Genetic variation and structure of *Sclerotinia sclerotiorum* populations from soybean in Brazil

Anthony Pannullo, Zhian N. Kamvar, Thomas J.J. Miorini, James R. Steadman, and Sydney E. Everhart

Department of Plant Pathology, University of Nebraska, Lincoln, NE 68583

Current address of first author: Department of Microbiology and Immunology, University of Iowa, Iowa City, IA 52242

Corresponding author: S. Everhart, E-mail: everhart@unl.edu

ABSTRACT

Pannullo, A., Z.N. Kamvar, T.J.J. Miorini, J.R. Steadman, and S.E. Everhart. xxxx. Population structure of *Sclerotinia sclerotiorum* from Brazil. Phytopathology Tropical Plant Pathology xxx: xxx-xxx.

The clonal, necrotrophic plant pathogen, *Sclerotinia sclerotiorum* is the causal agent of white mold on soybean, causing significant losses for Brazilian farmers each year. While assessments of population structure and clonal dynamics can be beneficial for determining effective management strategies, few studies have been performed. In this paper, we present a broad-scale population genetic analysis with 11 microsatellite loci of 94 isolates of *S. sclerotiorum* from soybean fields in nine Brazilian states (N=74) with Argentina (N=5) and the United States

(N=15) as outgroups. Genotyping identified 87 multilocus genotypes with 81 represented by a single isolate. The pattern of genetic diversity observed suggested populations were not strongly differentiated because despite the high genetic diversity, there were few private alleles/genotypes and no multilocus genotypes were identified in both South and North America while one multilocus genotype was shared between Argentina and Brazil. Pairwise analysis of molecular variance between populations in Brazil revealed nine out of 15 pairs significantly different ($P > 0.05$). The population from the U.S. was most strongly differentiated in across all measures of population differentiation. Overall, our results found evidence for gene flow across populations with a moderate amount of population structure within states in Brazil. We additionally found shared genotypes across populations in Brazil and Argentina, suggesting that sclerotia may be transferred across states either through seeds or shared equipment. This represents the first population genetic study to cover a wide area in Brazil.

Additional keywords: white mold, Sclerotinia stem rot, soybean, *Glycine max*

INTRODUCTION

Sclerotinia sclerotiorum is a notorious plant pathogen that has been the focus of numerous studies due to the ability to infect several economically important crops. Over 400 species of plants are susceptible to *S. sclerotiorum* infection (Cubeta et al. 1997). In the United States (U.S.), loss to disease caused by *S. sclerotiorum* were estimated to be $252M per year for soybean, sunflower, canola, dry beans, and pulse crops (U.S. Canola Association, 2014). However, this is an economically important pathogen wherever soybean are grown. For example, a 2006 report from Brazil, one of the world's top soybean producers, reported a loss of 200,000 metric tons to white mold caused by *S. sclerotiorum* (Wrather et al. 2010). Host plants can be infected by *S. sclerotiorum* via a necrotrophic mechanism initiated either carpogenically via ascospores capable of infecting senescing plant parts, such as flowers, and myceliogenically via mycelium produced from sclerotia present in the soil (Kohn 1995). As nutrients become scarce, actively growing mycelium of *S. sclerotiorum* will form sclerotia, which can remain viable within the soil for 4-5 years (Kohn 1995), and up to eight years (Grogan 1979).

One strategy typically included as part of an integrated disease management approach includes use of cultivars with greater tolerance to disease and application of fungicides. Due to the primarily necrotrophic mechanism of infection by *S. sclerotiorum*, there are no known host plants with complete resistance, thus application of fungicides is recommended for effective disease control. Overuse and misuse of fungicides, however, drive concerns of fungicide resistance emergence. Risk of fungicide resistance, however, depends both on the mode of action of the fungicide and level of genetic diversity within the pathogen population exposed to fungicides. Populations that have a high amount of genetic diversity due to outcrossing are more likely to develop resistance to chemicals such as fungicides due to the increased probability of

resistance genes being present in the population (McDonald and Linde 2002). Because population diversity is directly related to adaptation, knowledge of the genetic structure of a given population should be obtained in order to avoid selection for more virulent strains and assess risk (McDonald 1997). It is also important to understand the driving biological factors that may augment genetic diversity, such as selection pressure(s) and mode of reproduction.

Over the past few decades, the population structure of *S. sclerotiorum* has been extensively documented from different host crops and from different regions in the world. Most studies reported populations are predominantly clonal in structure (Kohn 1995; Kohli and Kohn 1998; Carbone et al. 1999; Hambleton et al. 2002; Ekins et al. 2011). Although *S. sclerotiorum* is capable of both heterothallic and homothallic reproduction, evidence of extensive outcrossing has not been observed, with most reports of evidence for sexual outcrossing reported for individual fields and variably across years (Kamvar et al. 2017). Thus, evidence supports the notion that most ascocarp formation observed in fields are the product of reproduction via self-fertilization (Kohli and Kohn 1998). Many of the population genetic studies that provided the foundational work in population analysis of *S. sclerotiorum* indicated a clonal structure. Genetic analysis of these populations typically showed that most individuals represented a few multilocus genotypes (MLGs). Several studies, however, reported populations that did not conform to this expectation, such as in China (Attanayake et al. 2013), Brazil (Gomes et al. 2011), Iran (Hemmati et al. 2009), the UK (Li et al. 2009), the U.S. (Malvárez et al. 2007), and New Zealand (Carpenter et al. 1999). The exact cause for some populations being clonal, while others are not, remains unclear because it could be due to true biological events or an artifact of sampling design, marker choice, or method of statistical analysis; determining the cause is not possible across different studies.

A majority of the population studies of *S. sclerotiorum* were performed on populations in the U.S. (Lehner and Mizubuti 2017). Brazil is another major producer of economic crops commonly infected by *S. sclerotiorum*, such as soybean (Lehner et al. 2015), which makes it an ideal comparison of *S. sclerotiorum* populations. It has been hypothesized that environments with a mild winter and variety of potential hosts, such as in Brazil, could lead to greater diversity among *S. sclerotiorum* populations (Malvárez et al. 2007). Though it is not possible to directly compare results from previous *S. sclerotiorum* population studies between temperate and tropical climates due to the differences in the number of studies, types of molecular markers that have been used (Lehner and Mizubuti 2017), and differences in primer labeling and fragment analysis. A recent study of *S. sclerotiorum* collected from soybean fields in the Brazilian state of Minas Gerais found a clonal population structure (Lehner et al. 2017). This is contrary to previous reports in Brazil, which did not identify predominantly clonal populations (Gomes et al. 2011; Litholdo Júnior et al. 2011). Another study identified populations in Brazil that had high genotypic diversity, though these populations were not found to exhibit outcrossing; rather, it was hypothesized that random mutations contributed to the observed genetic variation (Lehner et al. 2015).

Given that some studies are finding clonal populations while others describe diverse populations, it is important for more studies to examine a variety of populations using identical markers in order to acquire the most complete picture of the global population structure of *S. sclerotiorum* (Lehner and Mizubuti 2017). Locations that have been previously identified as having significant genetic diversity should be of primary interest as these populations are more likely to develop novel phenotypic traits including aggressiveness, fungicide resistance, or host resistance. Moreover, diverse and sexually reproducing pathogen populations may represent the

worldwide origin. This research seeks to describe the genetic and genotypic diversity present in populations of *S. sclerotiorum* from a wide sampling of soybean fields throughout six states in Brazil. Because of Brazil's warmer climate, which allows for a greater number of hosts year-round, greater genetic diversity is expected within populations relative to other populations studied previously.

