest class of R-genes is the nucleotide binding site, leucine rich repeat (NBS-LRR) class. This class of R-gene proteins function to recognize effector proteins (directly or indirectly) and activate downstream signaling pathways that in turn activate plant defenses. Recognition of effectors often results in activation of a separate signaling cascade from PTI that results in a hypersensitive response, HR. This process is known as ETI, or Effector Triggered Immunity (Flor, 1955; Keen 1990; Staskawicz *et al.*, 1995). HR results in programmed cell death, preventing the spread of biotrophic pathogens, such as *P. pachyrhizi* (Scofield *et al.*, 1996; Tang *et al.*, 1996). In direct interactions of R-gene proteins and secreted effector proteins, this

interaction is known as a gene-for-gene interaction (Flor, 1971). As not all effectors interact with R-gene proteins, the ones that do are known as Avirulence proteins. This interaction is referred to as an R-AVR gene interaction. In *P. pachyrhizi*, HR occurs after penetration and haustorial formation begins inside the host plant's mesophyll cells, hindering colonization of *P. pachyrhizi* (Szabo and Bushnell, 2001). As a result of the specificity of the interaction between resistance and effector genes, gene-for-gene types of resistance are often broken as a result of small mutations to, or complete loss of an effector gene in the pathogen population (McDonald and Linde, 2002). These small mutations mean that R-gene proteins cannot recognize the pathogen, therefore, a HR is not induced. Other models of interaction between R-gene proteins and effectors exist, such as the guard model, decoy model, and the integrated decoy model. In the guard model, interaction between the R and AVR proteins is indirect, through detection of an effector's host target protein called the "guardee" (Dangl and Jones, 2001). The decoy model suggests that an R-protein detects modification to another plant protein that has the sole purpose of mimicking the effector target protein (van der Hoorn and Kamoun, 2008). The integrated decoy suggests domains are integrated into NLRs that play the role of the decoy (Cesari *et al.*, 2014; Le Roux *et al.*, 2015; Sarris *et al.*, 2015). These models increase in the complexity of interaction between the R-AVR gene and generally are thought of as more durable as they can decrease selection pressure on the pathogen.

Single R-genes have not historically provided durable resistance on their own but can be stacked, or pyramided, through breeding efforts to combine traits from multiple parents (Francis *et al.*, 2012). This method of combining resistance has provided a more robust and durable level of resistance for many crop plant species via reducing the selection pressure on the pathogen to overcome individual resistance genes (Maphosa *et al.*, 2012; Bhor *et al.*, 2015). A natural

example of R-gene pyramiding is found in a soybean Japanese cultivar Hyuuga which has *Rpp3* and a second unlinked gene located near *Rpp5* (Kendrick *et al.*, 2011). Traditional breeding efforts to create these R-gene pyramids is time consuming, requiring many generations of hundreds of plants that are phenotyped for resistance and then crossed (Salomon and Sessa, 2012; Langenbach et al., 2016). Traditional breeding can also introduce unwanted traits, so breeding efforts can be more efficiently facilitated via marker assisted selection which allows for breeding with more directed introgression of wanted traits (Langenbach, 2016). Resistance is likely to still be overcome with static R-gene pyramids for the same reason as single gene resistance (McDonald, 2014).

Single dominant resistance genes that confer race specific resistance to *P. pachyrhizi* have not been durable in the field, breaking down within a few years after deployment (Hartman *et al.,* 2005). Seven dominant resistance genes, designated *Rpp1-7,* as well as 6 additional alleles, have been discovered and mapped to seven different loci in soybeans (*Rpp*1, 2, 3, 4, 5, 6, 7) (Bromfield and Hartwig 1980; Bromfield *et al.*, 1980; Chakraborty *et al.* 2009; Childs *et al.*, 2018; Garcia *et al.*, 2008; Hartwig and Broomfield 1983; Hartwig, 1986; Hyten *et al.*, 2007, 2009; King *et al.* 2015, 2017; Li *et al.*, 2012; McLean and Byth 1980; Ray *et al.*, 2011; Silva *et al.*, 2008). These dominant R-genes confer an immune or a resistant response. R-genes that confer resistance to all races of *P. pachyrhizi* have not been identified. Additionally, none of the current soybean accessions have resistance to all isolates of *P. pachyrhizi* (Hartman *et al.*, 2005; Bonde *et al.*, 2006; Monteros *et al.*, 2007). Resistance genes in the field historically have a limited average lifespan of around five years, only conferring resistance until they are overcome by the pathogen (Garcia *et al.*, 2008). Some recessive resistance genes such as *rpp2* and *rpp5* also provide resistance to *P. pachyrhizi* and may be exploited in breeding efforts. It is likely there

are more genes to exploit in wild progenitors of soybean; *Glycine soja, G. tomentella,* as well as a several of traits from the USDA Soybean Germplasm collection of over 20,000 globally collected varieties (Childs *et al.*, 2018).

Unlike the qualitative, "on" or "off", nature of single dominant R genes, plants also have quantitative trait loci (QTL) that contribute a variable degree of resistance to a pathogen. QTL provides additive effects of resistance and can be located across the genome. Many of these genes are thought to have arisen from, or are closely related to, R-genes. For example, resistance to *Colletotrichum graminicola* 1 (*Rcg1*), a large effect QTL, that confers resistance to anthracnose stalk rot was found to encode for an NB-LRR R-gene (Frey, 2006). In the case of SBR resistance, there are "slow rusting" accessions (SRE-Z-11A, SRE-Z-11B, SRE-Z-15A,) that produce a lower infection frequency, smaller lesions, and reduced sporulation in comparison to susceptible accessions (Tukamuhabwa and Maphosa, 2010). QTL have been identified in soybean through forward genetic screening and activation tagged soybean plants (Mathieu *et al.*, 2009). There is evidence of yield increases of 30-60% when using "slow rusting" accessions during years with high pathogen pressure, giving credence to the idea of using these accessions to reduce SBR epidemics (Tukamuhabwa and Maphosa, 2010). Additionally, the polygenic nature of partially disease resistant soybean accessions reduces selection pressure on *P. pachyrhizi* to select compatible pathogen races (Arias *et al.*, 2008). As a result, QTL have the capacity to provide durable resistance, and some genes have been found to provide a spectrum of resistance to multiple pathogen races of SBR (Long *et al.*, 2006; Langenbach *et al.*, 2016). Comparative transcriptome, as well as proteome analysis have been used to identify these QTL and, in some cases, provide resources for soybean varieties with tolerance to high SBR pressure and produce an acceptable yield (Langenbach *et al.*, 2016).

Population Genomics and Structure

To inform disease management practices for any disease, there is requisite knowledge needed about the pathogen in question. One form of this knowledge is through the study of a pathogen's population structure and its genomic diversitywhich are a measure the variation and the specific organization of a population with regards to its genetic makeup. This can include aspects such as genetic drift, recombination, mutations, selection pressures, demographics, all of which are the combined effects of evolution in a population (Grünwald *et al.,* 2016).

The general population diversity and structure of SBR has been assessed with microsatellite markers, the internal transcribed spacer region (ITS) and single genes including the ADP ribosylation factor (ARF) (Twizeyimana *et al*., 2011, Freire *et al*., 2008, 2012; Zhang *et al*., 2012). These studies indicated little population structure over large geographic areas and that most variation was maintained within each location sampled. These studies also predicted mechanisms for long distance dispersal, lining up with the current model of urediniospore wind dispersal. However, there have been no robust population genomics analyses on SBR beyond simple sanger sequencing and sampling small regions of the genome or single genes. Additionally, there have been no population genomics analyses of SBR that have focused on the southeastern US, indicating a gap in knowledge.

The current trends in population genomics and population structure analyses include discriminant analysis of principal components (DAPC) (Miller et al. 2020). This combines discriminant analysis with principal components analysis (PCA) which is useful for inferring divergence and distribution of genetic clusters in population data (Sousa and Hey, 2013). This method is used, along with K-means clustering, to determine the appropriate number of

populations (K) in a given number of samples. PCA is another analysis method which measures the maximum level of variance between samples in abstract principle components. Population structure and genetic diversity of many rusts, plants, and other pathogens have been characterized using next generation sequencing. Puccinia rusts on wheat, including *Puccinia striiformis* and *P. triticina,* have been characterized globally (Thach et al, 2016, Ali et al., 2014, Kolmer et al., 2020). The genetic diversity and structure of other plant pathogens has been characterized using genotype-by-sequencing (Halpern et al., 2020).

REFERENCES

Aime MC, MacTaggart AR, Mondo SJ, Duplessis S. 2017. Phylogenetics and phylogenomics of rust fungi. Advances in Fungal Genetics 100: 267– 307.

Aime M, Toome M, McLaughlin D. 2014. The Pucciniomycotina. In: D McLaughlin, JW Spatafora, eds. The Mycota VII Part A. Berlin/Heidelberg, Germany: Springer, 271– 294.

Akamatsu, H., Yamanaka, N., Yamaoka, Y., Soares, R. M., Morel, W., Ivancovich, A. J. G., ... and Suenaga, K. 2013. Pathogenic diversity of soybean rust in Argentina, Brazil, and Paraguay. Journal of General Plant Pathology, *79*(1), 28-40.

Ali, S., Gladieux, P., Leconte, M., Gautier, A., Justesen, A. F., Hovmøller, M. S., ... and de Vallavieille-Pope, C. 2014. Origin, migration routes and worldwide population genetic structure of the wheat yellow rust pathogen *Puccinia striiformis* f. sp. *tritici*. PLoS pathogens, *10*(1), e1003903.

Arias, C. A. A., Toledo, J. F. F., Almeida, L. A., Pipolo, G. E. S., Carneiro, R. V., Abdelnoor, R. V., *et al.* 2008. "Asian rust in Brazil: varietal resistance," in *Facing the Challenge of Soybean Rust in South America*, eds H. Kudo, K. Suenaga, R. M. S. Soares, and A. Toledo (Tsukuba: JIRCAS), 29–30.

American Soybean Association 2018. https://ndsoygrowers.com/wp-content/uploads/2018/09/2018ASA-SoyStats.pdf.

Bhor, T. J., Chimote, V. P., and Deshmukh, M. P. 2015. Molecular tagging of Asiatic soybean rust resistance in exotic genotype EC 241780 reveals complementation of two genes. Plant Breed. 134, 70–77. doi: 10.1111/pbr.12240

Bonde, M.R. 1988. A comparison of isoenzymes of *Phakopsora pachyrhizi* from the eastern hemisphere and the new world. Phytopathology, 78, 1491–1494.

Bonde MR, Nester SE, Austin CN, Stone CL, Frederick RD, Hartman GL, andMiles MR 2006. Evaluation of virulence of *Phakopsora pachyrhizi* and *P. meibomiae i*solates. Plant Dis 90:708–716. doi: 10.1094/PD-90-0708

Bromfield, K. R. 1984. Soybean rust, Monograph (American Phytopathological Society), No. 11 American Phytopathological Society. St. Paul, MN.

Bromfield, K.R. and Hartwig, E.E. 1980. Resistance to soybean rust and mode of inheritance. Crop. Sci. 20, 254–255.

Bromfield, K. R., Melching, J. S., and Kingsolver, C. H. 1980. Virulence and aggressiveness of Phakopsora *pachyrhizi* isolates causing soybean rust. *Phytopathology*, *70*(1), 17-21.

Bradley, C. A. 2007. "Fungicide resistance management in soybean," in *Using Foliar Fungicides to Manage Soybean Rust*, eds A. E. Dorrance, M. A. Draper, and D. E. Hershman (Columbus, OH: Land-Grant Universities Cooperating NCERA-208 and OMAF), 57–60.

Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P. 2014. A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. Front. Plant Sci. 5:606. doi: 10.3389/fpls.2014.00606

Chakraborty, N., Curley, J., Frederick, R. D., Hyten, D. L., Nelson, R. L., Hartman, G. L., Diers, B. W. 2009. Mapping and confirmation of a new allele at *Rpp1* from soybean PI 594538A conferring RB lesion-type resistance to soybean rust. *Crop Sci* 49:783–790. doi: doi: 10.2135/cropsci2008.06.0335

Childs, S.P., King, Z.R., Walker, D.R. *et al.* 2018. Discovery of a seventh Rpp soybean rust resistance locus in soybean accession PI 605823. Theor Appl Genet 131:27. doi: 10.1007/s00122-017-2983-4

Childs, S. P., Buck, J. W., and Li, Z. (2018). Breeding soybeans with resistance to soybean rust (Phakopsora pachyrhizi). *Plant Breeding*, *137*(3), 250-261.

Dalla Lana, F., Paul, P.A., Godoy, C.V., Utiamada, C.M., da Silva, L.H.C., Siqueri, F.V., Forcelini, C.A., Jaccoud-Filho, D.D.S., Miguel-Wruck, D.S., Borges, E.P. and Juliatti, F.C. 2018. Meta-analytic modeling of the decline in performance of fungicides for managing soybean rust after a decade of use in Brazil. Plant Disease, *102*(4), 807-817.

Dangl, J. L., and Jones, J. D. 2001. Plant pathogens and integrated defence responses to infection. Nature, *411*(6839), 826.

Dashiell, K. E., Bello, L. L., and Root, W. R. 1987. Breeding soybeans for the tropics. Soybean for the tropics. Research, production and utilisation. Wiley, Chichester, 3-16.

Dean R, van Kan JAL, Pretorius ZA, Hammond-Kosack KE, di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J *et al.* 2012. The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* 13: 414– 430.

Delves, A. C., Mathews, A., Day, D. A., Carter, A. S., Carroll, B. J., and Gresshoff, P. M. 1986. Regulation of the soybean-Rhizobium nodule symbiosis by shoot and root factors. Plant Physiology, *82*(2), 588-590.

Dodds, P. N., and Rathjen, J. P. 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. Nature Reviews Genetics, *11*(8), 539-548.

Dracatos, P. M., Haghdoust, R., Singh, D., Park, R. F. 2018. Exploring and exploiting the boundaries of host specificity using the cereal rust and mildew models. New Phytologist 218:453– 462.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one, *6*(5), e19379.

Erickson, D. R. (Ed.). 2015. Practical handbook of soybean processing and utilization. Elsevier.

Figueroa, M., Hammond-Kosack, K. E., Solomon, P. S. 2017. A review of wheat diseases – a field perspective. Molecular Plant Pathology 19: 1523– 1536.

Fischer, M. C., Henck, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484, 186– 194.

Flor, H. H. 1955. Host-parasite interaction in flax rust-its genetics and other implications. Phytopathology 45:680-685.

Flor, H. H. 1971. Current status of the gene-for-gene concept. Annu. Rev. Phytopathology. 9, 275-296.

Francis, D. M., Merk, H. L., and Namuth-Covert, D. 2012. Gene pyramiding using molecular markers. Plant Breed. Genomics. Available at: http://articles.extension.org/pages/32465/gene-pyramiding-using-molecular-markers.

Freeman, B. C., and Beattie, G. A. 2008. An overview of plant defenses against pathogens and herbivores. The Plant Health Instructor. doi: 10.1094/PHI-I-2008-0226-01

Frederick, R. D., Snyder, C. L., Peterson, G. L., and Bonde, M. R. 2002. Polymerase chain reaction assays for the detection and discrimination of the soybean rust pathogens *Phakopsora pachyrhizi* and *P. meibomiae*. Phytopathology 92:217-227.

Freire, M.C.M., de Oliveira, L.O., de Almeida, A.M.R., Schuster, I., Moreira, M.A., Liebenberg, M.M. and Mienie, C.M.S. 2008. Evolutionary history of *Phakopsora pachyrhizi* (the Asian soybean rust) in Brazil based on nucleotide sequences of the internal transcribed spacer region of the nuclear ribosomal DNA. Genet. Mol. Biol. 31, 920–931.

