

# **Population structure and phenotypic variation of *Sclerotinia sclerotiorum* from dry bean (*Phaseolus vulgaris*) in the United States**

**Zhian N. Kamvar<sup>1</sup>, Bimal Sajeewa Amaradasa<sup>1</sup>, Rachana Jhala<sup>1</sup>, Serena McCoy<sup>1</sup>, James Steadman<sup>1</sup>, and Sydney E. Everhart<sup>1</sup>**

**<sup>1</sup>Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE 68583**

Current address of second author: Plant Pathology Department, University of Florida, Gainesville, FL 32611

Current address of third author: Nebraska Center for Virology, University of Nebraska-Lincoln, Lincoln, NE 68583

Corresponding author:

Sydney E. Everhart<sup>1</sup>

Email address: everhart@unl.edu

## **ABSTRACT**

The ascomycete pathogen *Sclerotinia sclerotiorum* is a necrotrophic pathogen on over 400 known host plants, and is the causal agent of white mold on dry bean. Currently, there are no known cultivars of dry bean with complete resistance to white mold. For more than 20 years, bean breeders have been using white mold screening nurseries with natural populations of *S. sclerotiorum* to screen new cultivars for resistance. It is thus important to know if the genetic diversity in populations of *S. sclerotiorum* within these nurseries a) reflect the genetic diversity of the populations in the surrounding region and b) are stable over time. Furthermore, previous studies have investigated the correlation between mycelial compatibility groups (MCG) and multilocus haplotypes (MLH), but none have formally tested these patterns. We genotyped 366 isolates of *S. sclerotiorum* from producer fields and white mold screening nurseries surveyed over 10 years in 2003–2012 representing 11 states in the United States of America, Australia, France, and Mexico at 11 microsatellite loci resulting in 165 MLHs. Populations were loosely structured over space and time based on analysis of molecular variance and discriminant analysis of principal components, but not by cultivar, aggressiveness, or field source. Of all the regions tested, only Mexico (n=18) shared no MLHs with any other region. Using a bipartite network-based approach, we found no evidence that the MCGs accurately represent MLHs. Our study suggests that breeders should continue to test dry bean lines in several white mold screening nurseries across the US to account for both the phenotypic and genotypic variation that exists across regions.

## **INTRODUCTION**

*Sclerotinia sclerotiorum* (Lib.) de Bary is an ascomycete plant pathogen with a worldwide distribution (Bolton et al., 2006). This is a necrotrophic pathogen that is primarily homothallic (self-fertilization) and has the ability to survive for more than five years in soil using melanized survival structures called sclerotia (Bolton et al., 2006; Sexton et al., 2006). It causes disease on more than 400 plant species belonging to 75 families (Boland & Hall, 1994) including crops of major economic importance such as sunflower (*Helianthus spp.*), soybean (*Glycine max L.*), canola (*Brassica napa L.*, *Brassica campestris L.*), and dry bean (*Phaseolus vulgaris L.*) (Bolton et al., 2006).

On dry bean, *S. sclerotiorum* is the causal agent of white mold, a devastating disease that can be yield-limiting in temperate climates (Steadman, 1983). All above-ground tissues (flowers, stems, leaves, pods) are susceptible to infection, first appearing as wet lesions with white mycelial tufts, and then bleaching as the tissue senesces (Steadman, 1983; Bolton et al., 2006). For many years, white mold has been the most serious dry bean disease in the Northwestern United States (Otto-Hanson et al., 2011; Knodel et al., 2012, 2015, 2016). The impact of white mold on the dry bean industry in the Northwestern United States

45 alone has been estimated at a loss of 140 kg/ha with just 10% disease incidence (Ramasubramaniam et al.,  
46 2008).

47 Currently, there are no commercially available resistant cultivars of dry bean (Otto-Hanson et al.,  
48 2011). Organized breeding efforts have used a common-garden approach with white mold screening  
49 nurseries in dry bean production areas across the United States with additional sites in Australia, France,  
50 and Mexico (Steadman et al., 2003, 2004, 2005, 2006; Otto-Hanson & Steadman, 2007, 2008; McCoy &  
51 Steadman, 2009). These white mold screening nurseries use no chemical or cultural treatments against *S.*  
52 *sclerotiorum* and employ standardized protocols for screening new cultivars for resistance to white mold  
53 (Steadman et al., 2003; Otto-Hanson et al., 2011). These protocols included three established cultivars  
54 used for comparison in the trials: Beryl (great northern bean, susceptible), Bensi (a.k.a. Ex Rico, navy  
55 bean, low susceptibility), and G122 (cranberry bean, partial resistance) (Tu & Beversdorf, 1982; Steadman  
56 et al., 2005; Otto-Hanson et al., 2011). It was previously shown that aggressiveness (the severity of disease  
57 symptoms on the host) is significantly different across white mold screening nursery sites in separate  
58 geographic regions (Otto-Hanson et al., 2011). The genetic structure and mode of reproduction in these  
59 populations, however, is currently unknown.