MATERIALS AND METHODS

***Sclerotinia sclerotiorum* Isolation.** We selected 94 isolates of *S. sclerotiorum* for this project, which represented multiple states in Brazil where soybean is produced and included additional isolates from the U.S. and Argentina as out-groups for comparisons. To obtain isolates from multiple states throughout Brazil, we requested monosporic sclerotia from researchers at ESALQ/USP, the University of Passo Fundo, and the Federal University of Viçosa. We also requested sclerotia from a colleague in Argentina. Isolates obtained from the same field were often collected from cull piles generated from sorting seeds after harvest and are thus considered a random sample of isolates present within the field. Isolates from the U.S. were selected from the culture collection of J.R.S. at Department of Plant Pathology – University of Nebraska – Lincoln. Among the U.S. isolates, we selected isolates across various years and locations, where one was from Colorado and the other 14 were selected from various locations in Nebraska. All isolates were collected as sclerotia from infected soybean plants. A subset of 69 isolates were initially used for MCG pairings and DNA isolated for genotyping, and then later included an additional 25 isolates from Brazil for genotyping and genetic analysis; no MCG was determined for this additional set.

Reactivation of *Sclerotinia sclerotiorum*. Sclerotia from each isolate were surface sterilized to prevent contamination in later steps. Sterilization was performed by placing sclerotia from each isolate in a 1.7 mL centrifuge tube (BioExpress, Kaysville, UT) along with 500 µL of 1:1 6% sodium hypochlorite and ddH_2O, and then vortexed for five minutes. Sodium hypochlorite was decanted and 500 µL of ddH_2O added, which was vortexed for three minutes. Sclerotia were then transferred to a new 1.7 mL centrifuge tube with 500 µL of ddH_2O, which was vortexed for three minutes. Finally, sclerotia were placed on a sterile paper towel and allowed to dry in a class II laminar flow hood for 3-5 minutes. Three sclerotia per isolate were deposited on water agar plates (16 g of Bacto agar per liter of H_2O, Becton, Dickinson and Company, Sparks, MD) and incubated in the dark at 23 °C for 48 hours. Plugs of mycelia were taken at the edge of mycelial growth and transferred to PDA plates (39 g of Potato Dextrose Agar per liter of H_2O, Becton, Dickinson and Company, Sparks, MD).

Mycelial Compatibility Groups. Isolates from the U.S. (15) and Argentina (5) were paired in all possible combinations for MCG determination. For Brazilian isolates, a subset of 49 isolates were selected and paired on DS medium in all possible combinations. Reactions of mycelial growth between isolates grown on indicator media was used to identify mycelial compatibility groups (MCGs). It was expected that compatible reactions would form between isolates with the same MCG, whereas incompatible reactions would form between isolates of different MCGs in a mutually exclusive manner, such that no two MCGs would result in compatible pairings. For pairings, Diana Sermons (DS) medium (Cubeta et al. 2001) was used, which contained 40 g malt extract broth (Sigma-Aldrich, Co., St. Louis, MO), 20 g NaCl (Cochran Road, Solon, OH), 5 g Bacto peptone (Becton, Dickinson and Company, Sparks, MD), 15 g Bacto Agar, 80 µl red food

dye (McCormick & Co., inc., Hunt Valley, MD), 80 µl yellow food dye (McCormick & Co., inc., Hunt Valley, MD) in 1 liter of ddH_2O. A 6-mm plug of each isolate from PDA plates was taken from the advancing margin of the mycelia and placed mycelium-side down in the center of a DS medium plate and incubated for 48 hours. For the MCG, a 6-mm plug of each isolate from DS medium plates was taken from the advancing margin of the mycelia and placed mycelium-side down in the center of one half of a DS medium plate. A second isolate was then placed on the opposite half of the plate, which spaced inoculation plugs approx. 2.5 cm apart. Isolates were incubated for 48 h at room temperature and MCG evaluations were performed at 5 and 10 days later by three different people. Mycelial reactions were recorded as incompatible when an apparent line of demarcation, a barrage line, and reduced growth between colonies was observed. Isolates were considered compatible if mycelium of isolates grew together, without a barrage line. When two isolates were determined to share the same MCGs, only one isolate was selected for further pairings to represent the MCG. Next, isolates from the U.S. and Argentina were paired with isolates from Brazilian. All pairings were conducted at least twice, including self-pairing.

DNA Purification. DNA of 69 isolates was purified in 2015 and DNA of the remaining 26 was purified in 2016. For the initial set of 69, DNA was purified from actively growing cultures. Three 6-mm mycelial plugs were excised from the margin of colonies grown on PDA media and were inoculated into an Erlenmeyer flask containing 100 mL potato dextrose broth (24 g of Difco Bacto Potato Dextrose Broth per liter of ddH2O, Becton, Dickinson and Company, Sparks, MD). After 3–5 days, mycelium was collected by vacuum filtration, washed with ddH2O, lyophilized, and pulverized manually in whirl-pak bags. Lyophilized mycelia were then stored in

microcentrifuge tubes at -20 °C until needed for DNA extraction. DNA from 25mg of pulverized mycelia was purified using a phenol-chloroform extraction method followed by alcohol precipitation, suspending the DNA in 200µl TE (Sambrook et al. 1989). Suspended DNA was stored at 4 °C until genotyping. DNA quantification and purity were determined by a spectrophotometric analysis, following which, DNA was stored at -20 °C until later use.

DNA from the remaining 26 isolates was purified from mycelia. To generate mycelial tissue for DNA purification, each isolate was grown on top of a cellophane membrane (0.45 µm pores, Bio-Rad, Hercules, CA) over regular PDA Petri plates. This was done to allow the mycelium to absorb nutrients through the pores of the cellophane without growth into the medium itself, thus allowing a more rapid collection of mycelium free of medium. After inoculations, isolates were incubated in the dark at 25 °C for three days. Mycelia of each isolate was harvested from the cellophane and placed into sterile 1.7 mL centrifuge tubes, which were stored at -80 °C until a total of 24 samples were ready for DNA purification. Prior to DNA extraction, tissue homogenization was performed by first transferring into 2 mL centrifuge tubes (Fisher Scientific, Hampton, NH) that contained two pre-sterilized 4 mm glass beads (Fisher Scientific, Hampton, NH) and approximately 1 mg of sterile sand (Fisher Scientific, Hampton, NH). Mycelial tissue was kept frozen, a solution of 400 µL of AP1 buffer provided in the DNeasy Plant Mini Kit (Qiagen, Germantown, MD) was added and samples then disrupted for 40 seconds using the MP FastPrep 24 (MP Biomedicals, Solon, OH). After homogenization, the remainder of the extraction was performed using the DNeasy Plant Mini Kit according to manufacturer instructions. DNA quantification and purity was determined using a spectrophotometric analysis. Concentrations ranged from 30-200 ng/µL and 260/280 ratios approximately 1.8 and greater were considered pure. DNA was stored at -20 °C for later use.

Microsatellite Genotyping. Microsatellite genotyping was performed using 11 primer pairs developed previously (Sirjusingh and Kohn 2001) to obtain a multilocus genotype (MLG) for each isolate. The PCR protocol was modified to use a third, fluorescently labeled CAG primer (CAGTCGGGCGTCATCA) that incorporates a FAM dye into the amplicon. A batch solution was made so that each PCR reaction contained 2.5 µL 5X buffer with $MgCl_2$, 0.25 µL of 10mM dNTPs, 0.5625 µL of 10 µM FAM-CAG, 0.0625 µL of 10 µM forward primer, 0.625 µL of 10 µM reverse primer, 0.125 µL of Q5 High Fidelity Hotstart Taq polymerase (New England Biolabs, Ipswich, MA), 1 µL of genomic DNA, and 7.375 µL of ddH_2O. Total volume for a single reaction was 13 µL. Thermal cycler (Mastercycler Nexus, Eppendorf) conditions were an initial denaturation for 3 min at 94 °C, following which there were 20 cycles consisting of denaturation for 30s at 94 °C, annealing for 20s at 60 °C that stepped down 0.5 °C in each cycle, and extension for 30s at 72 °C, and then 15 cycles of denaturation for 30s at 94 °C, annealing for 20s at 50 °C, and extension for 30s at 72 °C, with the final extension lasting six minutes.