Frey, T.J. 2006. Finemapping, cloning, verification, and fitness evaluation of a QTL, Rcg1, which confers resistance to *Colletotrichum graminicola* in maize. Ph.D. diss. Univ. of Delaware, Newark, DE.

Garcia, A., Calvo, E. S., Kiihl, R. A. S., Harada, A., Hiromoto, D. M., and Vieira, L. G. E. 2008. Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. Theor Appl Genet 117, 545–553. doi: 10.1007/s00122-008-0798-z

Godoy, C. V. 2011. "*Phakopsora pachyrhizi*: the performance of soybean rust fungicides over years and regions in Brazil," in Modern Fungicides and Antifungal Compounds VI, eds H. W. Dehne, H. B. Deising, U. Gisi, K. H. Kuck, P. E. Russell, and H. Lyr (Braunschweig: Deutsche Phytomedizinische Gesellschaft e.V. Selbstverlag), 203–209.

Godoy, C. V., Bueno, A. F., and Gazziero, D. L. P. 2015. Brazilian soybean pest management and threats to its sustainability. Outlooks Pest Manag. 26, 113–117. doi: 10.1564/v26_jun_06

Godoy, C. V., Seixas, C. D. S., Soares, R. M., Marcelino-Guimaraes, F. C., Meyer, M. C., Costamilan, L. M. 2016. Asian soybean rust in Brazil: past, present, and future. Pesquisa Agropecuaria Brasileira 51:407–421. doi: 10.1590/S0100-204X2016000500002

Goellner, K., Loehrer, M., Langenbach, C., Conrath, U., Koch, E.,and Schaffrath, U. 2010. *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. Molecular Plant Pathology 11:169-177. doi: 10.1111/J.1364-3703.2009.00589.X

Grünwald, N. J., McDonald, B. A., and Milgroom, M. G. 2016. Population genomics of fungal and oomycete pathogens. Annual Review of Phytopathology, 54, 323-346.

Guicherit, E., Bartlett, D., Dale, S. M., Haas, H. U., Scalliet, G., Walter, H., et al. 2014. "Solatenol-the second generation benzonorbornene SDHI carboxamide with outstanding performance against key crop diseases," in Modern Fungicides and Antifungal Compounds VII, eds H. W. Dehne, H. B. Deising, B. Fraaije, U. Gisi, D. Hermann, A. Mehl, et al. (Braunschweig: Deutsche Phytomedizinische Gesellschaft e.V. Selbstverlag), 67–72.

Halpern, H. C., Qi, P., Kemerait, R. C., and Brewer, M. T. 2020. Genetic diversity and population structure of races of Fusarium oxysporum causing cotton wilt. G3: Genes, Genomes, Genetics, 10(9), 3261-3269.

Harmon, P. F., Momol, M. T., Marois, J. J., Dankers, H., and Harmon, G. L. 2005. Asian soybean rust caused by *Phakopsora pachyrhizi* on soybean and kudzu in Florida. Plant health progress, 2005, 1-4.

Hartman, G. L., Miles, M. R., and Frederick, R. D. 2005. Breeding for resistance to soybean rust. Plant Dis 89:664–666. doi: 10.1094/PD-89-0664

Hartwig E. E., and Bromfield K. R. 1983. Relationships among three genes conferring specific resistance to rust in soybeans. Crop Sci 23:237–239.

Hartwig, E. E. 1986. Identification of a fourth major gene conferring resistance to soybean rust. Crop Sci. 26, 1135–1136. doi: 10.2135/cropsci1986.0011183X002600060010x

Hyten D. L., Hartman G. L., Nelson R. L., Frederick R. D., Concibido V. C., Narvel J. M., and Cregan P. B. 2007. Map location of the Rpp1 locus that confers resistance to soybean rust in soybean. Crop Sci 47:837–840. doi: 10.2135/cropsci2006.07.0484

Isard, S.A., Gage, S.H., Comtois, P. and Russo, J.M. 2005. Principles of the atmospheric pathway for invasive species applied to soybean rust. Bioscience, 55, 851–861.

Japan International Research Center for Agricultural Sciences (JIRCAS). 2019. Laboratory manual for studies on soybean rust resistance. 1-50.
https://www.jircas.go.jp/en/publication/manual_gudeline/30

Keen, N. T. 1990. Gene-for-gene complementarity in plant-pathogen interactions. Annual review of genetics, 24(1), 447-463.

Kendrick, M. D., Harris, D. K., Ha, B.-K., Hyten, D. L., Cregan, P. B., Frederick, R. D., 2011. Identification of a second Asian soybean rust resistance gene in Hyuuga soybean. Phytopathology 101, 535–543. doi: 10.1094/PHYTO-09-10-0257

Kemen, E., Kemen, A. C., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., et al. 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. Mol. Plant Microbe Interact. 18, 1130–1139. doi: 10.1094/MPMI-18-1130

Killgore, E., Heu, R. and Gardner, D.E. 1994. First report of soybean rust in Hawaii. Plant Dis. 78, 1216–1216.

King, Z. R., Harris, D. K., Pedley, K. F., Song, Q., Wang, D., Wen, Z., Buck, J. W., Li, Z., and Boerma, H. R. 2015. A novel *Phakopsora pachyrhizi* resistance allele (Rpp) contributed by PI 567068A. Theor Appl Genet 129:517–534. doi: 10.1007/s00122-015-2645-3

King, Z. R., Childs, S. P., Harris, D. K., Pedley, K. F., Buck, J. W., Boerma, H. R., and Li, Z. 2017. A new soybean rust resistance allele from PI 423972 at the Rpp4 locus. Mol Breed 37:62. doi: 10.1007/s11032-017-0658-0

Klosowski, A. C., May De Mio, L. L., Miessner, S., Rodrigues, R., and Stammler, G. 2015. Detection of the F129L mutation in the cytochrome b gene in Phakopsora pachyrhizi. Pest Manag. Sci. 72, 1211–1215. doi: 10.1002/ps.4099

Koch, E., Ebrahimnesbat, F. and Hoppe, H.H. 1983. Light and electron microscopic studies on the development of soybean rust (*Phakopsora pachyrhizi* Syd) in susceptible soybean leaves. Phytopathol. Z. 106, 302–320.

Koch, E. and Hoppe, H.H. 1988. Development of infection structures by the direct-penetrating soybean rust fungus (*Phakopsora pachyrhizi* Syd) on artificial membranes. J. Phytopathol. 122, 232–244.

Kolmer, J.A., Herman, A., Ordoñez, M.E. 2020. Endemic and panglobal genetic groups, and divergence of host-associated forms in worldwide collections of the wheat leaf rust fungus

Puccinia triticina as determined by genotyping by sequencing. Heredity 124**,** 397–409. doi: 10.1038/s41437-019-0288-x

Kumundini, S, Godoy, CV, Board, JE, Omeilan, J, and Tollenaar, M. 2008. Mechanisms involved in soybean rust-induced yield reduction. Crop Sci. 48:2334-2342. doi: 10.2135/cropsci2008.01.0009

Langenbach, C., Campe, R., Beyer, S., Mueller, A., and Conrath, U. 2016. Fighting Asian Soybean Rust. Frontiers in Plant Science. 7, 797. doi: 10.3389/fpls.2016.00797

Levy, C. 2005. Epidemiology and chemical control of soybean rust in southern Africa. Plant Dis. 89, 669–674.

Le Roux, C., Jauneau, A., Camborde, L., Trémousaygue, D., Kraut, A., Zhou, B. 2015. A receptor pair with an integrated decoy converts pathogen disabling of transcription factors to immunity. Cell 161, 1074–1088.doi: 10.1016/j.cell.2015.04.025

Li, F., Upadhyaya, N.M., Sperschneider, J. 2019. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. Nat Commun 10, 5068. doi: 10.1038/s41467-019-12927-7

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics, 26(5), 589-595.

Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27(21), 2987-2993.

Li, S., Smith, J. R., Ray, J. D., Frederick, R. D. 2012. Identification of a new soybean rust resistance gene in PI 567102B. Theor Appl Genet 125:133–142. doi: 10.1007/s00122-012-1821-y

Li, X., Esker, P. D., Pan, Z., Dias, A. P., Xue L., and Yang, X. B. 2010. The uniqueness of the soybean rust pathosystem: an improved understanding of the risk in different regions of the world. Plant Dis 94, 796–806. doi: 10.1094/pdis-94-7-0796

Link, T. I., Lang, P., Scheffler, B. E., Duke, M. V., Graham, M. A., Cooper, B., Tucker, M. L., van de Mortel, M., Voegele, R. T., Mendgen, K. 2014. The haustorial transcriptomes of Uromyces appendiculatus and *Phakopsora pachyrhizi* and their candidate effector families. Molecular Plant Pathology 15, 379– 393.

Lorrain, C., Gonçalves dos Santos, K. C., Germain, H., Hecker, A. and Duplessis, S. 2019. Advances in understanding obligate biotrophy in rust fungi. New Phytol, 222: 1190-1206. doi: 10.1111/nph.15641

Long, J., Holland, J. B., Munkvold, G. P., and Jannink, J.-L. 2006. Responses to selection for partial resistance to crown rust in oat. Crop Sci. 46, 1260–1265. doi: 10.2135/cropsci2005.06-0169

Maltby, L., Brock, T. C. M., and van den Brink, P. J. 2009. Fungicide risk assessment for aquatic ecosystems: importance of interspecific variation, toxic mode of action, and exposure regime. Environ. Sci. Technol. 43, 7556–7563. doi: 10.1021/es901461c

Maphosa, M., Talwana, H., and Tukamuhabwa, P. 2012. Enhancing soybean rust resistance through Rpp2, Rpp3 and Rpp4 pair wise gene pyramiding. African J. Agric. Res. 7, 4271–4277. doi: 10.5897/AJAR12.1123

Mathieu, M., Winters, E. K., Kong, F., Wan, J., Wang, S., and Eckert, H. 2009. Establishment of a soybean (Glycine max Merr. L) transposon-based mutagenesis repository. Planta 229, 279–289. doi: 10.1007/s00425-008-0827-9

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal, 17(1), 10-12.

McDonald, B. A. 2014. Using dynamic diversity to achieve durable disease resistance in agricultural ecosystems. Trop. Plant Pathol. 39, 191–196. doi: 10.1590/S1982-56762014000300001

McDonald, B. A. and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Ann.Rev. Phytopathology 40: 349-379.

McLean, R. J., and Byth, D. E. 1980. Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. Aust. J. Agric. Res. 31: 951-956.

Möller, M., and Stukenbrock, E. H. 2017. Evolution and genome architecture in fungal plant pathogens. Nature Reviews Microbiology, 15(12), 756.

Monteros, M. J., Missaoui, A. M., Phillips, D. V., Walker, D. R., and Boerma, H. R. 2007. Mapping and confirmation of the “Hyuuga” red-brown lesion resistance gene for Asian soybean rust. Crop Sci. 47, 829–834. doi: 10.2135/cropsci06.07.0462

Mueller, T. A., Miles, M. R., Morel, W., Marois, J. J., Wright, D. L., Kemerait, R. C., Levy, C., and Hartman G. L. 2009. Effect of fungicide and timing of application on soybean rust severity and yield. Plant Dis. 93, 243–248. doi: 10.1094/PDIS-93-3-0243

Mueller J. D., Koenning S. R., Kemerait R. C., Phipps P. M. 2008. Soybean Rust Management in the Mid-Atlantic Region. Clemson, USA: Clemson University Extension Service publication.

Nagar, R., Schwessinger, B. 2018. High purity, high molecular weight DNA extraction from rust spores via CTAB based DNA precipitation for long read sequencing. protocols.io dx.doi.org/10.17504/protocols.io.n5ydg7w.

Ono, Y., Buritica, P. and Hennen, J. F. 1992. Delimitation of *Phakopsora Physopella*, and *Cerotelium* and their species on *leguminosae*. Mycol. Res 96: 825-850.

Panstruga, R. 2003. Establishing compatibility between plants and obligate biotrophic pathogens. Curr. Opin. Plant Biol. 6:320-326.

Park, R. F., and Wellings, C. R. 2012. Somatic hybridization in the *Uredinales*. Annual review of phytopathology, 50, 219-239.

Persoons, A., Morin, E., Delaruelle, C., Payen, T., Halkett, F., Frey, P., De Mita, S., and Duplessis, S. 2014. Patterns of genomic variation in the poplar rust fungus *Melampsora larici-populina* identify pathogenesis-related factors. Frontiers in Plant Science 5: 450.

Petit-Houdenot, Y., and Fudal, I. 2017. Complex interactions between fungal avirulence genes and their corresponding plant resistance genes and consequences for disease resistance management. Frontiers in plant science, 8, 1072.

Petre, B., Joly, D. L., and Duplessis, S. 2014. Effector proteins of rust fungi. Frontiers in plant science, 5, 416.

Pham, T. A., Miles, M. R., Frederick, R. D., Hill, R. D., Hill, C. B., and Hartman, G. 2009. Differential responses of resistant soybean entries to isolates of *Phakopsora pachyrhizi.* Plant Dis 93:224–228. doi: 10.1094/PDIS-93-3-0224

Pivonia, S., and Yang, X. B. 2004. Assessment of potential year round establishment of soybean rust throughout the world. Plant Disease 88: 523-529.

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J. L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one, 7(2), e32253.

Pretorius, Z. A., Kloppers, R. J., and Frederick, R. D. 2001. First Report of soybean rust in South Africa. Plant Dis, 85: 1288.

Qi, M., Grayczyk, J. P., Seitz, J. M., Lee, Y., Link, T. I., Choi, D., Pedley, K., Voegele, R., Baum, T., and Whitham, S. A. 2018. Suppression or activation of immune responses by predicted secreted proteins of the soybean rust pathogen *Phakopsora pachyrhizi*. Molecular Plant–Microbe Interactions 31: 163– 174.

Qi, M., Link, T. I., Muller, M., Hirschburger, D., Pudake, R. N., Pedley, K. F., Braun, E., Voegele, R. T., Baum, T. J., and Whitham, S. A. 2016. A small cysteine-rich protein from the asian soybean rust fungus, *Phakopsora pachyrhizi*, suppresses plant immunity. PLoS Pathogens 12: e1005827.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Reviews Microbiology 10: 417.

Rau, A., and Maugis-Rabusseau, C. 2018. Transformation and model choice for RNA-seq co-expression analysis. Briefings in bioinformatics, 19(3), 425.

Ray, J. D., Smith, J. R., Morel, W., Bogado, N., and Walker, D. R. 2011. Genetic resistance to soybean rust in PI567099A is at or near the Rpp3 locus. J. Crop Improv 25:219–231. doi: 10.1080/15427528.2011.555833

Rehmany, A. P., Gordon, A., Rose, L. E., Allen, R. L., Armstrong, M. R., Whisson, S. C., and Beynon, J. L. 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. The Plant Cell, 17(6), 1839-1850.

Reis, E. M., Deuner, E., Zanatta, M., Reis, E. M., Deuner, E., and Zanatta, M. 2015. In vivo sensitivity of *Phakopsora pachyrhizi* to DMI and QoI fungicides. Summa Phytopathol. 41, 21–24. doi: 10.1590/0100-5405/1975

Robinson, M. D., and Oshlack, A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology, 11(3), R25.

Sabre-barcode-demultiplexing: https://github.com/najoshi/sabre. Accessed 11 Aug 2019.

Saksiriat W., Hoppe HH. 1991. Teliospore germination of soybean rust fungus (Phakopsora pachyrhizi Syd.). J Phytopathol 132:339-342.