60 Understanding genetic relationships and reproduction behavior of *S. sclerotiorum* populations is  
61 beneficial for breeders seeking to develop new resistant cultivars for worldwide deployment (Milgroom,  
62 1996; McDonald & Linde, 2002). In particular, genetically diverse populations with high rates of sexual  
63 reproduction are more likely to overcome host resistance. Most populations of *S. sclerotiorum* are  
64 predominantly clonal with low genetic diversity and have a large degree of population fragmentation  
65 (Kohli et al., 1995; Cubeta et al., 1997; Kohli & Kohn, 1998; Carbone & Kohn, 2001; Ekins et al., 2011;  
66 Attanayake et al., 2012). Some studies, however have found populations that show signatures of sexual  
67 reproduction (Atallah et al., 2004; Sexton & Howlett, 2004; Attanayake et al., 2013; Aldrich-Wolfe et al.,  
68 2015).

69 Nearly all population genetic studies of *S. sclerotiorum* employ a macroscopic assay to determine  
70 mycelial compatibility, the ability for fungal hyphae from different colonies to appear to grow together  
71 without forming a barrier of dead cells between them (known as a barrage line, Fig. S1B) (Leslie, 1993;  
72 Sirjusingh & Kohn, 2001). Mycelial compatibility has been used as a proxy for vegetative compatibility,  
73 a fungal trait controlled by several independent genes controlling the ability for two hyphae to fuse and  
74 grow as a single unit (Fig. S1A) (Leslie, 1993; Schafer & Kohn, 2006). Because of the genetic connection  
75 to vegetative compatibility, two isolates that are mycelially compatible were considered clones (Leslie,  
76 1993); but correlation with genetic markers, such as microsatellites, have shown mixed results (Ford et al.,  
77 1995; Micali & Smith, 2003; Jo et al., 2008; Attanayake et al., 2012; Papaioannou & Typas, 2014; Lehner  
78 et al., 2017). Thus, the relationship between mycelial compatibility groups and clonal genotypes remains  
79 unclear.

80 In this study, we analyze and characterize the genetic and phenotypic diversity of 366 *S. sclerotiorum*  
81 isolates collected between 2003 and 2012 from dry bean cultivars among different geographic locations  
82 in the Australia, France, Mexico, and the United States. We wanted to know if the *S. sclerotiorum*  
83 populations from white mold screening nurseries were representative of the fields within the same region.  
84 As these nurseries were not treated with any chemical or cultural control of white mold, we hypothesized  
85 that these nurseries would represent the natural population of *S. sclerotiorum*. Furthermore, we wanted  
86 to investigate the potential effect of cultivar on genetic diversity of the pathogen by assessing three dry  
87 bean cultivars with different levels of resistance, Beryl (great northern bean, susceptible), Bensi (navy  
88 bean, low susceptibility), and G122 (cranberry bean, partial resistance) (Otto-Hanson et al., 2011). We  
89 additionally wanted to determine categorical or phenotypic variables that best predicted genetic structure  
90 and if there was correlation between multilocus haplotype and mycelial compatibility group. Knowing  
91 what variables predict genetic structure can help direct breeding efforts. By investigating these aims, we  
92 will effectively describe the population structure of *S. sclerotiorum* in the USA and make available our  
93 database of isolates for use in future dry bean breeding efforts.

## 94 MATERIALS AND METHODS

### 95 Isolate collection

96 Several (156) of the isolates used for this study were collected as reported in previous studies using  
97 the same methods (Otto-Hanson et al., 2011). Broadly, isolates were collected from two sources: white  
98 mold screening nurseries (wmn) or producer fields. White mold screening nurseries were 5m x 10m in

size and maintained without application of fungicides to observe natural incidence of white mold. The early nursery plots were incorporated with a basal dressing of N:P:K = 1:3:2 and side dressing of 0:3:2 during the growing season (Steadman et al., 2003).