Amplicons were diluted by adding 1 µL PCR product to 9 µL of sterile ddH_2O, following which, 1.5 µL of the diluted product was added to 10 µL of Hi-Di formamide and 0.3 µL of LIZ 600 size standard (ThermoFisher Scientific, Waltham, MA). Samples were denatured at 95 °C for five minutes, frozen, and then sent to the Plant-Microbe Genomics Facility (Ohio State University, Columbus, OH) for fragment analysis via capillary electrophoresis on the ABI 3730 DNA analyzer (Applied Biosystems, Carlsbad, CA).

Data Analysis. Chromatograms from capillary electrophoresis were imported and analyzed using GeneMapper software (version 4.1, Applied Biosystems, Foster City, CA). The genotype of each sample was determined based upon the expected fragment size at each locus. Data were

recorded in Excel, formatted in GenAlEx file format, and imported into R (version 3.4.3; R Core Team 2017) using the package *poppr* (version 2.6.1; Kamvar et al. 2014) for population genetic analysis. All analyses were performed in RStudio (version 1.1.383, RStudio Team, 2015).

We assessed the power of the 11 loci to accurately detect all multilocus genotypes by calculating a genotype accumulation curve in *poppr*. This curve calculates the number of multilocus genotypes (MLG) observed sampled from an increasing number (up to *n* - *1*) of loci. The resulting curve can be used to determine how many loci are needed to obtain diminishing marginal gains, which would indicate there is saturation of MLG detection and additional loci are not needed. To characterize genetic and genotypic variation within populations, several statistics were calculated. Genetic diversity for each population was analyzed using the locus_table() function present in *poppr*. This function provided the average number of alleles per locus and Nei's unbiased gene diversity, *h* (Nei 1978). We used the calculation of *h* to derive the effective number of alleles (A_e) for each population, which is calculated as $\frac{1}{r}\sum 1/(1-h_j)$ where *r* is the number of loci and h_j is the gene diversity at the j^{th} locus (Brown and Weir 1983). We identified the number of private alleles (that is, the number of alleles unique to each population) with the function private_alleles(). In addition, we calculated genotypic diversity using both the exponentiation of the Shannon-Weiner index (e^H; Shannon 1948) and Stoddart and Taylor's index (*G*; Stoddart and Taylor 1988). Both of these statistics are expected to be equal to the number of MLGs when all MLGs are equally abundant and give greater weight to less and more abundant genotypes, respectively. We calculated evenness, which is a ratio between e^H and *G* (*G* - 1)/(e^H - 1) and ranges from 0 to 1 where 1 indicates that all genotypes are equally abundant and 0 indicates one dominant genotype (E_5; Pielou 1975; Legovic 1991); these statistics were summarized using the function poppr() in *poppr*.

We used several complementary methods to characterize population structure. An analysis of molecular variance (AMOVA; Excoffier et al. 1992) was included to determine if any of the populations could be considered significantly different at the continent, country, and state levels. The AMOVA was performed using the amova() function available in the *pegas* package with 1,000 random permutations evaluate significance (version 0.10; Paradis 2010). We used a pairwise matrix of Bruvo's distance (calculated with the function bruvo.dist() in *poppr*) in place of the standard squared Euclidean distance as the input for AMOVA (Bruvo et al. 2004). This distance measure employs a stepwise mutation model, which uses the repeat motif to compare two alleles at a single locus to determine the number of mutational steps as a quantitative measure of genetic distance; combined over all loci, this represents an estimate of the genetic distance between two individuals. In addition, we calculated AMOVA between all possible pairs of populations with the above methods to assess population differentiation on a finer scale.

To provide a visual representation of the genetic distance between isolates, and to determine if the isolates were clustered into clades that corresponded with geographic proximity, we created a neighbor-joining tree and minimum spanning network from Bruvo's distance using the packages *poppr*, *ape* (version 5.0; Paradis et al. 2004), and *igraph* (version 1.1.2; Csardi, GNepusz 2006), with isolates color-coded according to the state/country of origin. The bruvo.boot() function in *poppr*, was used to calculate the genetic distances and bootstrap analysis using 1,000 replicates. We used the *poppr* function bruvo.msn() to create a minimum spanning network (MSN) from Bruvo's distance in order to visualize the relationship between the different MLGs and identify how prevalent each MLG was across populations.

Evidence of population structure was analyzed further by utilizing discriminate analysis of principle components (DAPC; Jombart et al. 2010). This method provides another way to

determine population clustering by geographic location. DAPC transforms the data into the principal components before discriminant analysis is performed, which will maximize between-group variability and minimize within-group variability. The DAPC was performed with the dapc() function in the *adegenet* package for R (version 2.1.1; Jombart 2008; Jombart and Ahmed 2011) in which we used the state/country of origin for each isolate to define groups. The number of principal components retained was determined using the function xvalDapc() available in *adegenet*, which was used with 500 replicates to determine the optimal number of PCs to prevent the overfitting of data.

Data Availability All data and scripts used for analyses are deposited in the Open Science Framework (https://osf.io/e4uph/; Pannullo et al. 2018) and GitHub (https://github.com/everhartlab/brazil-sclerotinia-2017).

RESULTS

A total of 94 isolates were used for this study, which included 75 isolates from 20 locations in six states in Brazil where soybean is grown: Goiás (18), Rio Grande do Sul (16), Paraná (15), Bahia (13), Minas Gerais (7), and Mato Grosso do Sul (5). For comparisons, we also genotyped 5 isolates from a soybean cull pile from an unspecified location in Argentina and 15 isolates from soybean in the U.S., represented by 14 isolates from 11 locations in Nebraska and 1 isolate from Colorado. Isolates in the Brazil populations were collected as sclerotia from soybean fields from 2009 to 2014 and isolates from Argentina were collected from soybean in 1996 and those from the U.S. were collected between 1977 and 2012.

Mycelial Compatibility Groups. A sub-set of 69 isolates from Brazil (49), Argentina (5), and U.S. Midwest (15) were paired on DS medium, which identified 25 MCGs (Figure 1). Close to

half (11) of these MCG were represented by a single isolate and, therefore, were incompatible with all other isolates. The remaining 14 MCG were represented by two or more isolates, meaning that 84.1% of the 69 isolates tested were compatible with at least one other isolate. The single most abundant MCG was represented by 16 isolates (23.18% of isolates tested), which were isolates collected in 10 locations in 5 states in Brazil. Two isolates from Brazil and one isolate from Nebraska were in the same MCG; no other isolates from different countries were compatible.