Salomon, D., and Sessa, G. 2012. “Biotechnological strategies for engineering plants with durable resistance to fungal and bacterial pathogens” in Plant Biotechnology and Agriculture - Prospects for the 21st Century-, eds A. Altman and P. M. Hasegawa (Cambridge, MS: Academic Press), 329–342. doi: 10.1016/B978-0-12-381466-1.00021-3

Saunders, D. G. O., Win, J., Cano. L. M., Szabo, L. J., Kamoun, S., and Raffaele, S. 2012. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. PLoS ONE 7: e29847.

Sarris, P. F., Duxbury, Z., Ma, Y., Segonzac, C., Sklenar, J., Derbyshire, P., et al. 2015. A plant immune receptor detects pathogen effectors that target WRKY transcription factors. Cell 161, 1089–1100. doi: 10.1016/j.cell.2015.04.024

Schneider, R. W., Hollier, C. A., Whitam, H. K., Palm, M. E., McKemy, J. M., Hernandez, J. R., Levy, L. and DeVries-Paterson, R. 2005. First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. Plant Dis 89:774.

Schafer, J. F., Roelfs, A. P. 1985. Estimated relation between numbers of urediniospores of *Puccinia graminis f. sp. tritici* and rates of occurrence of virulence. Phytopathology 75: 749-750.

Scherm, H., Christiano, R. S. C., Esker, P. D., Del Ponte, E. M., and Godoy, C. V. 2009. Quantitative review of fungicide efficacy trials for managing soybean rust in Brazil. Crop Prot. 28, 774–782. doi: doi: 10.1016/j.cropro.2009.05.006

Schmitz, H. K., Medeiros, C.-A., Craig, I. R., and Stammler, G. 2014. Sensitivity of *Phakopsora pachyrhizi* towards quinone-outside-inhibitors and demethylation-inhibitors, and corresponding resistance mechanisms. Pest Manag. Sci. 70, 378–388. doi: 10.1002/ps.3562

Scofield, S. R., Tobias, C. M., Rathjen, J. P., Chang, J. H., Lavelle, D. T., Michelmore, R. W. and Staskawicz, B. J. 1996. Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. Science 274: 2063-2065.

Sikora, E. J., Delaney, D. P., Delaney, M. A., Lawrence, K. S., Pegues, M. 2009. Evaluation of sequential fungicide spray programs for control of soybean rust. Plant Health Prog. doi: 10.1094/PHP-2009-0402-01-RS

Silva, D. C., Yamanaka, N., Brogin, R. L., Arias, C. A., Nepomuceno, A. L., Di Mauro, A. O. 2008. Molecular mapping of two loci that confer resistance to Asian rust in soybean. Theor Appl Genet 117:57–63. doi: 10.1007/s00122-008-0752-0

Simões, K., Hawlik, A., Rehfus, A., Gava, F., and Stammler, G. 2018. First detection of a SDH variant with reduced SDHI sensitivity in *Phakopsora pachyrhizi*. Journal of Plant Diseases and Protection, 125(1), 21-26.

Stark, R., Grzelak, M., Hadfield, J. 2019. RNA sequencing: the teenage years. Nature Reviews Genetics. doi: 10.1038/s41576-019-0150-2

Staskawicz, B. J., Ausubel, F. M., Baker, B. J., Ellis, J. G., and Jones, J. D. 1995. Molecular genetics of plant disease resistance. Science, 268(5211), 661-667.

Stergiopoulos, I., and de Wit, P. J. 2009. Fungal effector proteins. Annual review of phytopathology, 47, 233-263.

Stukenbrock, E. H. 2013. Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. New Phytologist 199: 895– 907.

Szabo, L. J. and Bushnell, W. R. 2001. Hidden robbers: The role of fungal haustoria in parasitism of plants. PNAS 98: 7654-7655.

Tang, X., Frederick, R. D., Zhou, J., Halterman, D. A., Jia, Y. and Martin, G. B. 1996. Initiation of plant disease resistance by physical interaction of AvrPto and Pto Kinase. Science 274:2060-2063.

Thach, T., Ali, S., de Vallavieille-Pope, C., Justesen, A. F., and Hovmøller, M. S. 2016. Worldwide population structure of the wheat rust fungus *Puccinia striiformis* in the past. Fungal Genetics and Biology, 87, 1-8.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I. V., Dubchak, I. 2014. The genome portal of the Department of Energy Joint Genome Institute.Nucleic Acids Res. 2014,42(1):D26-31.

Torkamaneh, D., Laroche, J., Bastien, M., Abed, A., and Belzile, F. 2017. Fast-GBS: a new pipeline for the efficient and highly accurate calling of SNPs from genotyping-by-sequencing data. BMC bioinformatics, 18(1), 5.

Tukamuhabwa, P., and Maphosa, M. 2010. State of Knowledge on Breeding for Durable Resistance to Soybean Rust Disease in the Developing World. Rome: FAO.

Tyler, B. M., and Rouxel, T. 2012. Effectors of fungi and oomycetes: their virulence and avirulence functions and translocation from pathogen to host cells. Molecular Plant Immunity.

Upadhyaya, N. M., Mago, R., Staskawicz, B. J., Ayliffe, M. A., Ellis, J. G., Dodds, P. N. 2014. A bacterial type III secretion assay for delivery of fungal effector proteins into wheat. Molecular Plant–Microbe Interactions 27: 255– 264.

Uppalapati, S. R., Ishiga, Y., Doraiswamy, V., Bedair, M., Mittal, S., Chen, J., et al. 2012. Loss of abaxial leaf epicuticular wax in *Medicago truncatula* irg1/palm1 mutants results in reduced spore differentiation of anthracnose and nonhost rust pathogens. Plant Cell 24, 353–370. doi: 10.1105/tpc.111.093104

van der Hoorn, R. A. L., and Kamoun, S. 2008. From guard to decoy: A new model for perception of plant pathogen effectors. Plant Cell 20:2009-2017.

Ve, T., Williams, S. J., Catanzariti, A. M., Rafiqi, M., Rahman, M., Ellis, J. G., Hardham, A. R., Jones, D. A., Anderson, P. A., Dodds, P. N., and Kobe, B. 2013. Structures of the flax-rust effector AvrM reveal insights into the molecular basis of plant-cell entry and effector-triggered immunity. Proceedings of the National Academy of Sciences, 110(43), 17594-17599.

Verweij, P. E., Snelders, E., Kema, G. H. J., Mellado, E., and Melchers, W. J. G. 2009. Azole resistance in *Aspergillus fumigatus*: a side-effect of environmental fungicide use? Lancet. Infect. Dis. 9, 789–795. doi: 10.1016/S1473-3099(09)70265-8

Vittal, R., Yang, H. C., and Hartman, G. L. 2012. Anastomosis of germ tubes and migration of nuclei in germ tube networks of the soybean rust pathogen, *Phakopsora pachyrhizi*. European Journal of Plant Pathology, 132(2), 163-167.

Voegele, R. T., Mendgen, K. 2011. Nutrient uptake in rust fungi: how sweet is parasitic life? Euphytica 179: 41– 55.

Wightwick, A., Walters, R., Allinson, G., Reichman, S., and Menzies, N. 2010. "Environmental risks of fungicides used in horticultural production systems," in Fungicides, ed. O. Carisse (Rijeka: InTech), 273–304.

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., Armstrong, M. R., Grouffaud, S., van West, P., Chapman, S., Hein, I., Toth, I. K., Pritchard, L. and Birch, P. R. J. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature, 450, 115–118.

Wolf, RB, Cavins, JF, Kleiman, R, Black LT 1982. Effect of temperature on soybean seed constituents: Oil, protein, moisture fatty acids, amino acids and sugars. Journal of the American Oil Chemists' Society, 59(5):230-232. doi: 10.1007/BF02582182

Yorinori, J. T., Paiva, W. M., Frederick, R. D., Costamilan, L. M., Bertagnolli, P. F., Hartman, G. L., Godoy, C. V., Nunes, J. Jr. 2005. Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. Plant Dis 89:675–677. doi: 10.1094/PD-89-0675.

Zhang, X. C., Freire, M. C. M., Le, M. H., De Oliveira, L. O., Pitkin, J. W., Segers, G., Concibido, V. C., Baley, G. J., Hartman, G. L., Upchurch, G. and Pedley, K. F., 2012. Genetic diversity and origins of *Phakopsora pachyrhizi* isolates in the United States. Asian Journal of Plant Pathology, 6(3), pp.52-65.

CHAPTER 2

POPULATION STRUCTURE AND GENETIC DIVERSITY OF PHAKOPSORA PACHYRHIZI IN THE SOUTHEASTERN UNITED STATES

Clary, J. and Smith, S. M. To be submitted to *G3: Genes, Genomes, Genetics.*

Abstract

Phakopsora pachyrhizi is an obligate biotrophic rust pathogen that causes soybean rust (SBR) on soybeans and other leguminous plants. SBR is considered one of the most devastating foliar diseases of soybeans globally, causing yield losses of up to 90% in favorable conditions. To reduce reliance on fungicides for SBR management, breeding efforts are underway to produce durable and broad-spectrum host resistance. Increased knowledge of the general population structure and diversity of SBR can assist decisions in plant breeding. The genetic structure and diversity of SBR in the southeastern United States collected over a period of 9 years was assessed. Genotype by sequencing (GBS) was used to identify single nucleotide polymorphisms (SNPs) in 49 SBR field isolates. Two distinct clusters were revealed for the *P. pachyrhizi* isolates based on K-means hierarchical clustering, DAPC, and a neighbor-joining (NJ) tree. The two clusters were present in Georgia, but Alabama, Florida, and Louisiana only contained the larger cluster. A genotype accumulation curve using approximately 54 SNPs revealed very little genetic diversity between the isolates. Isolates did not cluster by locality or year, despite clustering into two distinct populations, and there was little evidence of sexual or parasexual recombination. The results of the DAPC, neighbor joining tree, and Gst values of the two clusters provided evidence for a second more recent introduction of SBR into the southeastern US.

Introduction

Phakopsora pachyrhizi is an agronomically important biotrophic rust fungus which causes soybean rust (SBR) disease (Goellner *et. al.*, 2010). SBR threatens soybean production globally having been introduced to every soybean-producing continent in the world, with the continental

US being the last introduction in 2004 (Schneider *et al.*, 2005). In 2004, SBR was also found on Kudzu (*Pueraria lobata*) in Florida where the pathogen is thought to overwinter (Harmon *et al.*, 2005). A total of 340 million metric tons of soybean are produced globally, with the US producing over one third at 120 million metric tons (ASA, 2018). SBR has caused crop losses of more than $10 billion USD since the first epidemic in Brazil in 2001 (Yorinori *et al.*, 2005; da Silva *et al*., 2014). Though epidemics have not been as widespread in the continental US, SBR causes intermittent issues for farmers in the Southeastern US. It has been hypothesized that the reason for the less widespread problems has been due to environmental conditions and low inoculum introduction into the US (Christiano and Scherm, 2007). The main strategy for combating this disease in the US is through the use of fungicide inputs as *P. pachyrhizi* resistant varieties have not been released commercially. In countries such as Brazil, *P. pachyrhizi* has historically overcome genetically resistant varieties within a few years of deployment (Hartman *et al.*, 2005, Paul *et al.*, 2015).

Disease control through fungicide use has proven to be costly. In Brazil, three treatments per season costs the growers ~2 billion USD annually (Godoy *et al.*, 2016). Chemicals in demethylation inhibitors (DMI) and quinone outside inhibitor (QoI) classes, as well as the newer highly active succinate dehydrogenase inhibitors (SDHI) class, are an effective means of managing SBR given proper timing of application (Mueller *et al.*, 2009; Godoy, 2011; Guicherit *et al.*, 2014; Childs *et al.*, 2019). A predominant concern of fungicide usage is in the potential for pathogen populations to develop insensitivity. Fungicide insensitivity has been observed with DMIs caused by either point mutations in, or overexpression of, the *cyp51* gene (Scherm *et al.*, 2009; Barbosa *et al.*, 2013; Schmitz *et al.*, 2014; Reis *et al.*, 2015). QoI insensitivity has also been observed in *P. pachyrhizi*, as a result of a *cyt b* mutation (F129L) (Klosowski *et al.*, 2015).

Since the introduction of SDHI use for SBR in 2013, this class of fungicides has become increasingly used and at least one isolate of *P. pachyrhizi* has been found to be insensitive (Godoy *et al.*, 2015; Simões *et al.*, 2018). Widespread usage of highly effective fungicides with single modes of action causes selective pressure on a pathogen to develop resistance to the fungicide. This effect is compounded by the polycyclic nature of *P. pachyrhizi* (Bradley, 2007). Early detection and forecasting are required for efficient disease management. As a result, it is important to understand how *P. pachyrhizi* is behaving in the Southeastern US in regards to its diversity and population structure. These data will provide insight on *P. pachyrhizi's* ability to locally adapt to US environmental conditions and determine if *P. pachyrhizi* is diverging rapidly from earlier isolates identified closer to the introduction of soybean to the US, and if there have been more recent introductions of *P. pachyrhizi* into the US since 2004.

Genotyping-by-sequencing (GBS) is high throughput reduced-representation genome sequencing in which genomic DNA is digested with restriction enzymes, and short fragments are amplified and sequenced via next-generation sequencing (Elshire *et al.* 2011; Andrews *et al.* 2016). This allows for the discovery and identification of thousands of single nucleotide polymorphisms (SNPs) while reducing the costs associated with sequencing. Effectively, this means that a large number of samples can be sequenced, providing a high resolution of genetic diversity while keeping costs low. This method uses restriction endonucleases that target a small portion of the genome, making it a low cost, high throughput, and simple one step method for polymorphism discovery and genotyping (Elshire *et al.*, 2011). The predominant reason for utilizing this sequencing method over others, is that the *P. pachyrhizi* genome is large for a rust (~1 Gb), and it has large sections of repetitive DNA that have made whole genome sequencing prohibitive in the past (Link *et al.*, 2014).

The ultimate success of disease management through host plant resistance depends largely on the knowledge of pathogen diversity. Therefore, genetic improvement of SBR resistance requires comprehensive information on the genetic diversity of the pathogen. To investigate the genetic diversity and population structure of SBR, 49 *P. pachyrhizi* isolates from four southeastern states (Georgia, Florida, Alabama, and Louisiana) collected over ten years (2008-2017) were assessed with genotype-by-sequencing. Identified SNPs from each isolate were compared to the *P. pachyrhizi* reference genome for a more robust analysis. We hypothesize that *P. pachyrhizi* isolates found in the southeastern US are divergent, rich in genetic polymorphisms and have a distinct population structure.

Materials and Methods

Soybean Isolate Collection and Urediniospore Increases

Forty-nine *P. pachyrhizi* isolates were collected from multiple localities in Georgia (42 counties), Florida (two counties), Alabama (two counties), and Louisiana (two counties) from 2009 to 2017 (Table 2.1.). The collections were predominately from sentinel plots in South Georgia. Two Florida isolates, two Louisiana isolates, two Alabama isolates, as well as isolates from north Georgia were included. While the Georgia isolates are bulk field isolates, meaning that they are potentially multi genotype isolates which were collected across multiple leaves from a field, the Florida, Louisiana, and Alabama isolates are single uredial isolates maintained on a susceptible soybean variety and at least three rounds of uredial reinoculations. Florida, Louisiana, and Alabama isolates were provided by David Walker (USDA-ARS Urbana, Illinois), Patricia Bollich (Louisiana State University) and Ed Sikora (Auburn University).