Sampling was carried out by collecting sclerotia from diseased tissue in zig-zag transects across field plots. Because sampling depended on disease incidence, the number of samples isolated varied from year to year. Although the nursery locations were the same over sampling years, sampling plots within a location varied for sampling years.

Sclerotia of *S. sclerotiorum* were collected over several years from grower fields and/or wmn in 11 states of the Australia, France, Mexico, and the United States (Table S1). After collection, sclerotia were stored in Petri plates lined with filter paper, then stored at 20 °F or -4 °C. Sclerotia were surface-sterilized with 50% Clorox bleach (at least 6% NaOCl, The Clorox Company, Oakland, CA) solution for 3 min, and double rinsed with ddH<sub>2</sub>O for 3 min. The sterilized sclerotia were then placed on water agar plates (16g of Bacto agar per liter of ddH<sub>2</sub>O, BD Diagnostic Systems, Sparks, MD), with four to five sclerotia of each isolate separated on each plate and stored on the counter top at room temperature for 5 to 6 days. An 8-mm plug from a 5- or 6-day-old culture was transferred from the advancing margin of the mycelia onto a plate of Difco potato dextrose agar (PDA at 39 g/liter of ddH<sub>2</sub>O) (Otto-Hanson et al., 2011). In combination with the 156 isolates described previously, we collected 210 isolates for a total of 366 isolates (Otto-Hanson et al., 2011).

### **Mycelial Compatibility**

MCG was determined as described previously through co-culturing pairs of 2-day-old isolates 2.5 cm apart on Diana Sermons (DS) Medium (Fig. S1) (Cubeta et al., 2001). Incompatibility of different MCGs resulted in formation of a barrage line accompanied by formation of sclerotia on either side of the barrage line, indicating the limits of the isolates' growth (Kohn et al., 1990; Leslie, 1993; Otto-Hanson et al., 2011). Isolates were compared in a pairwise manner for each site and then representatives among sites were compared to determine mycelial compatibility groups by scoring compatible and incompatible interactions (Otto-Hanson et al., 2011). No MCGs were compatible with any other MCG.

### **Aggressiveness**

Aggressiveness of each isolate was assessed using a straw test as described in Otto-Hanson et al. (2011) that rated necrotic lesion size (Petzoldt & Dickson, 1996; Teran et al., 2006). Briefly, the straw test uses 21-day-old G122 plants as the host in a greenhouse setting. Clear drinking straws cut to 2.5 cm and sealed were used to place two mycelial plugs of inoculum on the host plant after removing plant growth beyond 2.5 cm above the fourth node. Measurements of the necrotic lesion were taken 8 days later using the Modified Petzoldt and Dickson scale of 1–9, where 1 is no disease and 9 is plant death (Petzoldt & Dickson, 1996; Teran et al., 2006).

### **Microsatellite genotyping**

Prior to DNA extraction, isolates were grown on PDA and plugs were subsequently transferred to Potato Dextrose Broth (PDB) where they were grown until there was significant mycelial growth, but before the mycelial mat became solidified (4–5 days). Each mycelial mat was collected in a filtered Büchner funnel, agar plugs removed, lyophilized and pulverized manually in Whirl-pak® HDPE sampling bags (Sigma-Aldrich, St. Louis, MO). Lyophilized mycelia was 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 and evaporation, suspending the DNA in 200μl TE (Sambrook et al., 1989). Suspended DNA was stored at 4 °C until genotyping.

We genotyped each *S. sclerotiorum* isolate using 16 microsatellite primer pairs developed previously (Sirjusingh & Kohn, 2001). PCR was carried out as described previously, using primers labeled with FAM fluorophore. Resulting amplicons were first resolved in a 1.5% agarose gel stained with ethidium bromide to ensure product was within the expected size range prior to capillary electrophoresis. Capillary electrophoresis (fragment analysis) of amplicons, with size standard GeneScan™ 500 LIZ®, was performed using an ABI 3730 genetic analyzer (Life Technologies Corporation, Carlsbad, CA) at the Michigan State University Genomic Sequencing Center (East Lansing, MI). Alleles were scored using PeakScanner version 1.0 (Life Technologies Corporation, Carlsbad, CA) and recorded manually in a spreadsheet.