Genotyping and Population Structure. Isolates were genotyped at 11 SSR loci and yielded an amplicon at no fewer than 9 of those loci (Table 1). The number of alleles per locus varied from two to 12, with an average of six alleles per locus. Within the entire data set, a total of 83 multilocus genotypes (MLG) were identified (Table 2), wherein 6 were clonal MLG. For subsequent analyses, we defined populations according to their state of origin in the case of the six populations from Brazil, the Midwest region of origin for isolates from Nebraska and Colorado in the U.S. and country of origin for Argentinian isolates. Calculating the effective number alleles showed several populations had an A_e value greater than two (Table 2), which included only those five populations lacking duplicate genotypes (ie. no clones). Populations from Brazil had 19 private alleles among the six populations, with an average of 3.17 private alleles per population and the greatest number observed in the Paraná population (A_p = 6). The only other population with greater number of private alleles was the Midwest, which had 7 private alleles. Gene diversity within populations was lowest in Minas Gerais (h = 0.235) and highest in Rio Grande do Sul (h = 0.579), with an average gene diversity among all populations

of 0.470. The average gene diversity for the Brazilian populations was 0.454, which is lower than the Midwest (h = 0.472).

Most populations had high values of genotype diversity, which was observed for Rio Grande do Sul, Paraná, Bahia, Mato Grosso do Sul, and Argentina (Table 2). In these instances, all MLG are equally abundant and thus necessarily have an E_5 = 1.0. The Midwest population—where one MLG represented 20% of the population sample—had the lowest genotypic diversity as determined by having the greatest difference in number of MLG (11) and estimated genotype diversity (G = 6.43 and e^H = 8.77) and evenness (E_5 = 0.698). To determine whether genotypic variation was structured within, between, or among the eight populations, we used an AMOVA (Table 3). Although this showed the majority of variation (69.0%) was explained by differences between individual genotypes (Error), significant differentiation between populations and continents was detected. There was 10.7% of genetic variation partitioned at the continent-level ($\phi_{CT} = 0.107, p < 0.001$) and 15.1% variation between states ($\phi_{SR} = 0.151, p = 0.043$). Although 7.91% of genetic variation was partitioned between countries ($\phi_{RC} = 0.0886$), this was not significant (p = 0.999). Pairwise comparison of Brazilian populations showed nine of the 15 pairs were significantly different (Table 4). The six pairs that showed no significant difference ($p > 0.05$) were: Goiás:Bahia ($\phi = 0.02$; p = 0.304), Goiás:MinasGerais ($\phi = 0.064$; p = 0.215), Goiás:Mato Grosso do Sul ($\phi = 0.143$; p = 0.066), Mato Grosso do Sul:Paraná ($\phi = 0.054$; p = 0.248), Mato Grosso do Sul:Bahia ($\phi = 0.173$; p = 0.061), and Mato Grosso do Sul:Rio Grande do Sul ($\phi = 0.030$; p = 0.300).

Examination of relationships between MLG was performed using a minimum spanning network, which shows connections between MLG that are most similar (Fig. 2). There were six clonal MLG found across the eight populations. Two of the six clonal MLG did not cross

345 populations, one of which was the largest clonal MLG (N = 5) found in the Midwest population.
346 The other four clonal MLG were found in Brazil and crossed between two to three regions:
347 Bahia:Minas Gerais, Goiás:Minas Gerais, Paraná:Bahia:Rio Grande do Sul, and Goiás:Rio
348 Grande do Sul. Most of these MLG were not abundant, represented by only 2 to 3 isolates. A
349 lower proportion of MLG in populations from South America (8.9%) were clonal compared to
350 North America (26.6%). Despite this, the average genetic distance between isolates from the
351 Midwest was 0.355, while the average genetic distance within Brazilian populations was 0.402.
352 Reticulations in the network, which indicated multiple equivalent paths between MLG, were
353 present only among MLGs from Brazilian populations. No MLG from the same population
354 formed an exclusive cluster. The population with the maximum number of connected MLG was
355 the Midwest population, which had seven directly connected MLG representing a total of 11
356 isolates. However, these MLG were on the periphery of the network and only two of these
357 connections represented a single mutational step. The most densely connected populations were
358 Minas Gerais and Goiás, but nevertheless, MLGs were spread throughout the network with no
359 clear clustering pattern based on population. This pattern was also observed in the dendrogram
360 (Fig. 3), where there were only 7 samples that were in clades with bootstrap support > 75%: five
361 from the same clone in the Midwest (266, 265, 264, 202, 143) and two from separate MLG in
362 Paraná (1014) and Argentina (293A), which differed by a missing allele at a single locus.
363 Structure within and between populations was characterized using a DAPC to visualize
364 variance between populations (Fig. 4). To prevent over-fitting the model, we calculated the
365 DAPC with the first seven principal components that captured 66.1% of the variation and then
366 retained the first four discriminant functions. We observed considerable overlap between South
367 American populations, while the Nebraska population was most distinct from all other

368 populations and separated from the other populations on the first discriminant axis, which
369 represented 59.1% of the discriminatory power.

370 Membership probability for each isolate as determined by the DAPC was used to further
371 visualize structure within and between these populations (Fig. 5). Genotypes of isolates from the
372 Midwest population had the highest average membership probability (74.0%) to their own
373 population where nine of the 15 isolate genotypes in the Midwest population had a re-assignment
374 probability greater than 92% and another four had a re-assignment probability greater than or
375 equal to 50%. Genotypes from the Minas Gerais population had the second highest average
376 membership probability (54.0%) to their own population, with three isolates having greater than
377 70% membership probability to Minas Gerais. Paraná also had two isolates with a membership
378 probability greater than 70% to their own state, with an average probability of 42.0%. All other
379 states in Brazil had fewer than two isolates with a membership probability greater than 70% for
380 belonging to their state of origin. The average probability of membership for these states were:
381 Argentina (36.1%), Goiás (34.7%), Rio Grande do Sul (30.0%), Bahia (20.6%), and Mato
382 Grosso do Sul (20.0%). In the South American populations, no isolate genotypes had a
383 membership probability greater than 86% to any one population. The Bahia population showed
384 two isolate genotypes had greater probability of assignment (>70%) to the Midwest population
385 than to any South American population, which was the only instance of isolates with a majority
386 of their assignment probability belonging to the North American population.
387

388 **DISCUSSION**

389 We sought to describe the structure of populations of *Sclerotinia sclerotiorum* from soybean
390 across a broad region of southern Brazil, with additional populations from Argentina and the
391 U.S. using microsatellite genotyping and MCG profiling. Our approach to assess the genetic and

genotypic diversity of Brazilian *S. sclerotiorum* populations used analyses performed at three levels in population hierarchy: continent > country > state. The analyses selected for this study allowed us to test hypotheses on the reproductive mode (genetic diversity) as well as the clonal dynamic across regions (genotypic diversity).

The number of MCGs identified in the sub-set of 69 isolates in our study was typical of other MCG profiling performed on *S. sclerotiorum* previously. For example, most isolates in our study were compatible with at least one other and nearly one quarter of all isolates were identified as having the same MCG. On the presumption that MCG is an indirect measure of population variation, it was unexpected for our study to find one isolate from Nebraska was in the same MLG as two isolates from Brazil. Previous studies have attempted to associate MCG and MLG directly in analyses, however, our previous population study of more than 350 *S. sclerotiorum* isolates showed no association between MCG and MLG (Kamvar et al. 2017). This may be due to the inconsistent nature of MCG typing, both in the method and as a biological phenomenon that is not free of affects by environment. For analyses of population structure, the utility of MCG is limited compared to molecular markers.

Genotyping 94 *S. sclerotiorum* isolates in the present study identified high genotypic diversity and few clonal MLGs. The highest proportion of clonality was observed in the Midwest population which had 11 MLG and one clonal MLG represented by five isolates. In contrast, within populations from Brazil and Argentina, there were few to no clonal MLG identified. Interestingly, among the six clonal MLG identified in Brazil, four of those were found in more than one population and one was found in three populations. In two of those cases, the populations with shared genotypes were from contiguous states, whereas in the other two, including the genotyping that was found in three states, none were contiguous. Overall, the

genetic and genotypic data suggested that there is gene flow between populations, especially those in South America, and lack of strong population differentiation.