Some *P. pachyrhizi* isolates in the collection did not have sufficient urediniospore mass to perform DNA extractions. Therefore, urediniospores were increased using a detached leaf assay with the universally susceptible variety, Williams 82. Three Williams 82 seeds were planted in 1 gallon pots that contained a soil mixture from SunGro Horticulture Inc. (Bellview, WA). Osmocote 15-9-12 was added on the top of the potting mix. The pots were placed in a Conviron growth chamber set at 21°C, 300 µmol* sec $^{-1}$ * m $^{-2}$ photosynthetically active radiation, and a day/night cycle of 16h/8h. Large trifoliate soybean leaves were removed from the soybean seedlings with scissors 21 days after planting for the detached leaf assay. The leaves were washed briefly in distilled water, and then the leaflets were individually cut ~1 cm from the base of the leaf to expose the vasculature. The leaflets were placed onto a 10 cm × 10 cm square petri dish with sterilized cheesecloth at the bottom saturated with sterile distilled water. Cheesecloth was folded on top of the bottom 1/8 to 1/6 of the leaflets to help keep them hydrated. The *P. pachyrhizi* urediniospore suspension was prepared by suspending 5-10 x 10^4 urediniospores mL^{-1} in ~20 mL of 0.04 % Tween20 solution and then vortexed to produce a homogenous spore suspension.

Leaflets were inoculated by brushing the spore suspension onto the abaxial side of the leaves with a paintbrush or by pipetting the spore suspension onto the leaflet and rubbing with a clean gloved finger. Paintbrush and pipette-based inoculations helped ensure inoculation consistency between samples (JIRCAS, 2019). The petri dishes with the inoculated leaflets were then individually wrapped with parafilm to prevent water evaporation and placed in a black plastic bag to block out the light. After 24 hours, the petri dishes were removed from the black plastic bag and stacked in a pyramid shape in a growth chamber with the same conditions described for the stock plants. Urediniospores were harvested from inoculated leaflets 14 to 21

days post inoculation (dpi) under a laminar flow hood with the air flow off to prevent air movement of the spores when removing them from the leaflets. Each isolate was harvested separately to prevent spore mixing. The leaflets were removed from the petri dish and gently tapped over a plastic weigh boat to dislodge urediniospores. Leaflets were also agitated with a small paintbrush to remove the remaining urediniospores. The leaflets were placed back into the petri dish, dishes were sealed with parafilm and returned to the growth chamber to continue producing urediniospores. This process continued for each SBR isolate until the leaflets stopped producing urediniospores. The harvested urediniospores were dried down in a dessicant chamber for 24 hours and then stored in 1.5 mL Eppendorf tubes at -20°C for short term storage and at -80°C for longer term storage.

Rust spore DNA extraction

DNA was extracted from each isolate in the SBR isolate collection (Table 2.1; Appendix A) using a physical and chemical lysis method developed to extract high molecular weight and high purity rust DNA (Nagar and Schwessinger, 2018). From 2-20 mg of urediniospores were used for DNA extractions. Urediniospores were placed in a precooled mortar with 1 g of autoclaved acid washed sand and liquid nitrogen and ground for 1-2 minutes. Urediniospores and sand were transferred to a 30 mL Beckman centrifuge tube with 5 mL of 2x CTAB extraction buffer, 140 µL of beta-mercaptoethanol and 10 µL of proteinase K. The solution was heated in a 65ºC water bath for 30 minutes and then placed on a benchtop shaker at 70 rpm for 60 minutes. A phase separation set followed with 5 mL of 24:1 chloroform:isoamyl alcohol, 30 minutes on the tabletop shaker at 70 rpm, and centrifugation in an Avanti Beckman centrifuge at 16,000x g and 20ºC for 5 minutes. The upper aqueous phase was transferred to a new centrifuge tube, treated

with RNase A (10 mg/mL) and placed on a benchtop shaker at 70 rpm for 75 minutes. The phase separation step was repeated, and the upper aqueous phase was transferred to a new centrifuge tube. The DNA was then precipitated with one volume of isopropanol. The resulting solution was centrifuged at 13,300x g for 10 minutes at 4ºC. The supernatant was discarded, and the resulting pellet was washed with 70% ethanol by centrifugation at 13,300x g for 5 minutes at 4ºC. The supernatant was discarded, and the ethanol wash was performed a second time. After discarding the supernatant, the DNA pellet was allowed to air dry in a laminar flow hood. The DNA pellet was resuspended in a 1x TE buffer. DNA quantity and quality was assessed for each isolate using the Nanodrop 1000 (Fischer Scientific), Qubit Fluorometer (Invitrogen), and gel electrophoresis. The DNA extracted from all *P. pachyrhizi* isolates was maintained at 4ºC until submitted for genotype-by-sequencing.

Genotype-by-Sequencing (GBS)

Genotype-by-Sequencing (GBS) provided sequence data for downstream genetic diversity analyses for 49 high quality *P. pachyrhizi* DNA samples for four Southeastern states (AL, FL, GA and LA). GBS is a method that uses enzyme-based complexity reduction. Though GBS can be used without a reference sequence, an annotated draft genome for *P. pachyrhizi* is available, making GBS even more robust for answering questions about *P. pachyrhizi* population genetics and structure, and will aid in molecular based breeding. A common restriction enzyme (*PstI*) and a rare cutter enzyme (*MspI*) was used to generate uniform libraries that consisted of a forward, barcoded, adapter and a reverse, Y, adapter on alternate ends of each fragment (Poland *et al.*, 2012). This two-enzyme method was used to reduce the complexity of the large and repetitive *P. pachyrhizi* genome. Library preparations were conducted for sequencing on an Illumina Novaseq

platform with 150 bp paired end reads. All library preparations and sequencing were performed at the Wisconsin Biotechnology Center at the University of Wisconsin. Additional DNA and library quality and quantity control measures were taken by the Wisconsin Biotechnology center to assure an accurate sequencing run.

Sequence Data Analysis

Single nucleotide polymorphisms (SNPs) were identified in the GBS data using the Fast-GBS v2 referenced based pipeline (Torkamaneh *et al.*, 2017). The annotated *P. pachyrhizi* draft genome generated by the Joint Genome Institute and Sérgio H. Brommonschenkel was used as a reference genome (*Phakopsora pachyrhizi* UFV02 Standard Draft). The sequencing reads for each isolate were aligned to the *P. pachyrhizi* draft genome. This pipeline accomplished six major tasks: 1) imported the raw sequence data, 2) demultiplexed the pooled reads, 3) filtered low quality reads and trimmed adapter sequences, 4) assembled or aligned the reads 5) identified polymorphisms at the sequenced loci, and 6) imputed missing data in the vcf file. Publicly available tools within the pipeline were used to accomplish six additional tasks: 1) Sabre (demultiplexing), 2) Cutadapt (read trimming and cleaning), 3) BWA (read mapping), 4) SAMtools (file conversion and indexing), 5) Platypus (post-processing of reads and variant calling), and 6) BEAGLE (Missing Data imputation and phasing of biallelic sites (Figure 2.1) (Sabre, 2013; Martin, 2011; Li and Durbin, 2010; Li, 2011; Rimmer *et al.*, 2014, Browning and Browning, 2007; 2016).

Population Genetic Analyses

Filtered VCF files were analyzed in R version 4.0.5 with the packages vcfR, poppr 2.0, ape, and

adegenet (Jombart and Ahmed 2011; Kamvar et al. 2015; Knaus and Grünwald 2017; R Core Team 2018; Paradis and Schliep, 2019). A Gst statistic as well as heterozygosity values were calculated across all 49 isolates. A genotype accumulation curve was calculated to determine the minimum number of SNPs needed to differentiate unique genotypes as well as how many multilocus genotypes were found among the isolates (Grünwald *et al*, 2003; Kamvar *et al.* 2015). A nei's genetic distance matrix was calculated and used to construct a neighbor-joining (NJ) distance based tree. This analysis was used to show the genetic relationships among individuals in the population (Kamvar *et al.* 2015). Gst values were then calculated to compare divergence of subsets of the isolate population, informed by the genetic distance based tree. Discriminant analysis of principal components (DAPC) and K-means hierarchical clustering were conducted to determine the optimal number of genetically distinct clusters (Jombart *et al.* 2010). Principal coordinates analysis (PCoA) was performed to assess the population structure of the *P. pachyrhizi* population. PCoA uses a distance matrix to further inform its ordination; Nei's genetic distance and Bray Curtis distances were calculated and compared. A mantel test was performed to show any relatedness between isolate location and genetic distance.

Data Availability

Supplementary file S1 contains supplemental figures and tables with descriptions. Sequence data and VCF files are available at GenBank and the accession numbers are listed in supplementary file S2.

Results and Discussion

SNP Calling, Data Filtering, and Genetic Diversity

Genotype-by-sequencing was used in the current study to analyze the genetic diversity of 49 *P. pachyrhizi* isolates collected from four southeastern states during a ten-year period. Raw sequence data contained 105,252 Mb of reads. The reads had a mean quality score of 35.41, and 91.78% of reads had a quality score of greater than 30. A quality score above 30 indicates less than 0.01% chance of inaccuracy in the base call (Ewing and Green, 1998). Each of the 49 isolates were run against BLAST and confirmed that the sequenced organism was *P. pachyrhizi* (Altschul *et al.,* 1990). The unfiltered vcf file produced after using the Fast-GBS v2 pipeline contained 139,781 SNPs. After filtering for quality, >20% missing information, biallelic sites only, indels, and the data imputation step with BEAGLE, 11,217 SNPs were retained. This SNP dataset was used for all of the analyses performed in this study. The genotype accumulation curve indicated that all 49 isolates were considered multilocus genotypes (Figure 2.2). The genotype accumulation curve generated in poppr plateaued at approximately 54 SNPs, meaning all predicted MLGs could be differentiated with 54 SNPs (Figure 2.2).

Nei's Gst and heterozygosity values for each locality were also calculated in R using the genetic_diff function in the poppr package. Gst was used as a measure of allelic frequencies across a population as a generalized statistic for Fst. For this biallelic model, the rule of thumb of 0.0-0.05, 0.05-0.15, and 0.15-0.25 representing low, moderate, and high genetic diversity, respectively, was used. The average heterozygosity across all isolates was 0.259 and Gst across all isolates was 0.064, indicating a moderate to low level of genetic diversity. Gst values could be skewed by within isolate genetic diversity as many of these are field isolates, justifying 0.064 as a low to moderate level of diversity. The observed Gst values are what would be expected with a largely clonal pathogen that rarely undergoes sexual or parasexual recombination in nature.

The common sexual stages of the *P. pachyrhizi* life cycle has not been fully characterized. *P. pachyrhizi* urediniospores have been observed in nature. However, all other spore types identified for other rust fungal pathogens have not been identified. Additionally, t is unclear if *P. pachyrhizi* can complete its lifecycle on a single host species or if it requires two or more host species to complete its lifecycle. There is evidence that *P. pachyrhizi* undergoes a parasexual cycle that involves somatic hyphal fusion and genetic recombination (Li *et al.*, 2019; Goellner, 2010). This suggests that the *P. pachyrhizi* parasexual lifecycle, in the absence of sexual stages, is a contributing factor to the moderate to low level of genetic diversity observed for the 49 *P. pachyrhizi* isolates.

The Gst were also visualized as a distribution with a violin plot (Figure 2.3). The violin plot indicated that heterozygosity values across localities as well as Gst were close to bimodal distributions. For each locality, the same bimodal distribution of heterozygosity was observed. In the Gst distribution, most alleles were in line with the reference genome, while only a small subset was fixed for alternate alleles. This subset of alternate alleles characterizes diversity in these isolates as compared to the reference genome.

Multilocus genotypes, as identified in this study, were difficult to interpret accurately. Given the population structure information from DAPC and the Neighbor-Joining analysis (Figures 2.4 and 2.5), it is unlikely that all 49 isolates should be considered their own MLGs. It's likely only two predominant genotypes are present in these data. Despite the problems with determining the exact number of MLG using the genotype accumulation curve, the number of SNPs needed to differentiate between isolates is clear. This indicates that previous studies which look at genetic diversity do not have the same statistical power to differentiate between genotypes when using one or two genes or genetic regions such as the ADP Ribosylation Factor

(ARF) or the Internal Transcribed Spacer region (ITS) (Zhang et al., 2012). This means that genetic diversity as measured in previous studies may be overlooking novel genetic diversity which would lead to a greater understanding of *P. pachyrhizi* behavior.

Population Structure and Diversity

Phylogenetic analysis using a neighbor-joining tree based on nei's genetic distance calculations groups the 49 *P. pachyrhizi* isolates into two distinct clusters that were not based year (Figure 2.4). The larger cluster contained 40 of the isolates, while the smaller cluster contained 9 of the isolates. The small cluster contained isolates from Moultrie and Attapulgus, GA from 2013 to 2016. The large cluster contained all isolates from LA, FL, AL, and all localities from Georgia that were collected between 2008 and 2017. These clusters were well-supported with 100% bootstrap values. All of the isolates in the small cluster were from South Georgia. Isolates from south Georgia were identified in the large and small clusters s, indicating that south Georgia is a potential inoculum source for *P. pachyrhizi* spread each year. *P. pachyrhizi* solates in the two clusters potentially were introduced into the US at different times. Alternatively, the smaller cluster may have been impeded in establishing itself in the US. To test this, the large and small clusters were subsampled and Gst statistics were calculated for each. The Gst values for the small and large clusters were 0.02 and 0.07, respectively. This indicates that the smaller cluster was less diverse, supporting the hypothesis of a more recent introduction or an impedence in establishment in the US. No distinct clustering was identified in this analysis based on year.

Use of the 11,217 SNPs selected with Fast-GBS v2, DAPC and k-means hierarchical clustering identified an optimal number of K=2 genetically distinct clusters (Figure 2.5). These clusters corresponded directly to the clusters formed in the neighbor-joining genetic distance

tree. Cluster 1 contained 40 isolates, while cluster 2 contained 9 isolates. All 49 isolates clustered with 100% posterior membership probability. In the DAPC analysis, other less likely K values were tested. With K=3, 4, 5, the larger cluster was separated, and the smaller cluster remained as a single cluster. The DAPC, and K-means hierarchical clustering, and neighbor-joining tree analysis provided evidence that *P. pachyrhizi* isolates separated into two distinct clusters in the southeastern United States. These two clusters have been maintained from at least 2013 to 2016 given that the smaller cluster contained isolates collected in those years. The small cluster also only contained isolates from counties in South Georgia, indicating that either this genotype may be less prevalent in the southeastern US, or supports that the smaller cluster is a more recent introduction to the US. The DAPC analysis also indicated that there was no mixing of isolates from the small and large clusters given any likely K value. This indicated that sexual or parasexual recombination did not occur between the two populations at a measurable rate. If sexual or parasexual recombination had occurred, more structure and greater genetic diversity would be expected in the *P. pachyrhizi* isolate populations.The principal coordinates analysis, created using a matrix of Nei's genetic distances, also clustered according to the same pattern observed in the NJ tree and the DAPC analysis, further supporting the two clusters (Figure 2.7). PCoA. typicallyuses a dissimilarity matrix such as Bray-Curtis, which uses dissimilarity measures between 0 and 1. There were no discernable differences in interpretation between the Bray-curtis and Nei's distance matrices (Appendix G). The first and second PCoA axes described 27% and 9% of the variance in the dataset, respectively. The large and small clusters, with 40 and 9 isolates, separated fully on the first axis of the PCoA. This study also revealed a low level of selection pressure for the *P. pachyrhizi* population in the southeastern US. High selection pressure would have been observed in these analyses as greater structure of the isolates,

as well as consistent outbreaks on farms in the southeastern US. Pathogens which undergo sexual reproduction typically exhibit higher Gst values, like *Septoria passerinii* which has a Gst value of 0.238 and has a highly structured population in the northern great plains (Lee and Neate, 2007). Other pathogens such as *Phytophtora infestans* exhibit population structure and diversity of sexually reproducing organisms and Gst values as high as 0.449 have been reported in wild populations (Grünwald *et al.*, 2001). Like *P. pachyrhizi*, predominately clonal organisms such as *Phaeoisariopsis griseola* can also exhibit Gst values as low as 0.004 (Mahuku *et al*., 2002).