151 **Data processing and Analysis**

152 All data processing and analyses were performed in a Rocker “verse” project container running R ver-  
153 sion 3.4.2 (Boettiger & Eddelbuettel, 2017; R Core Team, 2017) and are openly available and reproducible  
154 at <https://github.com/everhartlab/sclerotinia-366/>. Of the 16 microsatellite loci  
155 genotyped, five included compound repeats, which made it challenging to accurately/confidently bin alle-  
156 les into fragment sizes expected for each locus based on the described repeat motif. Loci with compound  
157 repeats were removed for the reported statistics. To ensure the integrity of the results we additionally  
158 processed these loci and included them in concurrent analyses. We assessed the power of our 11 markers  
159 by generating a genotype accumulation curve in the R package *poppr* version 2.5.0, looking for evidence  
160 of saturation, which would indicate that loci were sufficiently sampled to adequately represent the full set  
161 of haplotypes (Arnaud-Hanod et al., 2007; Kamvar et al., 2015). To avoid including isolates potentially  
162 collected from the same plant (which increases the probability of collecting sclerotia from the same point  
163 of infection more than once), data were clone-corrected on a hierarchy of Region/Source/Host/Year—  
164 meaning that duplicated genotypes were reduced to a single observation when they were collected in the  
165 same year from the same host cultivar located in the same source field (wmn or producer)—for subsequent  
166 analysis. We assessed haplotype diversity by calculating Stoddart and Taylor’s index ( $G$ ) (Stoddart &  
167 Taylor, 1988), Shannon’s index ( $H$ ) (Shannon, 1948), Simpson’s index ( $\lambda$ ) (Simpson, 1949), evenness  
168 ( $E_5$ ), and the expected number of multilocus haplotypes ( $eMLH$ ) (Hurlbert, 1971; Heck et al., 1975;  
169 Pielou, 1975; Grünwald et al., 2003). To assess the potential for random mating, we tested for linkage  
170 disequilibrium with the index of association,  $I_A$  and its standardized version,  $\bar{r}_d$  using 999 permutations  
171 (Brown et al., 1980; Smith et al., 1993; Agapow & Burt, 2001). Both haplotype diversity and linkage  
172 disequilibrium were calculated in *poppr* (Kamvar et al., 2014).

173 **Assessing Importance of Variables**

174 **Distance-based Redundancy Analysis**

175 A distance-based redundancy analysis (dbRDA) (Legendre & Anderson, 1999) was performed with  
176 the function *capscale()* in the *vegan* package version 2.4.4 (Oksanen et al., 2017). This method  
177 uses constrained ordinations on a distance matrix representing the response variable to delineate relative  
178 contribution of any number of independent explanatory variables. We used this method to delineate  
179 the phenotypic (Aggressiveness, Mycelial Compatibility Group (MCG)), geographic (Region, Host,  
180 Location), and temporal (Year) components in predicting genetic composition of the populations. The  
181 distance matrix we used as the response variable was generated using Bruvo’s genetic distance from  
182 clone-corrected data (procedure described above) as implemented in *poppr*, which employed a stepwise  
183 mutation model for microsatellite data (Bruvo et al., 2004; Kamvar et al., 2014). Because aggressiveness  
184 measures differed between isolates that were reduced to a single observation during clone-correction,  
185 aggressiveness was first averaged across clone-corrected isolates. To identify explanatory variable(s)  
186 correlated with genetic variation, a forward-backward selection process was applied with the *vegan*  
187 function *ordistep()*. An analysis of variance (ANOVA) was then performed to test for significance of  
188 the reduced model and marginal effects using 999 permutations. The *varpart()* function of *vegan* was  
189 used to determine variation partitioning of explanatory variables.

190 **Aggressiveness Assessment**

191 We used ANOVA to assess if aggressiveness (determined via straw test on a scale of 1–9 as described  
192 above) was significantly different with respect to Region, MCG, or multilocus haplotype (MLH). To  
193 minimize complications due to small sample sizes, we chose the top 10 MCGs, representing 56.5% of the  
194 isolates collected, the 10 most abundant MLHs representing 26.7% of the isolates, and populations with  
195 more than five isolates. If ANOVA results were significantly different at  $\alpha = 0.05$ , pairwise differences  
196 were assessed using Tukey’s HSD test ( $\alpha = 0.05$ ) using the *HSD.test()* function in the package  
197 *agricolae* version 1.2.8 (Mendiburu & Simon, 2015).