Calculating genetic distance between isolates using Bruvo's stepwise model allowed us to utilize the quantitative power of the microsatellite loci in further analyses. There was evidence of significant population differentiation between continents and populations. There was no evidence for differentiation between countries, which may simply be due to this comparison evaluating the difference between the country (Midwest, Brazil, Argentina) level, where Argentina was composed of few samples and had a high degree of similarity to populations in Brazil, as shown in the MSN and DAPC analyses.

Our assessment of population differentiation also showed significant differentiation between states. Pairwise comparison between populations in Brazil showed that there were significant differences between all populations, with the exception of six pairs of five states. Two of those six pairs (Bahia:Rio Grande do Sul and Goiás:Minas Gerais) were also states that had one clonal MLG shared between them. Mato Grosso do Sul, which was present in the other four pairs, shared no MLG with any other state, but its samples were spread across the MSN (Fig. 2) and the tree (Fig. 3). This is further evidence that there is weak differentiation between these populations. All but one of the states (Minas Gerais) had at least two private alleles (Table 2). The presence of private alleles is normally a good indicator of population differentiation, however none of these alleles were at high frequency in the populations (Attanayake et al. 2013). Our $\phi_{ST} = 0.287$, which is directly comparable to the ϕ_{ST} from both Lehner et al. (2015); $\phi_{ST} = 0.270$ and Gomes et al. (2011); $\phi_{ST} = 0.288$. These studies concluded that while there is population structure in Brazil, there still exists gene flow among populations. This is supported by the fact that genotypes are shared among populations in Brazil (Fig. 2).

Because there were few private alleles within populations and both the MSN and DAPC showed little evidence of segregation and structure among populations, it was unexpected to find that the majority of population pairs in Brazil were significantly different in AMOVA. The minimum spanning network created from Bruvo's genetic distance showed a more complicated pattern of genetic differentiation (Fig. 2). This finding was similar to previous minimum spanning networks presented in the literature, there appeared to be minimal clustering based on region (Lehner et al. 2015; Dunn et al. 2017). Moreover, edges representing smaller distances were near the center of the network and edges representing larger distances around the periphery, suggesting little differentiation between populations; the opposite of what we would expect with significant population differentiation (Attanayake et al. 2013). This result combined with the long terminal branch lengths, lack of strong population-level clustering, and low bootstrap support in the neighbor-joining tree (Fig. 3) suggest that no populations are purely clonal or sexual. The structure observed was likely due to a combination of clonal expansion and random mutation within some genotypes. While the first discriminant axes of the DAPC shows the clear separation of the Midwest from all other populations, the second axis does not show a random cluster, but rather a gradient from Argentina to Minas Gerais. While physically adjacent states are adjacent to each other in this gradient, this is merely a coincidence. The presence of shared genotypes across populations may also suggest transfer of material between populations, which could occur via human-mediated inoculum transfer, such as in infected seed (Botelho et al. 2013).

The question of reproductive mode for *S. sclerotiorum* populations was of interest because sexually outcrossing populations have a greater potential to adapt to fungicide pressure due to the greater number of unique genetic combinations (Milgroom 1996). Several measures

461 can be employed that can serve as proxies for detecting the presence of sexual reproduction; one
462 of the most widely used measures is the index of association, which measures multilocus linkage
463 disequilibrium (Brown and Weir 1983; Smith et al. 1993; Haubold et al. 1998). Several studies
464 of *S. sclerotiorum* have employed the use of this index and have found some populations to be
465 sexual (Aldrich-Wolfe et al. 2015) and some populations to be clonal (Lehner et al. 2015).
466 However, simulation studies have shown that sampling small numbers of individuals from
467 disparate populations can result in values of this index similar to those from sexually reproducing
468 populations (Prugnolle and De Meeus 2010). Thus, we concluded that we simply did not have
469 the power to be able to detect the presence of random mating.

470 While this study covers a broad area of Brazilian soybean production, we are limited due
471 to small sample sizes. This is mitigated by the fact that we have used microsatellite markers
472 common to other studies that can allow us to combine and harmonize these data in the future
473 (Gomes et al. 2011; Litholdo Júnior et al. 2011; Lehner et al. 2015, 2017). While we did include
474 isolates gathered from the U.S. and Argentina, the samples from Brazil were over-represented,
475 which prevented us from making any conclusions about the observed differences between
476 countries. Despite this, these samples gave us a way to assess the magnitude of differentiation
477 between Brazilian populations. This contrasts with other studies of Brazilian populations, which
478 focused on intensive sampling at geographically proximal populations (Gomes et al. 2011;
479 Litholdo Júnior et al. 2011; Lehner et al. 2015, 2017). While we found significant evidence for
480 genetic structure between populations, there appeared to be different levels of within-population
481 variation. Overall, our results show comparable levels of diversity with previous studies (Gomes
482 et al. 2011; Lehner et al. 2015), suggesting that geopolitical barriers do not accurately delineate
483 *S. sclerotiorum* populations. Future studies with larger sample sizes will be necessary to

effectively elucidate pathways of migration and diversity within Brazil. Moreover, any future study should attempt to include isolates from this and previous studies in Brazil so that the data may be harmonized and re-used, giving a broader picture of *S. sclerotiorum* across continents and around the world.

The presence of variability in genetic structure within and among populations in Brazil may suggest that migration of genotypes is indeed happening across growing regions. This is supported by the fact that other studies on Brazilian populations with larger sample sizes have reached the same conclusions (Lehner et al. 2015, 2017). Our analysis thus gives a glimpse at the broad genetic structure of *S. sclerotiorum* across Brazil, complementing the previous studies assessing local structure.

ACKNOWLEDGMENTS

We thank Rebecca Higgins and Rachana Jhala for assistance in the mycelial compatibility assays and curation of the culture collection. This work was funded in part by grant #58-5442-2-209 from the USDA-ARS National Sclerotinia Initiative to JRS/SEE and start-up funds from the University of Nebraska-Lincoln (UNL) to SEE. Additional financial support was provided by the Agricultural Research Division of the Institute of Agriculture and Natural Resources at UNL to AP. Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This project is based on research that was partially supported by the Nebraska Agricultural Experiment Station with funding from the Hatch Act (Accession Number 1007272) through the USDA National Institute of Food and Agriculture.

LITERATURE CITED

507 Aldrich-Wolfe L, Travers S, Nelson BD (2015) Genetic variation of *Sclerotinia sclerotiorum* from multiple
508 crops in the North Central United States. PLoS One 10:e0139188. doi:
509 10.1371/journal.pone.0139188

510 Attanayake RN, Carter PA, Jiang D, et al (2013) *Sclerotinia sclerotiorum* populations infecting canola from
511 China and the United States are genetically and phenotypically distinct. Phytopathology 103:750–
512 761. doi: 10.1094/PHYTO-07-12-0159-R

513 Botelho L da S, Zancan WLA, Machado J da C, Barrocas EN (2013) Performance of common bean seeds
514 infected by the fungus *Sclerotinia sclerotiorum*. J Seed Sci 35:153–160. doi: 10.1590/S2317-
515 15372013000200003

516 Brown AHD, Weir BS (1983) Measuring genetic variability in plant populations. Dev Plant Genet Breed
517 1:219–239. doi: 10.1016/B978-0-444-42226-2.50016-5

518 Bruvo R, Michelis N, D'Sousa T, Schulenberg H (2004) A simple method for calculation of microsatellite
519 genotypes irrespective of ploidy level. Mol Ecol 13:2101–2106