Farms in the southeastern US experience yearly pockets of infection, indicating that the long-distance wind dispersal of the pathogen, environmental conditions, and the inoculum source are not sufficient to cause widespread problems each year. Low selection pressure could be the result of a combination of four variables: 1) long distance spore dispersal by wind means that *P. pachyrhizi* spores are not guaranteed to make it to any particular location, 2) *P. pachyrhizi* spore concentrations in Florida and South Georgia are maintained at low levels, 3) genetically susceptible varieties are used, therefore selection for more pathogenic isolate variants in the southeastern US is not likely, and 4) Kudzu could be a potential stable host reducing selection pressure for soybean fields. Additionally, the lower Gst values in this *P. pachyrhizi* population are in line with other predominately clonal populations, and also indicate that more generalist populations could be selected (Gillespie, 1973). Challenges with *P. pachyrhizi* outbreaks are much more widespread in Brazil, costing billions annually for control. Brazil uses a combination of genetically resistant soybean varieties as well as large fungicide inputs to control SBR (Godoy *et al.*, 2016). The genetically resistant varieties alone do not midigate SBR outbreaks indicating that selection pressure and environmental conditions allow SBR to propagate unrestricted. In 2004, when *P. pachyrhizi* was first found in the US, it was hypothesized to be windblown to the

US on a hurricane (Bonde et al., 2006). The *P. pachyrhizi* isolate genotype, represented in this study by the large cluster in the neighbor-joining tree, is more likely the genotype that arrived in the US in 2004. Isolates in the smaller cluster, which are less diverse, potentially arrived during or before 2013. Additionally, isolates in the smaller cluster are less prevalent in the southeastern US and may be due to low levels of inoculum maintained in the fields or low sampling rate. This study indicated that the two SBR genotypes arrived together in 2004 or were introduced independently in 2004 and then again during or before 2013. These two genotypes were also identified together only in southern Georgia, indicating that south Georgia is a potential source of inoculum.

Conclusion

This study demonstrated that the population structure of *Phakopsora pachyrhizi* isolates in the southeastern United States is not based on year of isolate collection, but is correlated to locality. There is a correlation between each of the two clusters of isolates found in this study and with southern Georgia. These two clusters have been found in southern Georgia together for 4 years, meaning that this could be an innoculum source for infections across the southeastern US. These findings are supported by the NJ tree, the DAPC analysis, as well as the PCoA. Low to moderate levels of genetic diversity were observed for all *P. pachyrhizi* isolates, and two distinct clusters of *P. pachyrhizi* samples were identified in the LA, AL, FL, and GA isolates. A small cluster of isolates in the neighbor-joining tree were also identified in isolates collected in later years that had diverged from the larger cluster, suggesting a potential reintroduction occurred, or the two distinct genotypes were introduced to the continental US at the same time in 2004. Separation of the two clusters is supported by the NJ tree, the DAPC analysis, K-means clustering, and the

PCoA analysis. Further, differentiating these samples requires approximately 54 SNPs, indicating that previous studies that assessed housekeeping genes and regions such as ITS are not sufficient for capturing the genetic diversity or population structure of *P. pachyrhizi*. This study also indicated that there was no evidence of selection pressure on *P. pachyrhizi* that resulted in more virulent *P. pachyrhizi* isolates in the southeastern US. It was not clear how *P. pachyrhizi's* population diversity and structure contribute to pathogen fitness or disease phenotypes. Additional analysis should be conducted to characterize *P. pachyrhizi* in a biologically meaningful way to improve plant disease management. Genotype-by sequencing is a reliable tool that provides sufficient SNPs for accurate estimation of the genetic diversity between *P. pachyrhizi* isolates. Applying this method on a larger number of *P. pachyrhizi* isolates collected globally and over a greater time period, will provide a better understanding of the population dynamics and genotypic diversity of *P. pachyrhizi*.

References

American Soybean Association 2018. https://ndsoygrowers.com/wp-content/uploads/2018/09/2018ASA-SoyStats.pdf.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. "Basic local alignment search tool." J. Mol. Biol. 215:403-410.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., and Hohenlohe, P. A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics, *17*(2), 81-92.

Barbosa, G. F., da Cruz Centurion, M. A. P., Marin, B. T., and Barbosa, G. F. (2013). Effect of reduced fungicide doses on control of soybean Asian rust and bean yield. *Interciencia* 38, 347–352.

Bonde, M. R., Nester, S. E., Austin, C. N., Stone, C. L., Frederick, R. D., Hartman, G. L., and Miles, M. R. 2006. Evaluation of virulence of *Phakopsora pachyrhizi* and *P. meibomiae i*solates. Plant Dis 90:708–716. doi: 10.1094/PD-90-0708.

Bradley, C. A. 2007. "Fungicide resistance management in soybean," in *Using Foliar Fungicides to Manage Soybean Rust*, eds A. E. Dorrance, M. A. Draper, and D. E. Hershman (Columbus, OH: Land-Grant Universities Cooperating NCERA-208 and OMAF), 57–60.

Browning, S. R., and Browning, B. L. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. The American Journal of Human Genetics, 81(5), 1084-1097.

Browning, B. L., and Browning, S. R. 2016. Genotype imputation with millions of reference samples. The American Journal of Human Genetics, *98*(1), 116-126.

Childs, S. P., Buck, J. W., and Li, Z. 2018. Breeding soybeans with resistance to soybean rust (*Phakopsora pachyrhizi*). Plant Breeding, 137(3), 250-261.

Christiano, R. S. C., and H. Scherm. 2007. "Quantitative aspects of the spread of Asian soybean rust in the southeastern United States, 2005 to 2006." Phytopathology 97.11, 1428-1433.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one, *6*(5), e19379.

Ewing, B., and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome research, 8(3), 186-194.

Paul, C., Frederick, R. D., Hill, C. B., Hartman, G. L., and Walker, D. R. 2015. Comparison of pathogenic variation among *Phakopsora pachyrhizi* isolates collected from the United States and international locations, and identification of soybean genotypes resistant to the US isolates. Plant Disease, *99*(8), 1059-1069.

Gillespie, J. H. 1973. Natural selection with varying selection coefficients–a haploid model. Genetics Research, *21*(2), 115-120.

Godoy, C. V. 2011. "*Phakopsora pachyrhizi*: the performance of soybean rust fungicides over years and regions in Brazil," in *Modern Fungicides and Antifungal Compounds VI*, eds H. W. Dehne, H. B. Deising, U. Gisi, K. H. Kuck, P. E. Russell, and H. Lyr (Braunschweig: Deutsche Phytomedizinische Gesellschaft e.V. Selbstverlag), 203–209.

Godoy, C. V., Seixas, C. D. S., Soares, R. M., Marcelino-Guimaraes, F. C., Meyer, M. C., and Costamilan, L. M. 2016. Asian soybean rust in Brazil: past, present, and future. Pesquisa Agropecuaria Brasileira 51:407–421. doi: 10.1590/S0100-204X2016000500002

Goellner, K., Loehrer, M., Langenbach, C., Conrath, U., Koch, E., and Schaffrath, U. 2010. *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. *Molecular Plant Pathology* 11:169-177. doi: 10.1111/J.1364-3703.2009.00589.X

Grünwald, N. J., Flier, W. G., Sturbaum, A. K., Garay-Serrano, E., van den Bosch, T. B., Smart, C. D., Matuszak, J. M., Lozoya-Saldaña, H., Turkensteen, L. J. and Fry, W. E. 2001. Population structure of *Phytophthora infestans* in the Toluca valley region of central Mexico. Phytopathology, *91*(9), pp.882-890.

Grünwald, Niklaus J. 2003. "Analysis of genotypic diversity data for populations of microorganisms." Phytopathology 93.6: 738-746.

Guicherit, E., Bartlett, D., Dale, S. M., Haas, H. U., Scalliet, G., and Walter, H.2014. "Solatenol-the second generation benzonorbornene SDHI carboxamide with outstanding performance against key crop diseases," in *Modern Fungicides and Antifungal Compounds VII*, eds H. W. Dehne, H. B. Deising, B. Fraaije, U. Gisi, D. Hermann, A. Mehl, *et al.* (Braunschweig: Deutsche Phytomedizinische Gesellschaft e.V. Selbstverlag), 67–72.

Harmon, P. F., Momol, M. T., Marois, J. J., Dankers, H., and Harmon, C. L. 2005. Asian soybean rust caused by Phakopsora *pachyrhizi* on soybean and kudzu in Florida. Plant Health progress, 2005, 1-4.

Hartman, G. L., Miles, M. R., and Frederick, R. D. 2005. Breeding for resistance to soybean rust. Plant Disease, *89*(6), 664-666.

Kamvar, Zhian N., Jonah C. Brooks, and Niklaus J. Grünwald. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Frontiers in genetics 6: 208.

Klosowski, A. C., May De Mio, L. L., Miessner, S., Rodrigues, R., and Stammler, G. 2015. Detection of the F129L mutation in the cytochrome b gene in *Phakopsora pachyrhizi*. Pest Manag. Sci. 72, 1211–1215. doi: 10.1002/ps.4099

Knaus, B. J., and Grünwald, N. J. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources 17.1, 44-53.

Japan International Research Center for Agricultural Sciences (JIRCAS). 2019. Laboratory manual for studies on soybean rust resistance. 1-50.
https://www.jircas.go.jp/en/publication/manual_gudeline/30.

Jombart, T., S. Devillard, and F. Balloux. 2010 Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMCgenetics 11: 94.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27: 3070e3071.

Lee, S. H. and Neate, S. M. 2007. Population genetic structure of *Septoria passerinii* in Northern Great Plains barley. Phytopathology, *97*(8), pp.938-944.

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics, *26*(5), 589-595.

Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, *27*(21), 2987-2993.

Link, T., Seibel, C., Voegele, R. T. 2014. Early insights into the genome sequence of *Uromyces fabae*. Frontiers in Plant Science 5: 587.

Mahuku, G. S., Henríquez, M. A., Munõz, J., & Buruchara, R. A. 2002. Molecular markers dispute the existence of the Afro-Andean group of the bean angular leaf spot pathogen, *Phaeoisariopsis griseola*. Phytopathology, *92*(6), 580-589.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal, 17(1), 10-12.

Mueller, T. A., Miles, M. R., Morel, W., Marois, J. J., Wright, D. L., Kemerait, R. C., Levy, C., and Hartman G. L. 2009. Effect of fungicide and timing of application on soybean rust severity and yield. Plant Dis. 93, 243–248. doi: 10.1094/PDIS-93-3-0243

Nagar, R., Schwessinger, B. 2018. High purity, high molecular weight DNA extraction from rust spores via CTAB based DNA precipitation for long read sequencing. doi: 10.17504/protocols.io.n5ydg7w

Paradis E, and Schliep K 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35, 526-528.

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J. L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one, 7(2), e32253.

R Core Team 2018. R: A language and environment for statistical computing. Vienna, Austria. URL https://www.R-project.org/.

Reis, E. M., Deuner, E., Zanatta, M., Reis, E. M., Deuner, E., and Zanatta, M. 2015. In vivo sensitivity of *Phakopsora pachyrhizi* to DMI and QoI fungicides. *Summa Phytopathol.* 41, 21–24. doi: 10.1590/0100-5405/1975.

Sabre-barcode-demultiplexing: https://github.com/najoshi/sabre. Accessed 11 Aug 2019.

Schneider, R. W., Hollier, C. A., Whitam, H. K., Palm, M. E., McKemy, J. M., Hernandez, J. R., Levy, L. and DeVries-Paterson, R. 2005. First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. Plant Disease 89:774.

Scherm, H., Christiano, R. S. C., Esker, P. D., Del Ponte, E. M., and Godoy, C. V. 2009. Quantitative review of fungicide efficacy trials for managing soybean rust in Brazil. *Crop Prot.* 28, 774–782. doi: 10.1016/j.cropro.2009.05.006

Schmitz, H. K., Medeiros, C.-A., Craig, I. R., and Stammler, G. 2014. Sensitivity of *Phakopsora pachyrhizi* towards quinone-outside-inhibitors and demethylation-inhibitors, and corresponding resistance mechanisms. *Pest Manag. Sci.* 70, 378–388. doi: 10.1002/ps.3562

Silva, A. C. da. 2014. Essential oils from Hyptis marrubioides, Aloysia gratissima and Cordia verbenacea reduce the progress of Asian soybean rust. Acta Scientiarum. Agronomy 36. 159-166.

Simões, K., Hawlik, A., Rehfus, A., Gava, F., and Stammler, G. 2018. First detection of a SDH variant with reduced SDHI sensitivity in *Phakopsora pachyrhizi*. Journal of Plant Diseases and Protection, 125(1), 21-26.

Torkamaneh, D., Laroche, J., Bastien, M., Abed, A., and Belzile, F. 2017. Fast-GBS: a new pipeline for the efficient and highly accurate calling of SNPs from genotyping-by-sequencing data. BMC Bioinformatics, 18(1), 5.

Torkamaneh, D., Laroche, J., Boyle, B., and Belzile, F. 2019. Fast-GBS.v2: An Analysis Toolkit for Genotyping-by-Sequencing data. Genome.

Wickham, H. 2016. ggplot2-Elegant Graphics for Data Analysis. Springer International Publishing. Cham, Switzerland.

Yorinori, J. T., Paiva, W. M., Frederick, R. D., Costamilan, L. M., Bertagnolli, P. F., Hartman, G. L., Godoy, C. V., and Nunes, J. Jr. 2005. Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. *Plant Dis* 89:675–677. doi: 10.1094/PD-89-0675

Zhang, X. C., Freire, M. C. M., Le, M. H., De Oliveira, L. O., Pitkin, J. W., Segers, G., Concibido, V. C., Baley, G. J., Hartman, G. L., Upchurch, G. and Pedley, K. F., 2012. Genetic diversity and origins of *Phakopsora pachyrhizi* isolates in the United States. Asian Journal of Plant Pathology, 6(3), pp.52-65.

TABLES AND FIGURES

Table 2.1. Southeastern SBR isolate collection. Forty-nine isolates used in the study including collection dates and localities.