198 **Correlating Multilocus Haplotypes with Mycelial Compatibility Groups**

199 We wanted to assess if there was correlation between MLHs and MCGs. This was performed using a  
200 network-based approach where both MLHs and MCGs were considered nodes and the number of isolates  
201 in which they were found together was the strength of the connection between an MLH and and MCG  
202 node. The network-based approach allowed us to assess the associations between MLHs and MCGs.  
203 To construct the network, a contingency table was created with MLHs and MCGs and converted to a

204 directed and weighted edgelist where each edge represented a connection from an MCG to an MLH,  
205 weighted by the number of samples shared in the connection. This was then converted to a bipartite  
206 graph where top nodes represented MLHs and bottom nodes represented MCGs. To identify clusters of  
207 MLHs and MCGs within the network, we used the cluster walktrap community detection algorithm as  
208 implemented in the `cluster_walktrap()` function in *igraph* version 1.1.2 (Csardi & Nepusz, 2006;  
209 Pons & Latapy, 2006). This algorithm attempts to define clusters of nodes by starting at a random nodes  
210 and performing short, random “walks” along the edges between nodes, assuming that these walks would  
211 stay within clusters. For this analysis, we set the number of steps within a walk to four and allowed the  
212 algorithm to use the edge weights in determining the path. All of the resulting communities that had fewer  
213 than 10 members were then consolidated into one. Community definitions were used to assess the average  
214 genetic distance (as defined by Bruvo’s distance) within members of the community (Bruvo et al., 2004).

## 215 **Genetic Diversity**

### 216 **Population Differentiation**

217 We used analysis of molecular variance (AMOVA) with Bruvo’s genetic distance in *poppr* to test for  
218 differentiation between populations in wmn and producer fields from the same region and collected in  
219 the same year (Excoffier et al., 1992; Bruvo et al., 2004; Kamvar et al., 2014). To identify Regions with  
220 greater differentiation, we used discriminant analysis of principal components (DAPC) as implemented in  
221 *adegenet* version 2.1.0, assessing the per-sample posterior group assignment probability (Jombart, 2008).  
222 This method decomposes the genetic data into principal components, and then uses a subset of these as  
223 the inputs for discriminant analysis, which attempts to minimize within-group variation and maximize  
224 among-group variation (Jombart et al., 2010). To avoid over-fitting data, the optimal number of principal  
225 components was selected by using the *adegenet* function `xvalDapc()`. This function implements a  
226 cross-validation procedure to iterate over an increasing number of principal components on a subset  
227 (90%) of the data, trying to find the minimum number of principal components that maximizes the rate of  
228 successful group reassignment. To assess if cultivar had an influence on genetic diversity between wmn,  
229 we first subset the clone-corrected data to contain only samples from wmn and from the cultivars Beryl,  
230 Bansi, and G122 and tested differentiation using AMOVA and DAPC as described above. We additionally  
231 assessed population stability over time by calculating DAPC over the combined groups of Region and  
232 Year as described above.

### 233 **Analysis of Shared Multilocus Haplotypes**

234 We wanted to evaluate patterns of connectivity between shared multilocus haplotypes across geo-  
235 graphic regions. We first tabulated the multilocus haplotypes shared between at least two populations  
236 (defined as states or countries) with the *poppr* function `mlg.crosspop()` (Kamvar et al., 2014). From  
237 these data, we constructed a graph with populations as nodes and shared haplotypes as edges (connections)  
238 between nodes using the R packages *igraph* (Csardi & Nepusz, 2006), *dplyr* version 0.7.4 (Wickham et  
239 al., 2017), and *purrr* version 0.2.4 (Henry & Wickham, 2017). Each node was weighted by the fraction of  
240 shared MLHs. Each edge represented a single MLH, but because a single MLH could be present in more  
241 than one population, that MLH would have a number of edges equivalent to the total number of possible  
242 connections, calculated as  $(n*(n-1))/2$  edges where n represents the number of populations crossed. Edges  
243 were weighted by  $1 - P_{sex}$ , where  $P_{sex}$  is the probability of encountering the same haplotype via two  
244 independent meiotic events (Parks & Werth, 1993; Arnaud-Hanod et al., 2007). This weighting scheme  
245 would thus strengthen the connection of edges that represented genotypes with a low probability of being  
246 produced via sexual reproduction. We then identified communities (among the Regions) in the graph  
247 using the `cluster_optimal()` function from *igraph* (Csardi & Nepusz, 2006). The graph was plotted  
248 using the R packages *ggplot2* version 2.2.1 (Wickham, 2009) and *ggraph* 1.0.0 (Pedersen, 2017). To  
249 ensure that we captured the same community signal, we additionally performed this analysis including  
250 the five polymorphic markers described above.