520 Carbone I, Anderson JB, Kohn LM (1999) Patterns of descent in clonal lineages and their multilocus
521 fingerprints are resolved with combined gene genealogies. Evolution (N Y) 53:11–21. doi:
522 10.1111/j.1558-5646.1999.tb05329.x

523 Carpenter MA, Frampton C, Stewart A (1999) Genetic variation in New Zealand populations of the plant
524 pathogen *Sclerotinia sclerotiorum*. New Zeal J Crop Hortic Sci 27:13–21. doi:
525 10.1080/01140671.1999.9514075

526 Csardi, GNepusz T (2006) The igraph software package for complex network research. InterJournal
527 Complex Sy:1695. doi: 10.1109/ICCSN.2010.34

528 Cubeta MA, Cody BR, Kohli Y, Kohn LM (1997) Clonality in *Sclerotinia sclerotiorum* on Infected Cabbage
529 in Eastern North Carolina. Phytopathology 87:1000–1004. doi: 10.1094/PHYTO.1997.87.10.1000

530 Cubeta MA, Sermons DN, Cody BR (2001) Mycelial interactions of *Sclerotinia minor*. Phytopathology
531 91:S19. doi: 10.1094/phyto.2001.91.6.s1

532 Dunn AR, Kikkert JR, Pethybridge SJ (2017) Genotypic characteristics in populations of *Sclerotinia*
533 *sclerotiorum* from New York State, USA. Ann Appl Biol 170:219–228. doi: 10.1111/aab.12330

534 Ekins MG, Hayden HL, Aitken EAB, Goulter KC (2011) Population structure of *Sclerotinia sclerotiorum* on
535 sunflower in Australia. Australas Plant Pathol 40:99–108. doi: 10.1007/s13313-010-0018-6

536 Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance infered from metric distances
537 among DNA haplotypes: application to human mitochondrial DNA restricyion data. Genetics
538 131:479–491. doi: 10.1007/s00424-009-0730-7

539 Gomes EV, Breseguello L, Augusto M, et al (2011) Microsatellite Markers Reveal Genetic Variation within
540 Sclerotinia sclerotiorum Populations in Irrigated Dry Bean Crops in Brazil. J Phytopathol 159:94–99.
541 doi: 10.1111/j.1439-0434.2010.01724.x

542 Grogan RG (1979) *Sclerotinia* species: summary and comments on needed research. Phytopathology
543 69:908–910

544 Hambleton S, Walker C, Kohn LM (2002) Clonal lineages of *Sclerotinia sclerotiorum* previously known
545 from other crops predominate in 1999-2000 samples from Ontario and Quebec soybean. Can J

Plant Pathol 24:309–315. doi: 10.1080/07060660209507014

Haubold B, Travisano M, Rainey PB, Hudson RR (1998) Detecting linkage disequilibrium in bacterial populations. Genetics 150:1341–1348

Hemmati R, Javan-Nikkhah M, Linde CC (2009) Population genetic structure of Sclerotinia sclerotiorum on canola in Iran. Eur J Plant Pathol 125:617–628. doi: 10.1007/s10658-009-9510-7

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart T, Ahmed I (2011) adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. doi: 10.1093/bioinformatics/btr521

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genet 11:94. doi: doi: 10.1186/1471-2156-11-94

Kamvar ZN, Amaradasa BS, Jhala R, et al (2017) Population structure and phenotypic variation of *Sclerotinia sclerotiorum* from dry bean (*Phaseolus vulgaris*) in the United States. PeerJ 5:e4152. doi: 10.7717/peerj.4152

Kamvar ZN, Tabima JF, Grünwald NJ (2014) *Poppr* : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. doi: 10.7717/peerj.281

Kohli Y, Kohn LM (1998) Random association among alleles in clonal populations of Sclerotinia sclerotiorum. Fungal Genet Biol 23:139–149. doi: 10.1006/fgbi.1997.1026

Kohn LM (1995) The clonal dynamic in wild and agricultural plant-pathogen populations. Can J Bot 73:1231–1240. doi: 10.1139/b95-383

Legovic T (1991) Statistical ecology. A primer on methods and computing. John Wiley & Sons

Lehner MS, De Paula Júnior TJ, Del Ponte EM, et al (2017) Independently founded populations of *Sclerotinia sclerotiorum* from a tropical and a temperate region have similar genetic structure. PLoS One 12:e0173915. doi: 10.1371/journal.pone.0173915

Lehner MS, Mizubuti ESG (2017) Are *Sclerotinia sclerotiorum* populations from the tropics more variable than those from subtropical and temperate zones? Trop Plant Pathol 42:61–69. doi: 10.1007/s40858-016-0125-1

Lehner MS, Paula Júnior TJ, Hora Júnior BT, et al (2015) Low genetic variability in *Sclerotinia sclerotiorum* populations from common bean fields in Minas Gerais State, Brazil, at regional, local and micro-scales. Plant Pathol 64:921–931. doi: 10.1111/ppa.12322

Li Z, Wang Y, Chen Y, et al (2009) Genetic diversity and differentiation of *Sclerotinia sclerotiorum* populations in sunflower. Phytoparasitica 37:77–85. doi: 10.1007/s12600-008-0003-6

Litholdo Júnior CG, Gomes E V., Lobo Júnior M, et al (2011) Genetic diversity and mycelial compatibility groups of the plant-pathogenic fungus *Sclerotinia sclerotiorum* in Brazil. Genet Mol Res 10:868–877. doi: 10.4238/vol10-2gmr937

Malvárez G, Carbone I, Grünwald NJ, et al (2007) New populations of *Sclerotinia sclerotiorum* from lettuce in California and peas and lentils in Washington. Phytopathology 97:470–483. doi:

10.1094/PHYTO-97-4-0470

McDonald BA (1997) The population genetics of fungi: Tools and techniques. Phytopathology 87:448–453. doi: 10.1094/PHYTO.1997.87.4.448

McDonald BA, Linde C (2002) The population genetics of plant pathogens and breeding strategies for durable resistance. Euphytica 124:163–180. doi: 10.1023/A:1015678432355

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. Annu Rev Phytopathol 34:457–477. doi: doi:10.1146/annurev.phyto.34.1.457

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Pannullo A, Kamvar ZN, Miorini TJJ, et al (2018) Data and Analysis for Genetic variation and structure of *Sclerotinia sclerotiorum* populations from soybean in Brazil

Paradis E (2010) pegas : an R package for population genetics with an integrated – modular approach. Bioinformatics 26:419–420. doi: 10.1093/bioinformatics/btp696

Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290. doi: 10.1093/bioinformatics/btg412

Pielou EC (1975) Ecological Diversity. Wiley & Sons, New York

Prugnolle F, De Meeus T (2010) Apparent high recombination rates in clonal parasitic organisms due to inappropriate sampling design. Heredity (Edinb) 104:135–140. doi: 10.1038/hdy.2009.128

R Core Team (2017). R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria URL http//wwwR-project.org/ R Foundation for Statistical Computing

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor laboratory press

Shannon CE (1948) A mathematical theory of communication. SIGMOBILE Mob Comput Commun Rev 5:3–55. doi: http://doi.acm.org/10.1145/584091.584093

Smith JM, Smith NH, O'Rourke M, Spratt BG (1993) How clonal are bacteria? Proc Natl Acad Sci 90:4384–4388. doi: 10.1073/pnas.90.10.4384

Stoddart JA, Taylor JF (1988) Genotypic diversity: Estimation and prediction in samples. Genetics 118:705–711

Wrather A, Shannon G, Balardin R, et al (2010) Effect of diseases on soybean yield in the top eight producing countries in 2006. Plant Heal Prog doi 10:2008–2013. doi: 10.1094/PHP-2010-0125-01-RS