SBR Isolate Collection						
Isolate Sample Name	**Locality**	**Year Collected**		**Isolate Sample Name**	**Locality**	**Year Collected**
GA_Attapulgus_10	Attapulgus	2010		GA_GRIF_13	Griffin	2013
GA_Attapulgus_13	Attapulgus	2013		GA_GRIF_14	Griffin	2014
GA_Attapulgus_13.1	Attapulgus	2013		GA_GRIF_14.1	Griffin	2014
GA_Attapulgus_14	Attapulgus	2014		GA_Thomas_13	Thomas	2013
GA_Attapulgus_14.1	Attapulgus	2014		GA_Thomas_13.1	Thomas	2013
GA_Attapulgus_16	Attapulgus	2016		GA_Thomas_14	Thomas	2014
GA_Attapulgus_14.2	Attapulgus	2014		GA_Thomas_14.1	Thomas	2014
GA_Attapulgus_13.2	Attapulgus	2013		GA_Thomas_14.2	Thomas	2014
GA_Attapulgus_13.3	Attapulgus	2013		GA_Thomas_14.3	Thomas	2014
GA_Attapulgus_13.4	Attapulgus	2013		GA_Thomas_16	Thomas	2016
GA_Attapulgus_14.3	Attapulgus	2014		GA_Thomas_17	Thomas	2017
GA_Attapulgus_14.4	Attapulgus	2014		GA_Tifton_14	Tifton	2014

Isolate Sample Name	Locality	Year Collected		Isolate Sample Name	Locality	Year Collected
GA_Attapulgus_16.1	Attapulgus	2016		GA_Tifton_14.1	Tifton	2014
GA_Attapulgus_14.5	Attapulgus	2014		GA_Tifton_14.2	Tifton	2014
GA_Attapulgus_14.6	Attapulgus	2014		GA_Tifton_14.3	Tifton	2014
GA_Attapulgus_14.7	Attapulgus	2014		GA_Tifton_14.4	Tifton	2014
GA_Attapulgus_14.8	Attapulgus	2014		GA_Tifton_14.5	Tifton	2014
GA_Attapulgus_16.2	Attapulgus	2016		FL_Quincy_09	Quincy	2009
GA_Moultrie_13	Moultrie	2013		FL_Quincy_12	Quincy	2012
GA_Moultrie_14	Moultrie	2014		FL_Quincy_15	Quincy	2015
GA_Moultrie_14.1	Moultrie	2014		AL_Fairhope_08	Fairhope	2008
GA_Moultrie_14.2	Moultrie	2014		AL_Elmore_14	Elmore	2014
GA_Moultrie_14.3	Moultrie	2014		LA_Houma_15	Houma	2015
GA_Moultrie_14.4	Moultrie	2014		LA_Jeanette_15	Jeanette	2015
GA_Moultrie_14.5	Moultrie	2014				

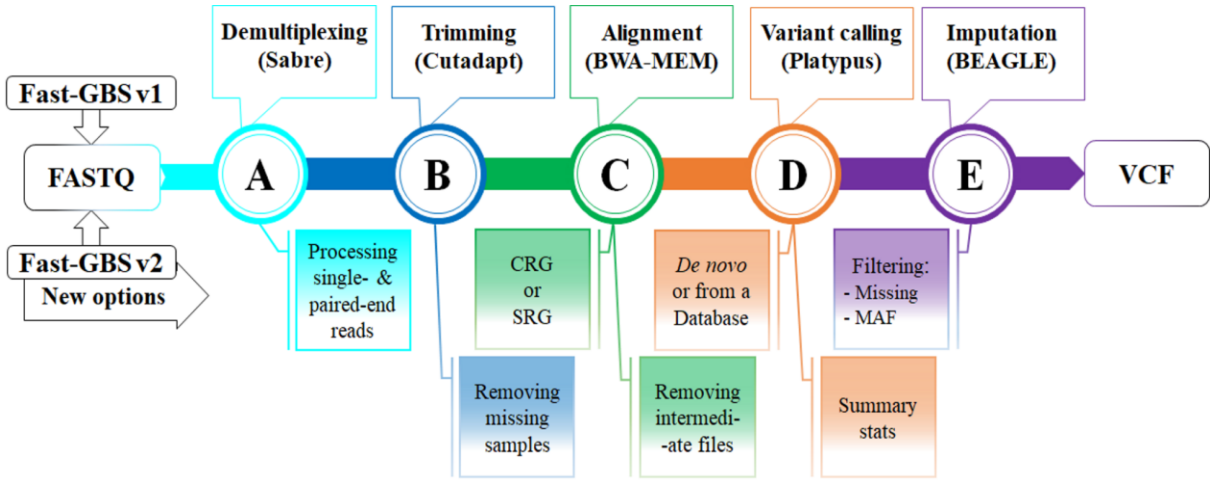


Figure 2.1. SBR isolate raw sequence data processing pipeline. A five step pipeline was used to process FastQ raw sequence data from the sequencing run to produce a VCF file for genetic diversity analysis: A) Demultiplexing, B) Trimming, C) Alignment, D) Variant Calling, and E) Data Imputation (Torkamaneh *et al.*, 2017).

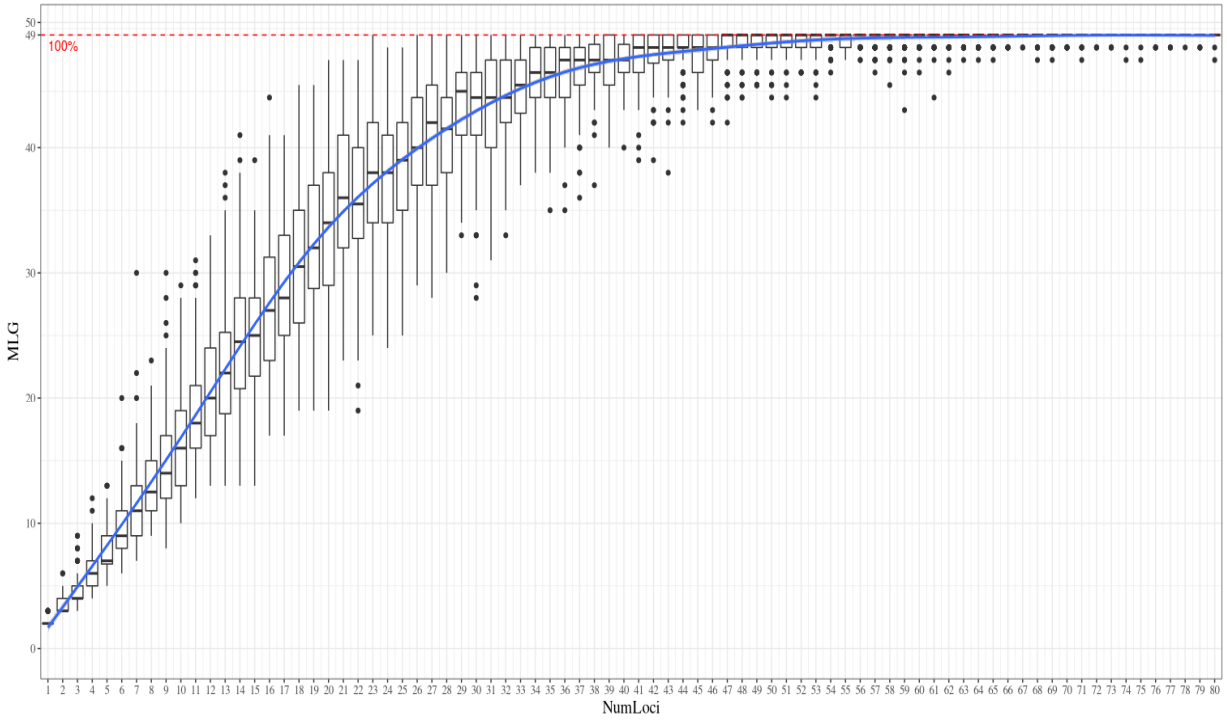


Figure 2.2. Genotype accumulation curve for 49 *P. pachyrhizi* southeastern isolates constructed in poppr (Version 2.9.1). Numbers on the x-axis indicate the number of loci. Numbers on the y-axis indicate the number of multilocus genotypes (MLG) found in the 49 samples of and the minimum number of SNPs to differentiate between each MLG.

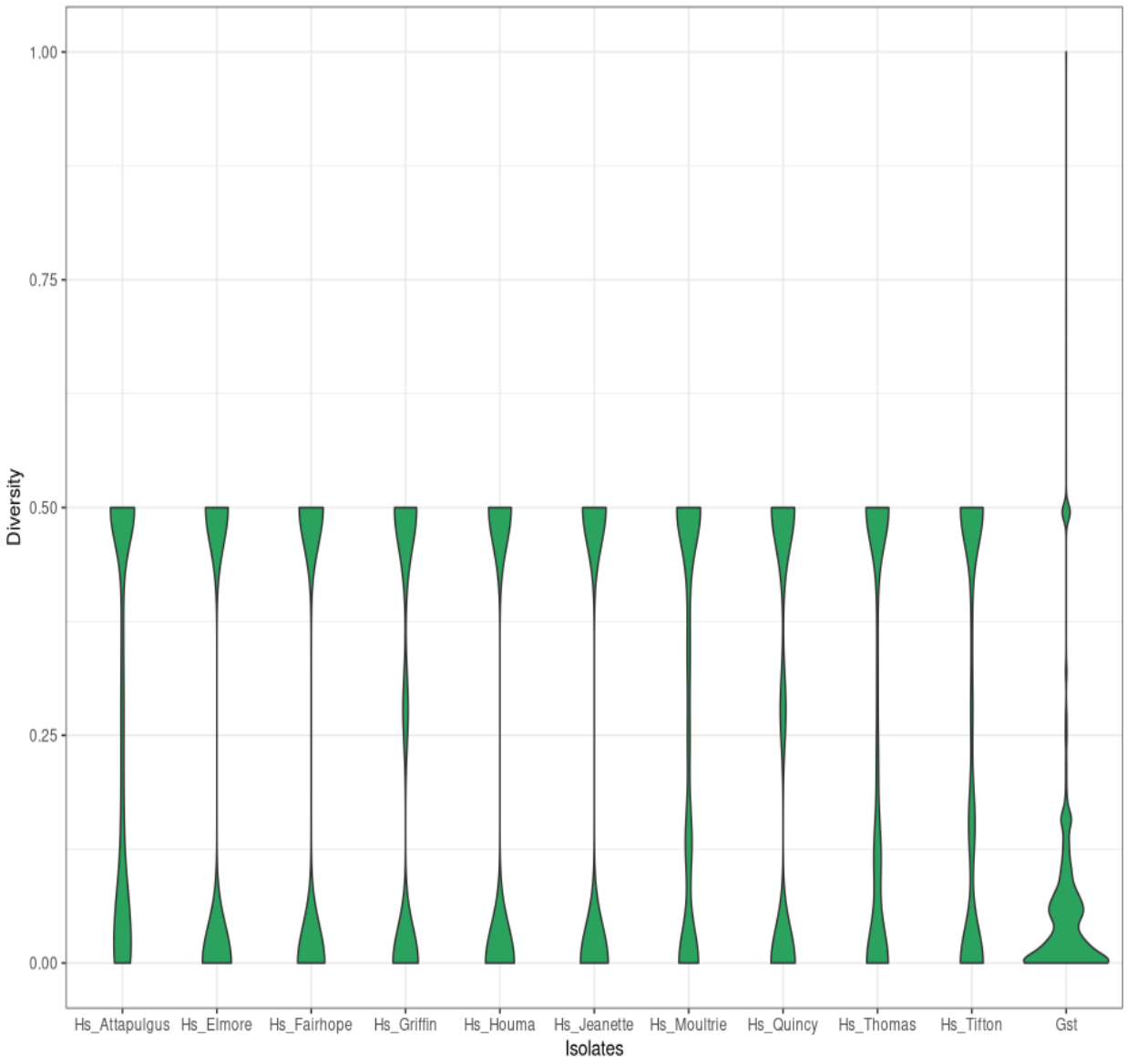


Figure 2.3. Violin plot representing heterozygosity averages for the 49 southeastern *P. pachyrhizi* isolates. The x-axis represents isolate location. The y-axis represents the heterozygosity distribution values for each locality (Hs=heterozygosity) as well as the overall distribution of genetic diversity (Nei's Gst) across the 49 samples.

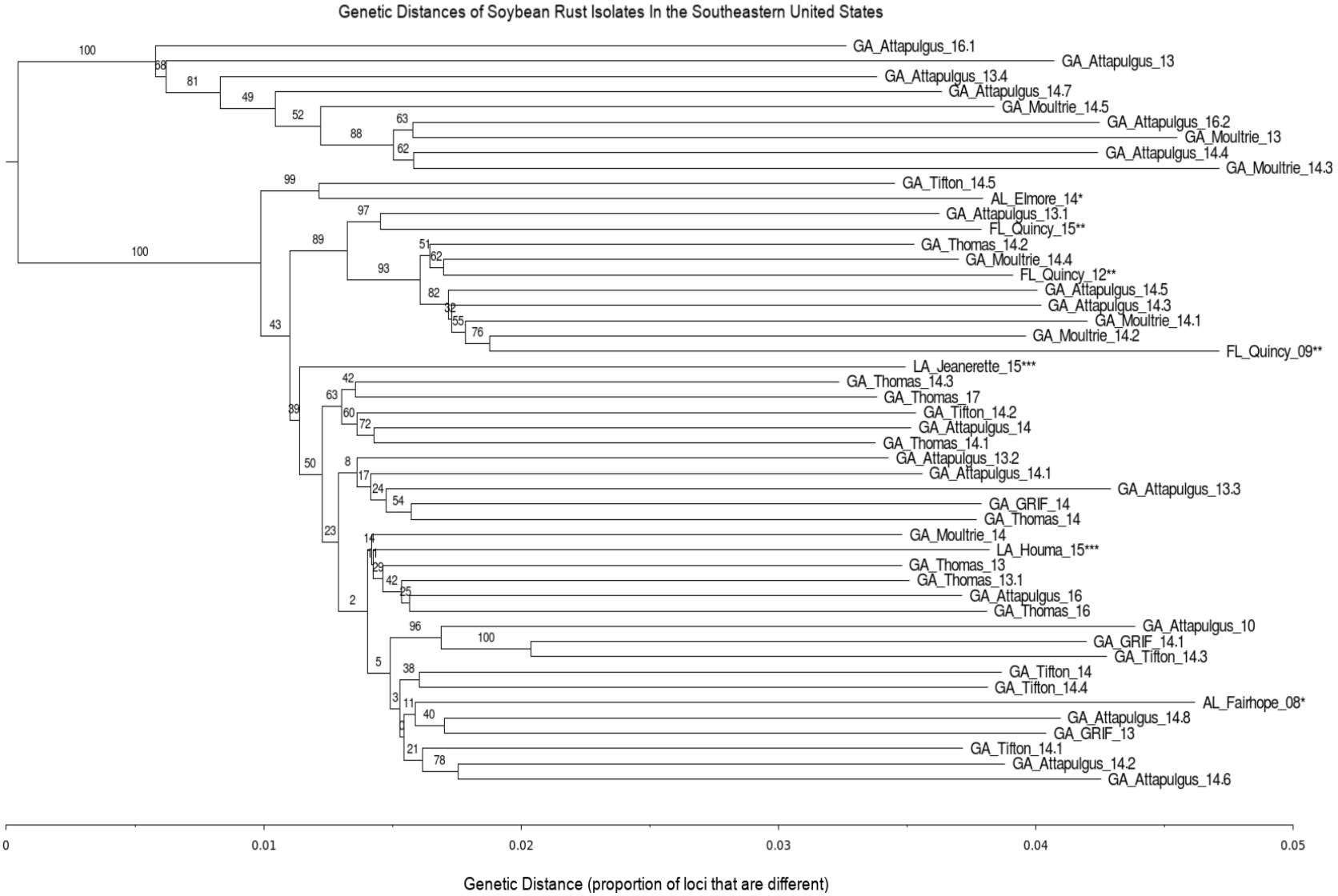

Genetic Distances of Soybean Rust Isolates In the Southeastern United States
GA_Attapulgus_16.1
GA_Attapulgus_13
GA_Attapulgus_13.4
GA_Attapulgus_14.7
GA_Moultrie_14.5
GA_Attapulgus_16.2
GA_Moultrie_13
GA_Attapulgus_14.4
GA_Moultrie_14.3
GA_Tifton_14.5
AL_Elmore_14*
GA_Attapulgus_13.1
FL_Quincy_15**
GA_Thomas_14.2
GA_Moultrie_14.4
FL_Quincy_12**
GA_Attapulgus_14.5
GA_Attapulgus_14.3
GA_Moultrie_14.1
GA_Moultrie_14.2
FL_Quincy_09**
LA_Jeanerette_15***
GA_Thomas_14.3
GA_Thomas_17
GA_Tifton_14.2
GA_Attapulgus_14
GA_Thomas_14.1
GA_Attapulgus_13.2
GA_Attapulgus_14.1
GA_Attapulgus_13.3
GA_GRIF_14
GA_Thomas_14
GA_Moultrie_14
LA_Houma_15***
GA_Thomas_13
GA_Thomas_13.1
GA_Attapulgus_16
GA_Thomas_16
GA_Attapulgus_10
GA_GRIF_14.1
GA_Tifton_14.3
GA_Tifton_14
GA_Tifton_14.4
AL_Fairhope_08*
GA_Attapulgus_14.8
GA_GRIF_13
GA_Tifton_14.1
GA_Attapulgus_14.2
GA_Attapulgus_14.6
0
0.01
0.02
0.03
0.04
0.05
Genetic Distance (proportion of loci that are different)

Figure 2.4. Neighbor-Joining distance tree of 49 *P. pachyrhizi* isolates collected from Georgia (GA), Louisiana (LA), Alabama (AL), and Florida (FL) created using the tree function in poppr and bitwise distance calculations. The tree was midpoint rooted. * indicates isolates from Alabama, ** indicates isolates from Florida, and *** indicates isolates from Louisiana. Numbers at the nodes are bootstrap values calculated with 1,000 replications.