## 251 **RESULTS**

252 A total of 366 isolates were collected from 2003 to 2012 (except 2006 and 2011) from diseased dry  
253 bean plants in 11 states in the United States as well as Australia, France, and Mexico (Table S1). With  
254 the 11 loci used in the analyses (Table 1), we observed a total of 165 MLHs (215 with 16 loci). These  
255 11 loci are found on 7 chromosomes in the *S. sclerotiorum* genome with a minimum distance of 55Kbp

256 between two loci on the same chromosome. Over 50% of the isolates came from four states, MI (62), ND  
 257 (60), WA (59), NE (47). Four regions had fewer than 10 isolates, Australia (6), WI (2), NY (1), ID (1).  
 258 We observed 87 MCGs, the most abundant of which ('MCG 5') was represented by 73 isolates over 37  
 259 MLHs (Fig. 1A,C).

260 The number of observed alleles per locus ranged from two to 10 with an average of 6.27 (Table  
 261 1). Locus 20-3, which contained only 2 alleles, showed low values of both  $h$  (0.0533) and evenness  
 262 (0.42), indicating that there was one dominant allele present. Analysis of the haplotype accumulation  
 263 curve showed no clear plateau for 11 or 16 loci (See section on 'Loading Data and Setting Strata' in the  
 264 MLG-distribution.md<sup>1</sup> file in the supplemental files (Kamvar et al., 2017)), indicating that we would  
 265 likely obtain more multilocus haplotypes if we were to genotype more loci.

266 After clone-correction on the hierarchy of Region/Source/Host/Year, a total of 48 isolates were  
 267 removed from the data set, resulting in 318 isolates representing 165 MLHs that were used in subsequent  
 268 analyses (Table 2). The results showed that, in terms of genotypic diversity ( $H$ ,  $G$ , and  $\lambda$ ), WA was the  
 269 most diverse population with both  $G$  (54.3) and  $e^H$  (55.3) being close to the observed number of MLGs  
 270 (56). This indicated that there are few duplicated genotypes in WA (Table 2). A more useful metric to  
 271 compare populations, however, is  $E_5$ , which scales from 0 to 1, where 1 indicates all unique genotypes  
 272 (Grünwald et al., 2003). Evaluating by  $E_5$  shows that both MI and NE exhibit lower than average values,  
 273 indicating that there are over-represented genotypes in the popualtions (table 2). When we look at Mexico,  
 274 we observed that it had relatively high values of  $E_5$  and genotypic diversity, but low richness, as measured  
 275 by  $eMLG$ . Moreover, Mexico had the lowest value for  $h$ , which is a measure of allelic diversity. Nearly all  
 276 populations showed evidence for linkage (Table 2), which serves as evidence for clonal reproduction or  
 277 other forms of non-random mating. The only exceptions are CA ( $P = 0.043$ ) and Australia ( $P = 0.052$ ).  
 278 Both of these populations showed only moderate significance with  $\bar{r}_d$  values of 0.03 and 0.12, respectively.

**Table 1.** Allelic diversity on full data set at loci used in this study.  $h$  = Nei's Gene Diversity (Nei, 1978). Average  $h$  = 0.583, average Evenness = 0.693, average no. alleles = 6.27

Locus	Range	Repeat Motif	No. alleles	$h$	Evenness
5-2	318–324	(GT)	4	0.45	0.62
6-2	483–495	(TTTTTC)(TTTTTG)(TTTTTC)	3	0.64	0.95
7-2	158–174	(GA)	7	0.73	0.76
8-3	244–270	(CA)	7	0.74	0.79
9-2	360–382	(CA)(CT)	9	0.35	0.41
12-2	214–222	(CA)	5	0.58	0.78
17-3	342–363	(TTA)	7	0.55	0.53
20-3	280–282	(GT)GG(GT)	2	0.05	0.42
55-4	153–216	(TACA)	10	0.72	0.66
110-4	370–386	(TATG)	5	0.76	0.91
114-4	339–416	(TAGA)	10	0.83	0.80

<sup>1</sup>Direct link: <https://github.com/everhartlab/sclerotinia-366/blob/master/results/MLG-distribution.md#loading-data-and-setting-strata>

**Table 2.** Genotypic diversity and Linkage Disequilibrium summary for geographic populations arranged by abundance after clone-correction by a hierarchy of Region/Source/Host/Year. Pop = Population, N = number of individuals (number of MLH in parentheses), eMLH = expected number of MLHs based on rarefaction at 10 individuals (standard error in parentheses), H = Shannon-Weiner Index, G = Stoddardt and Taylor's Index,  $\lambda$  = Simpson's Index,  $h$  = Nei's 1978 gene diversity,  $E_5$  = Evenness,  $\bar{r}_d$  = standardized index of association. An asterix indicates a significant value of  $\bar{r}_d$  after 999 permutations,  $P \leq 0.001$ .