Tables and Figures

Table 1. Origin and year of collection for 94 *Sclerotinia sclerotiorum* isolates from soybean used in the present study

Country	State/Region	Year(s) collected	N
Brazil	Goiás	2012, 2010	18
	Rio Grande do Sul	2012	16
	Paraná	2012, 2014	15
	Bahia	2012, 2010	13
	Minas Gerais	2009	7
	Mato Grosso do Sul	2012	5
Argentina	—	1996	5
U.S.	Midwest*	1977, 1990, 1994, 1996, 1999, 2012, 2009	15

* 14 isolates from Nebraska and one isolate from Colorado

Table 2. Variation in allelic and genotypic diversity statistics of eight populations from Brazil, U.S., and Argentina, which are ordered by sample size

				Allelic				Genotypic		
Continent	Country	Population	N	Alleles	A_p	A_e	*h*	e^H	G	E_5
S. America	—	—	79 (72)	5.36	33	2.22	0.525	68.9	64.3	0.932
S. America	Brazil	—	74 (67)	5.18	25	2.16	0.512	64.0	59.5	0.929
S. America	Brazil	Goiás	18 (17)	2.73	4	1.79	0.408	16.7	16.2	0.970
S. America		Rio Grande do Sul	16 (16)	3.27	3	2.49	0.579	16.0	16.0	1.00
S. America		Paraná	15 (15)	3.18	6	2.73	0.564	15.0	15.0	1.00
S. America		Bahia	13 (13)	2.64	4	2.01	0.461	13.0	13.0	1.00
S. America		Minas Gerais	7 (6)	1.64	0	1.41	0.235	5.74	5.44	0.937
S. America		Mato Grosso do Sul	5 (5)	2.18	2	2.35	0.479	5.00	5.00	1.00
S. America	Argentina	—	5 (5)	2.27	2	3.20	0.561	5.00	5.00	1.00
N. America	U.S.	Midwest*	15 (11)	3.00	7	2.01	0.472	8.77	6.43	0.698
Pooled	—	—	94 (83)	6.00	—	2.40	0.563	77.0	66.9	0.868

N: Number of isolates with number of haplotypes in parentheses
Alleles: mean number of alleles
A_e: Effective number of alleles
A_p: number of private alleles
h: Nei's 1978 Gene Diversity
e^H: Expected Number of Genotypes (exponent of Shannon's Index)
G: Expected Number of Genotypes (Stoddard and Taylor's index)
E_5: Evenness
* 14 isolates from Nebraska and one isolate from Colorado
—: Not Applicable

Table 3. Analysis of Molecular Variance (AMOVA) table.

Levels	df	SSD	σ	P	% variation	$\boldsymbol{\phi}$
Continent within Total	1	1.41	0.0131	0.000	10.7	$\phi_{CT} = 0.107$
Country within Continent	1	0.415	0.00964	0.999	7.91	$\phi_{RC} = 0.0886$
Population within Country	5	1.28	0.0150	0.043	12.3	$\phi_{SR} = 0.151$
Error	86	7.30	0.0841	-	69.0	$\phi_{ST} = 0.310$

Df = degrees of freedom

SSD = sum of squared differences (Bruvo's Distance)

σ = variance

P = p-value based on 999 permutations

$\boldsymbol{\phi}$ = Hierarchical $\boldsymbol{\phi}$ statistic

Table 4. Pairwise AMOVA for Brazilian States. Lower Triangle shows **ϕ** values while upper triangle shows p-values based on 1000 replications. Stars indicate significance level.

	Goiás	Paraná	Bahia	Rio Grande do Sul	Minas Gerais	Mato Grosso do Sul
Goiás	-	0.000**	0.304	0.011*	0.215	0.066
Paraná	0.166**	-	0.014*	0.044*	0.000**	0.248
Bahia	0.020	0.155*	-	0.023*	0.037*	0.061
Rio Grande do Sul	0.144*	0.096*	0.118*	-	0.018*	0.300
Minas Gerais	0.064	0.374**	0.244*	0.207*	-	0.007**
Mato Grosso do Sul	0.143	0.054	0.173	0.030	0.369**	-

** = $p < 0.01$

* = $p < 0.05$

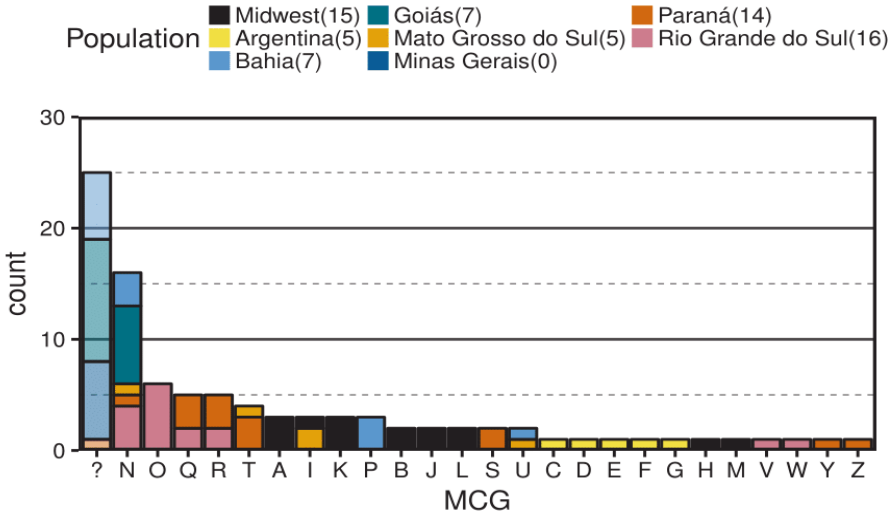

Figure 1. Frequency distribution of the 25 MCG identified (A-W, Y, Z) among 69 isolates characterized in this study. The 26 remaining isolates with no characterized MCG are shown as "?". Bar color/shade indicates population of origin. Numbers in legend refer to the number of samples tested in each population.

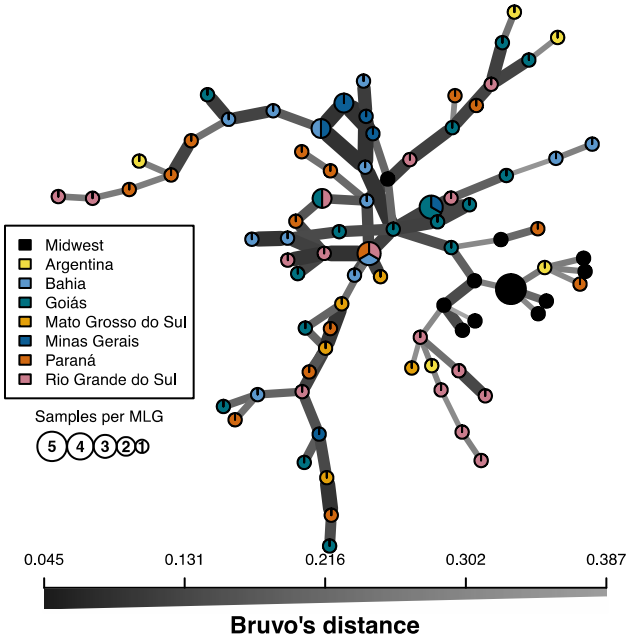

Figure 2. Minimum spanning network created using Bruvo's genetic distance representing the pairwise relatedness between multilocus genotypes (MLGs). Each individual MLG is represented by a node. The size of the node is relative to the number of isolates that possess the MLG. The node color corresponds to the origin of the individuals with that MLG, where MLG containing isolates with multiple origins are represented as pie charts scaled proportionally. The genetic distance between MLGs is represented by the thickness of the line joining the nodes, lines that are darker and thicker indicate shorter genetic distances, while the thinner, lighter lines represent larger genetic distances. Loops within the network represent multiple, equidistant paths between genotypes.