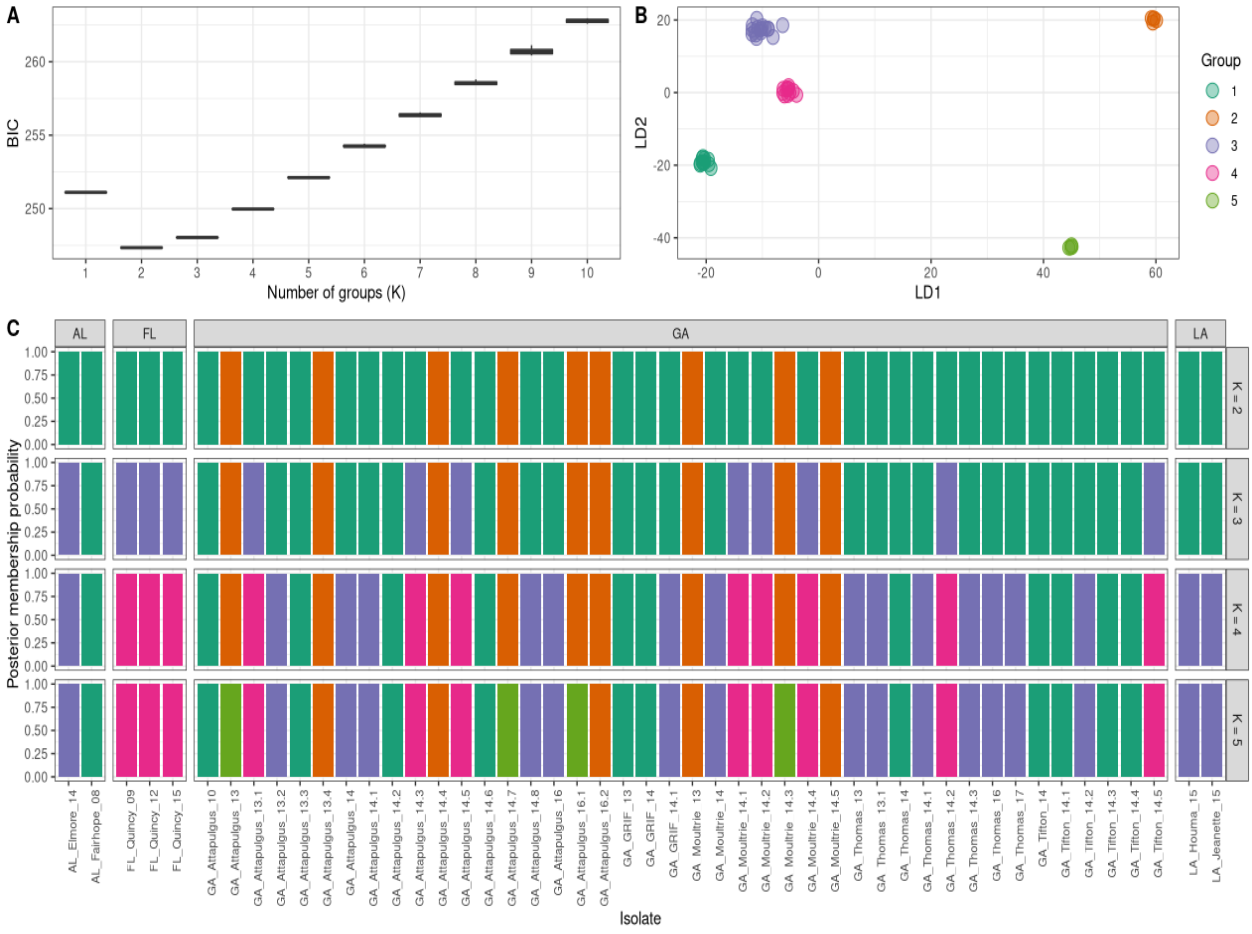


Figure 2.5. K-means hierarchical clustering and DAPC analysis for 49 *P. pachyrhizi* southeastern isolates. A) K-means hierarchical clustering with K=2 being the most likely K value. The number of K=x groups are represented on the x-axis and the y-axis represents the BIC values. This chart was interpreted by searching for a dip in BIC values, B) Likelihood values of isolates assigned to K=x populations. Each axis represents likelihood values of each K=x group from 2 to 5. The likelihood values are only positive on both axes for K=2 groups, C) DAPC analysis of posterior membership probability of each sample if K=2, 3, 4, and 5. Each row represents a different K value, as noted on the right side of the row. The rows are separated into columns based on state. The left side of the rows corresponds to posterior probability of membership information for each bar in the row. The bars in each row represent a single isolate

and all isolates were assigned 100% posterior membership probability to their respective assigned populations. The isolate names for each bar are located on the bottom x-axis.

Figure 2.6

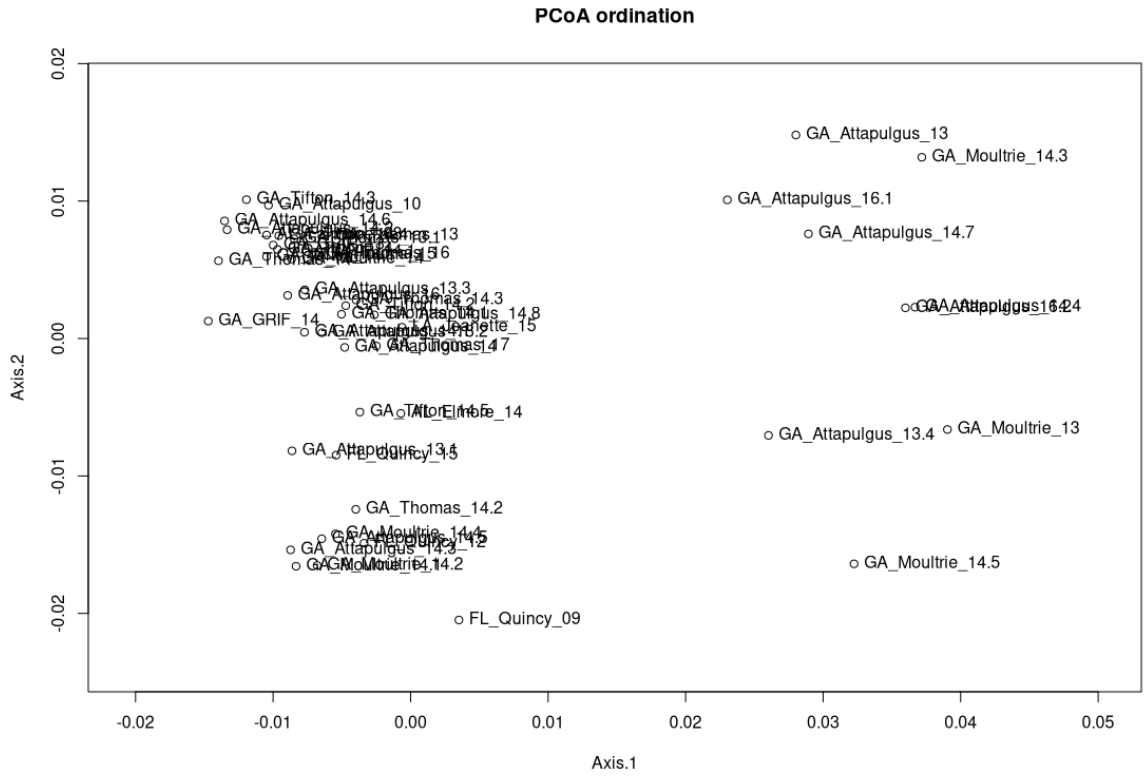


Figure 2.6. PCoA of 49 *P. pachyrhizi* southeastern isolates using nei's genetic distance matrix. Principal coordinates analysis showing population structure in poppr. The x and y axis represent ordination axes 1 and 2 which describe 27% and 9% of the variation in the isolates, respectively. Two clusters of isolates separate out on the first axis, and these populations correspond directly to the results in the DAPC and NJ tree. The larger cluster contains 40 isolates. The smaller cluster contains 9 isolates. PCoA was run using the pcoa command in the R package ape. Results were plotted using the R package biplot (Paradis and Schliep, 2019).

CHAPTER 3

SUMMARY

Phakopsora pachyrhizi, the causal agent of soybean rust (SBR), is one of the most agronomically devastating pathogen's globally. *P. pachyrhizi* can develop insensitivities to commonly used fungicides and overcome genetically resistant plant varieties. SBR has not been a widespread problem in the continental US since its introduction in 2004, but it has caused occasional yield losses for farmers across the southeastern US each year. More insight is needed with understanding SBR population structure and genetic diversity to best inform breeding strategies for the divergent selection of *P. pachyrhizi*. Understanding SBR's behavior since its introduction to the US would provide vital information for local adaptations of SBR to the US.

P. pachyrhizi isolates were collected across the Southeastern US, DNA was extracted and sequenced using GBS, and population genetics analyzed to better understand its structure and diversity. Forty-nine isolates were collected from LA, AL, FL, and GA from 2008 to 2017. SBR was not found to be structured by locality or by year of isolate collection alone. Moderate levels of genetic diversity were identified along with two distinct clusters of *P. pachyrhizi* isolates. We also found, through a neighbor-joining distance tree, that a small cluster of 9 isolates come from later years of isolate collection and have diverged from the larger cluster, indicating a potential reintroduction or that two distinct genotypes were introduced to the continental US at the same time in 2004. Further, differentiating these isolates takes approximately 54 SNPs according to a genotype accumulation curve, indicating that previous studies that assess housekeeping genes

and regions such as ITS, are not sufficient for capturing all of the genetic diversity or structure in a given SBR population.

The selection pressures imposed on SBR to continue to evolve and become more virulent in the US were not indicated in this study. It is possible that lower selection pressure was imposed on SBR in the US in comparison to other locations such as Brazil. This may be because SBR has not locally adapted to the US, the environmental conditions continue to be unfavorable for local adaptation, susceptible soybean varieties continue to be planted, and/or kudzu has become a more predominant host that stabilizes SBR's diversity in the southeastern US. It was not clear how SBR population diversity and structure contribute to pathogen fitness or disease phenotypes. An investigation should be conducted to characterize SBR in a biologically meaningful way to improve disease management strategies.

Some of the genetic diversity measured in this study could be related to the local adaptation of *P. pachyrhizi* in the southeastern United States. A study could be designed to measure if the power of divergent selection in the population of *P. pachyrhizi* outweighs the pressures imposed by gene flow and other evolutionary characteristics. A population with more gene flow relative to divergent selection, or even a population that was established more recently may be less able to locally adapt to a specific geographic area. Additionally, as the soybean cultivars grown in the US are highly susceptible to SBR, selection pressure that is imposed on the pathogen is less than in places such as Brazil which grow genetically resistant varieties. This suggests that higher selection pressures are not imposed that would drive divergent selection, and thus drive adaptive selection (Kawecki and Ebert, 2004). Additionally, the environment in the southeastern US may be hindering local adaptation of SBR, as all isolates will overwinter in areas that are above freezing conditions (Christiano and Scherm, 2007). A study could be

conducted that measures the fitness gradients of SBR found in the southeastern US and SBR found in Brazil. An approach involving experimental evolution of SBR populations in the southeastern US and Brazil utilizing traits desirable for local adaptation in the southeastern US such as lower temperatures could be a viable way to study these fitness gradients. Additionally, it would be beneficial to expose some of these isolates to proposed genetically resistant soybean varieties. It is possible that the pressures imposed by temporal variation, which select for more generalist populations, are higher than those imposed currently by the spatial variation from the Southeastern US to other more damaging populations of SBR such as in Brazil (Gillespie, 1973).

Some of the data generated was not analyzed, such as indels, which genes are being coded for that are divergent, as well as if there are rare alleles that are coding for any divergent phenotypic characteristics in SBR populations found in the southeastern US. These data could provide further insights into the behavior of SBR in the southeastern US, especially given the annotation of the draft genomes sequenced by JGI. A comparative population genomics study would be beneficial in understanding how these more recent isolates have diverged from parent populations in Brazil and Eastern Asia.

REFERENCES

Christiano, R. S. C., and Scherm, H. 2007. Quantitative aspects of the spread of Asian soybean rust in the southeastern United States, 2005 to 2006. Phytopathology, *97*(11), 1428-1433.

Gillespie, J. H. 1973. Natural selection with varying selection coefficients–a haploid model. Genetics Research, *21*(2), 115-120.

Kawecki, T. J., and Ebert, D. 2004. Conceptual issues in local adaptation. Ecology letters, *7*(12), 1225-1241.

APPENDICES

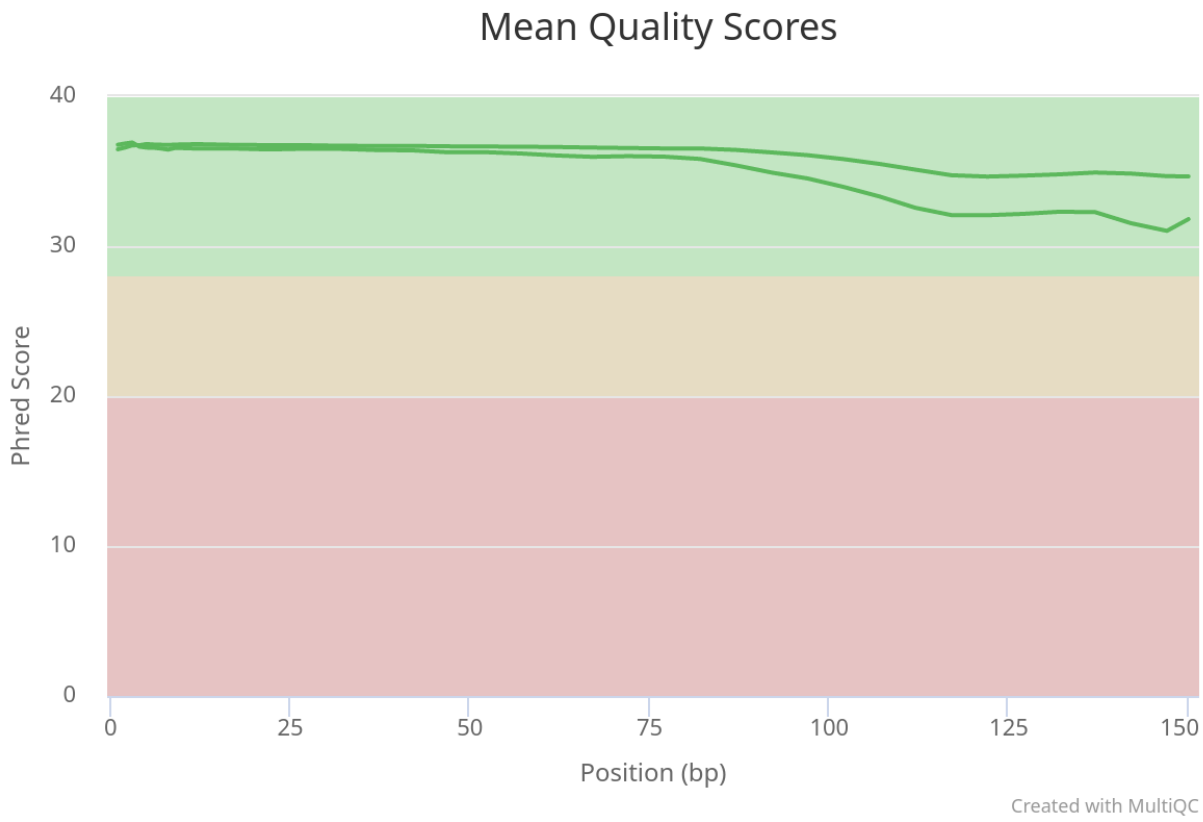
APPENDIX A. Quality and quantity of DNA obtained from 49 *P. pachyrhizi* southeastern isolates used for genotype-by-sequencing.