Pop	N	eMLH	H	G	$\lambda$	$E_5$	$h$	$\bar{r}_d$
WA	58 (56)	9.95 (0.23)	4.0	54.3	0.98	0.98	0.60	0.07*
MI	58 (43)	9.3 (0.79)	3.6	29.0	0.97	0.78	0.54	0.14*
ND	41 (35)	9.44 (0.73)	3.5	25.9	0.96	0.82	0.54	0.1*
NE	37 (28)	8.93 (0.94)	3.2	17.8	0.94	0.75	0.55	0.25*
CO	34 (28)	9.46 (0.67)	3.3	24.1	0.96	0.92	0.56	0.27*
France	21 (14)	8.5 (0.85)	2.6	12.6	0.92	0.95	0.48	0.11*
CA	18 (15)	9.12 (0.72)	2.7	13.5	0.93	0.94	0.51	0.03
OR	17 (13)	8.52 (0.85)	2.5	10.7	0.91	0.89	0.47	0.1*
Mexico	15 (9)	7.1 (0.85)	2.1	7.3	0.86	0.89	0.28	0.37*
MN	9 (7)	7 (0)	1.9	6.2	0.84	0.93	0.47	0.19*
Australia	6 (6)	6 (0)	1.8	6.0	0.83	1.00	0.48	0.12
WI	2 (2)	2 (0)	0.7	2.0	0.50	1.00	0.27	-
NY	1 (1)	1 (0)	0.0	1.0	0.00	NaN	NaN	-
ID	1 (1)	1 (0)	0.0	1.0	0.00	NaN	NaN	-