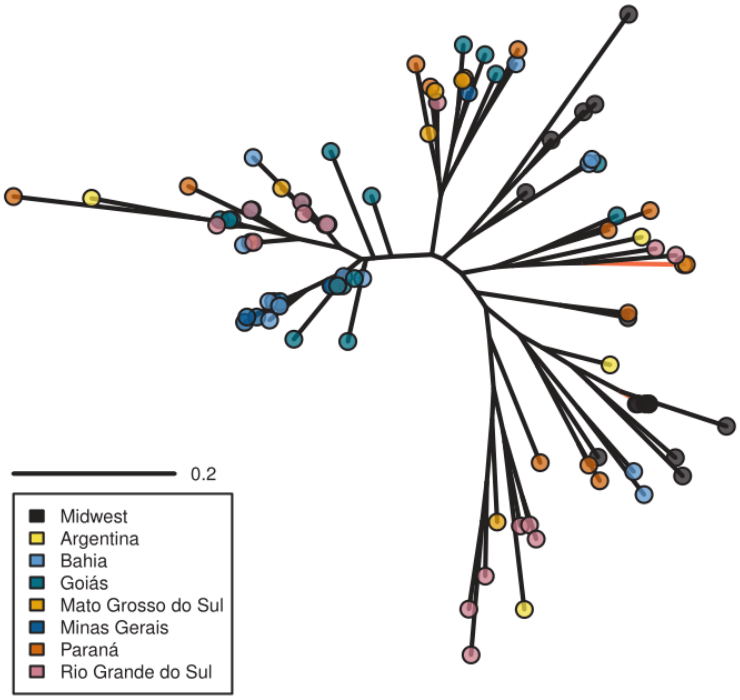

Figure 3. Unrooted, Neighbor-joining tree derived from Bruvo's distance for the 95 94 samples in the data, colored by population. Terminal nodes are jittered to display presence of duplicate genotypes. Shaded branches represent >75% bootstrap support.

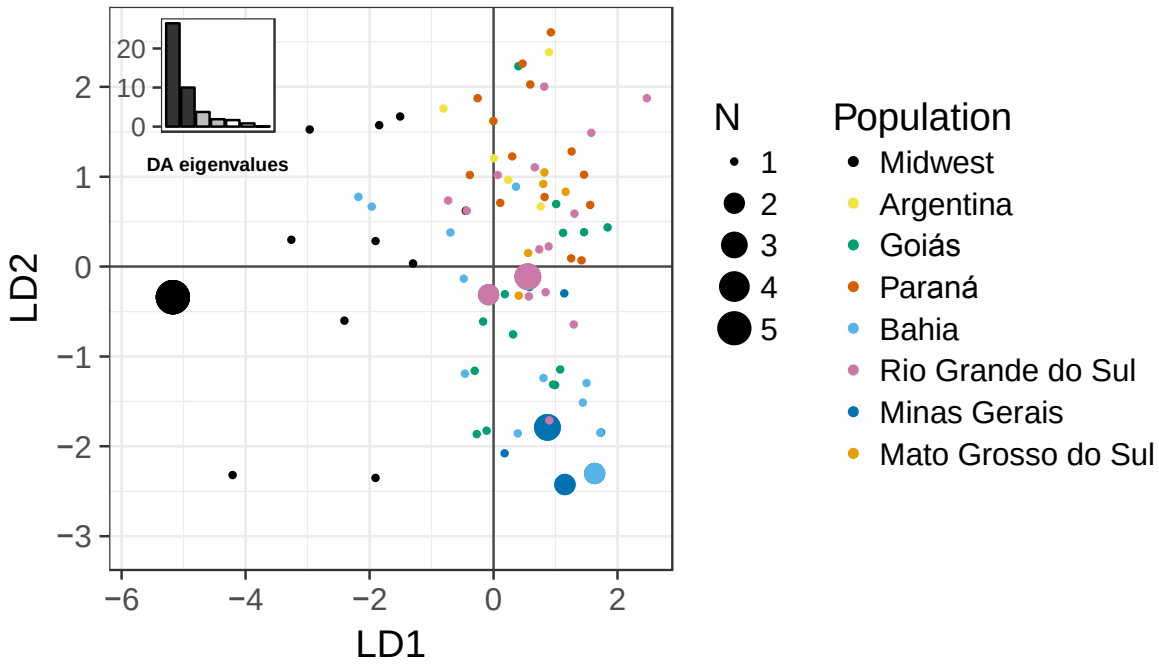

Figure 4. Biplot displaying the first two axes of a discriminant analysis of principal components (DAPC) based upon geographic population. Each dot represents an individual isolate (identical MLGs overlap and are indicated with larger dots) while the ellipses represent a 67% confidence interval for group membership for each of the populations based on a bi-variate normal distribution. Inset barplot represents the eigenvalues of the first six discriminant axes.

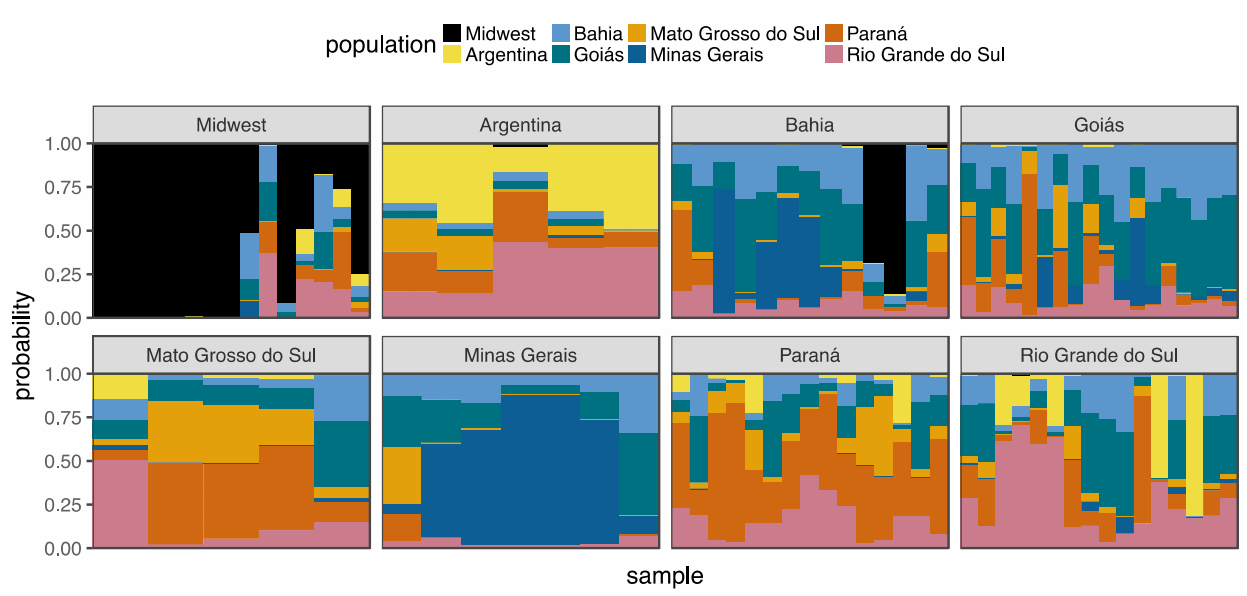

Figure 5. A stacked barplot showing the posterior population assignment probability for each individual based on the first four discriminant axes of the discriminant analysis of principal components (DAPC), representing 66.2% of the total observed variance in the data.