Isolate	Date Collected	NanoDrop Concentration	260/280	260/230
GA_Attapulgus_13	01/01/13	133	1.89	2.07
GA_Attapulgus_13.1	11/01/13	42.66	1.76	1.82
GA_Attapulgus_14	02/05/14	121.88	2	2.33
GA_Attapulgus_14.1	11/13/14	77.61	1.96	2.12
GA_Attapulgus_16	06/01/16	123.9	1.94	2.35
GA_Attapulgus_14.2	01/20/14	182	1.91	2.32
GA_Attapulgus_13.2	03/01/13	47.6	1.84	1.66
GA_Attapulgus_13.3	11/01/13	38.21	1.78	2.18
GA_Attapulgus_13.4	11/01/13	384.3	1.94	1.78
GA_Attapulgus_14.3	02/04/14	31.68	1.87	2.48
GA_Attapulgus_14.4	05/01/14	351.3	1.94	1.7
GA_Attapulgus_16.1	06/01/16	248.3	1.92	2.33
GA_Attapulgus_14.5	11/01/13	116.62	1.94	2.16
GA_Attapulgus_14.6	01/20/14	83.49	1.81	1.41
GA_Attapulgus_14.7	02/04/14	498.8	1.91	2.29
GA_Attapulgus_14.8	05/01/14	24.05	1.81	1.52
GA_Attapulgus_16.2	06/01/16	135.17	1.94	2.03
GA_Moultrie_13	11/11/13	219.1	1.95	1.95
GA_Moultrie_14	01/21/14	85.13	1.83	1.96

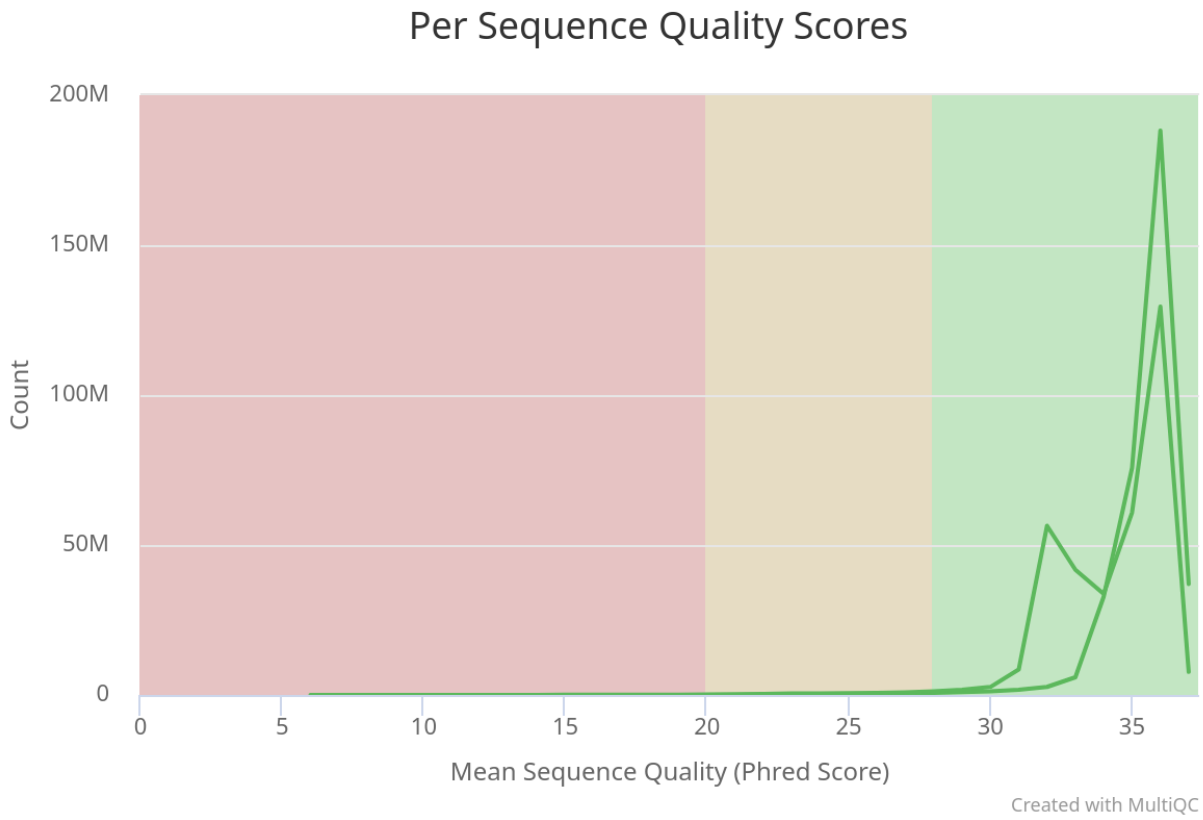
Isolate	Date Collected	NanoDrop Concentration	260/280	260/230
GA_Moultrie_14.1	02/07/14	110.52	1.95	2.36
GA_Moultrie_14.2	07/01/14	148.1	1.91	2.07
GA_Moultrie_14.3	07/03/14	345.8	1.95	2.08
GA_Moultrie_14.4	07/17/14	192	1.94	1.77
GA_Moultrie_14.5	10/31/14	242.9	1.92	1.91
GA_Griffin_13	11/01/13	32.61	1.76	1.81
GA_Griffin_14	02/04/14	69.8	1.87	2.09
GA_Griffin_14.1	07/01/14	13.6	1.61	2.15
GA_Attapulgus_10	09/01/10	45.79	1.79	1.69
GA_Thomas_13	02/01/13	37.27	1.9	2.33
GA_Thomas_13.1	11/01/13	74.83	1.96	2.12
GA_Thomas_14	01/20/14	112.6	1.94	1.82
GA_Thomas_14.1	02/04/14	126.1	1.94	1.79
GA_Thomas_14.2	05/01/14	70.26	1.81	1.69
GA_Thomas_14.3	11/13/14	51.73	1.84	1.94
GA_Thomas_16	06/01/16	64.39	1.86	2.26
GA_Thomas_17	10/01/17	28.15	1.75	1.61
GA_Tifton_14	02/05/14	225	1.92	2.29
GA_Tifton_14.1	05/01/14	51.05	1.67	1.17
GA_Tifton_14.2	06/12/14	120.68	1.88	2.14
GA_Tifton_14.3	07/03/14	232.2	1.94	2.12
GA_Tifton_14.4	07/03/14	89.68	1.9	2.02
GA_Tifton_14.5	07/17/14	24.52	1.74	1.47
FL_Quincy_09	2009	19.77	1.96	3.34
FL_Quincy_12	2012	29.39	1.89	1.68

Isolate	Date Collected	NanoDrop Concentration	260/280	260/230
FL_Quincy_15	2015	249.07	1.92	2.28
AL_Fairhope_08	2008	48.8	1.87	1.98
AL_Elmore_14	2014	37	1.85	1.96
LA_Houma_15	2015	91	1.97	1.77
LA_Jeanerette_15	2015	62	1.88	2.03

APPENDIX B. Sequence mean quality scores by position for 49 *P. pachyrhizi* southeasternisolates (green line). The x-axis represents nucleotide position, and the y-axis represents the mean sequence quality score (Phred score).



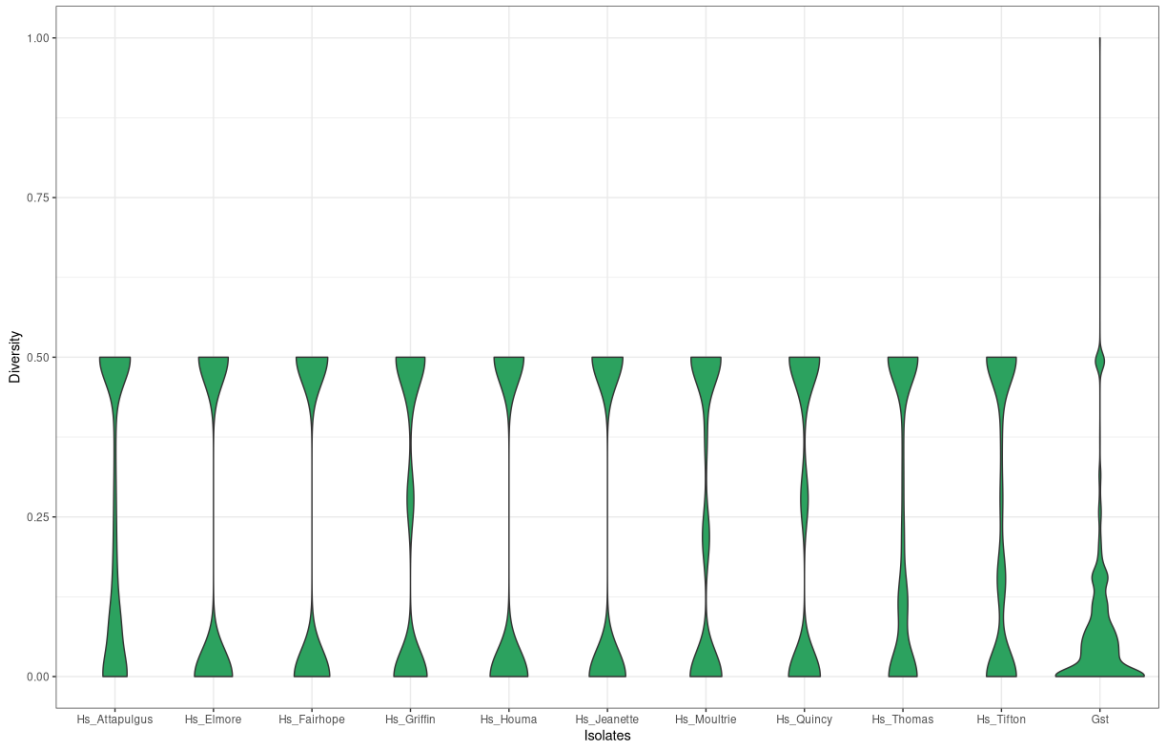
APPENDIX C. Per sequence quality scores for 49 *P. pachyrhizi* southeastern isolates. The x-axis represents the mean sequence quality score and the y-axis represents the number of nucleotides.The green peaks represent the mean quality score for the 49 *P. pachyrhizi* southeastern isolates as it relates to the number of nucleotides.



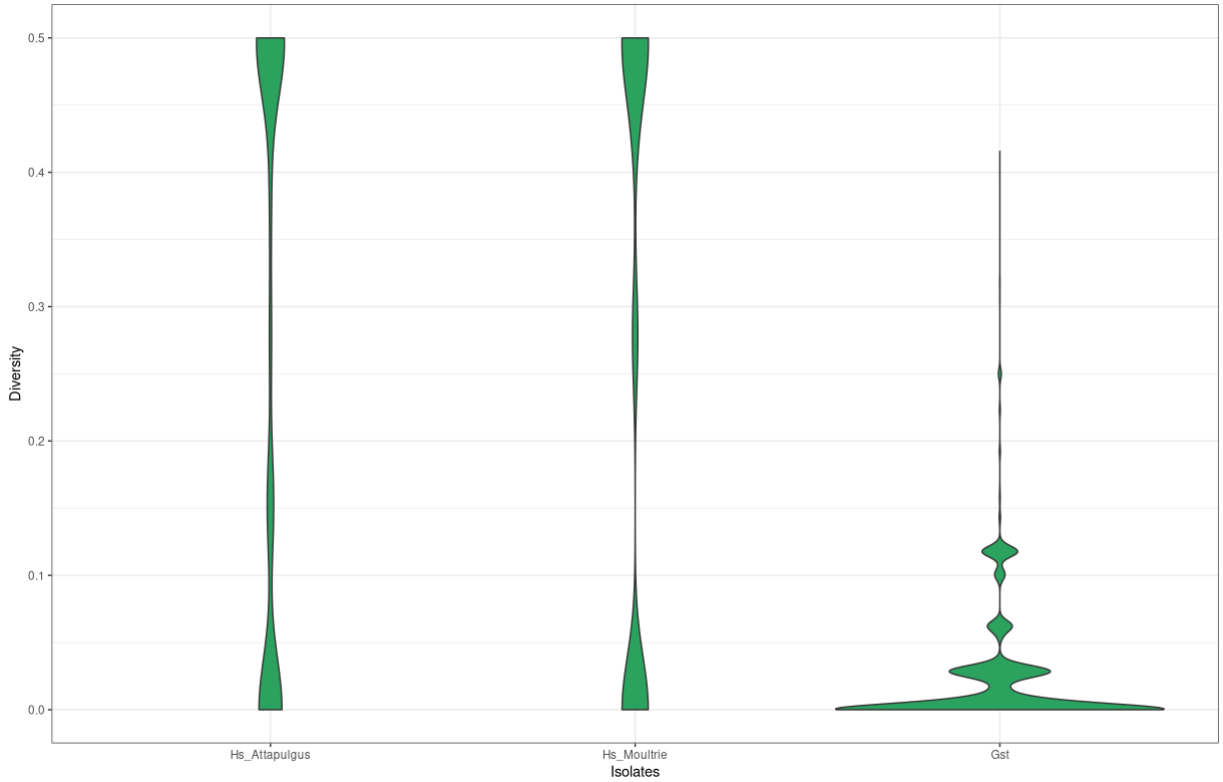
APPENDIX D. Average allelic heterozygosity proportions in each locality for the 49 *P. pachyrhizi* southeastern isolates. Ga = Georgia; Al= Alabama; Fl= Florida; La= Louisiana.

Cities Sampled	Avg Heterozygosity
Attapulgus_Ga	0.263
Griffin_Ga	0.243
Moultrie_Ga	0.264
Thomas_Ga	0.242
Tifton_Ga	0.244
Elmore_Al	0.219
Fairhope_Al	0.235
Quincy_Fl	0.252
Houma_La	0.220
Jeanerette_La	0.229

APPENDIX E. Violin plot representing heterozygosity averages for the 40 southeastern *P. pachyrhizi* isolates found in the large cluster of the NJ tree. The x-axis represents isolate location. The y-axis represents the heterozygosity distribution values for each locality (Hs=heterozygosity) as well as the overall distribution of genetic diversity (Nei's Gst) across the 40 samples.



APPENDIX F. Violin plot representing heterozygosity averages for the 9 southeastern *P. pachyrhizi* isolates found in the smaller cluster of the NJ tree. The x-axis represents isolate location. The y-axis represents the heterozygosity distribution values for each locality (Hs=heterozygosity) as well as the overall distribution of genetic diversity (Nei's Gst) across the 9 samples.



APPENDIX G. PCoA of 49 *P. pachyrhizi* southeastern isolates using Bray-Curtis dissimilarity matrix. Principal coordinates analysis showing population structure in poppr. The x and y axis represent ordination axes 1 and 2 which describe 27% and 9% of the variation in the isolates, respectively. Two clusters of isolates separate out on the first axis, and these populations correspond directly to the results in the DAPC and NJ tree. The larger cluster contains 40 isolates. The smaller cluster contains 9 isolates. PCoA was run using the pcoa command in the R package ape. Results were plotted using the R package biplot (Paradis and Schliep, 2019).