## 279 Variable Assessment

### 280 Variable Contributions

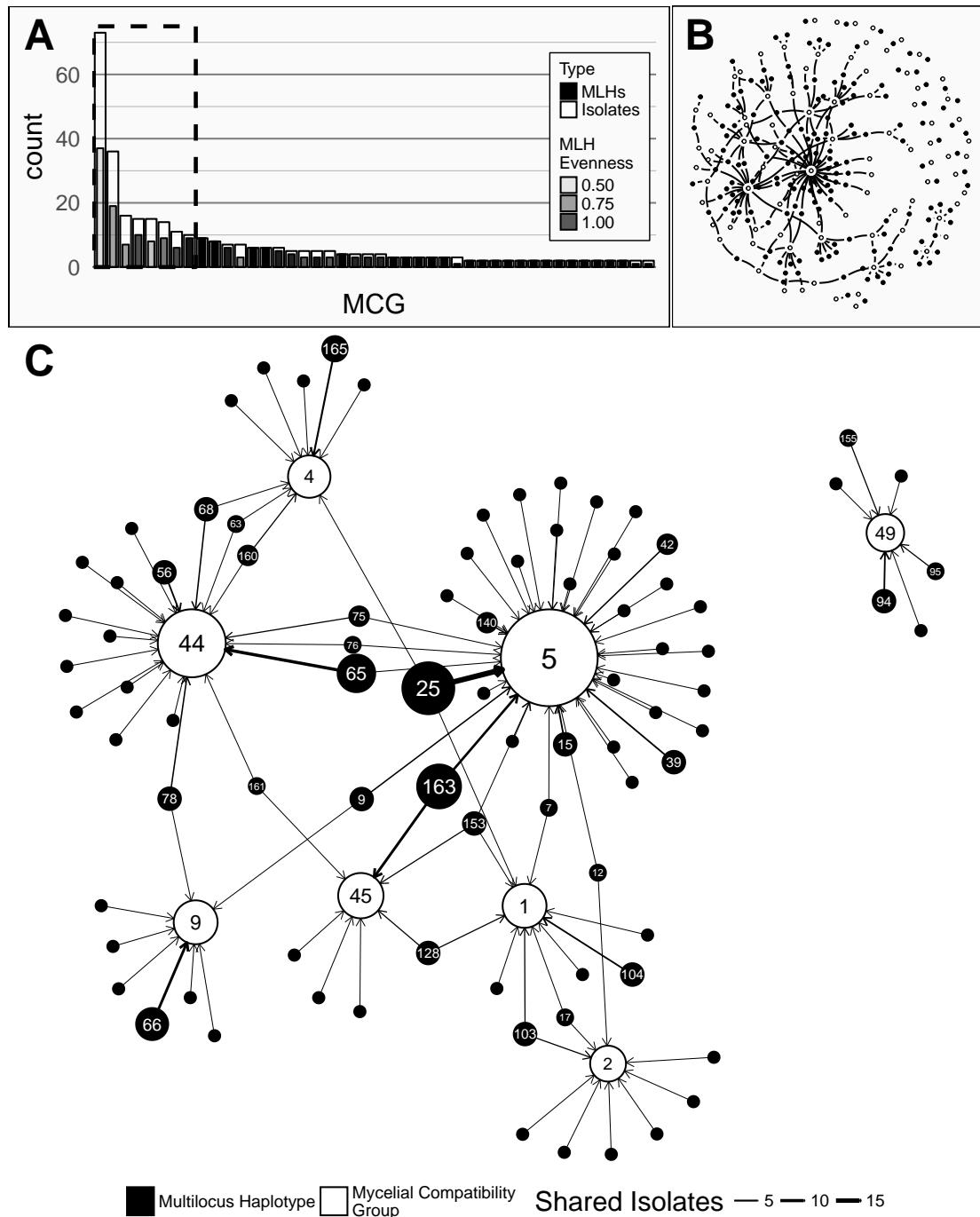
281 The forward-backward selection process of the dbRDA models on clone-corrected data revealed Year,  
 282 Region, Host, and MCG to be the optimal variables for the reduced model, accounting for 45% of the  
 283 total variation. ANOVA showed that the reduced model was significant with an adjusted  $R^2$  of 0.0675  
 284 ( $P = 0.001$ ). Assessment of the marginal effects showed that all variables significantly explained genetic  
 285 variation ( $P \leq 0.007$ ). We found that there was multicollinearity when MCG was combined with any  
 286 other variable, so repeated the analysis, dropping MCG from the list of potential predictors. From these  
 287 results, Year, Region, Host, and Aggressiveness were found to be optimal, accounting for 17.6% of the  
 288 total variation. ANOVA revealed significant effects with an adjusted  $R^2$  of 0.0325 ( $P = 0.001$ ). While  
 289 the marginal effect assessment revealed that Year, Region, and Host significantly explained variation  
 290 at  $P = 0.001$ , and Aggressiveness significantly explained variation at  $P = 0.039$ . Much of the variation  
 291 appeared to be driven by isolates from Mexico and 2005 (Fig. 2). Variance partitioning of the independent  
 292 variables without MCG indicated aggressiveness to be the least influential factor with 0.1% contributing  
 293 to explaining the variation of molecular data, whereas the combination of variables accounted for 3.3%.

### 294 Aggressiveness

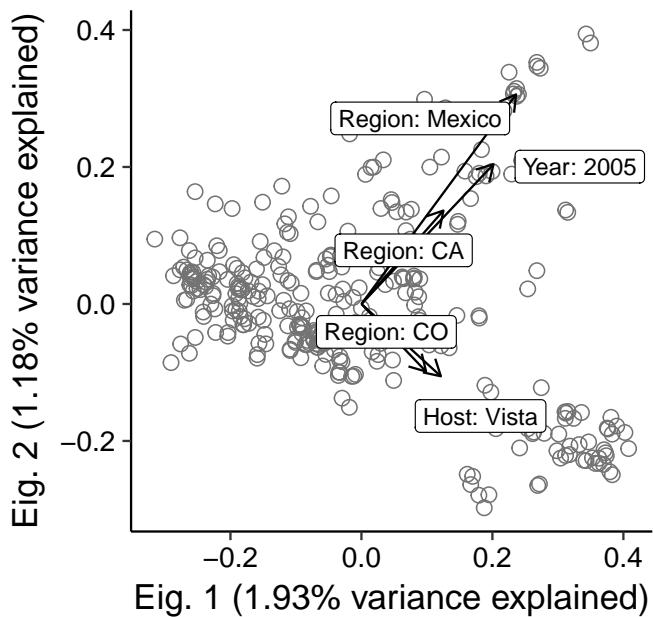
295 Aggressiveness of the isolates ranged from 1.4 to 7.9 with a mean of 5.02 and median of 4.85. The  
 296 group mean averages were 4.88, 5.13, and 5.19 for Region, MCG, and MLH, respectively. A strip plot  
 297 showing the distribution of severity across these three variables simultaneously can be seen in Fig. S2.  
 298 Our assessment of aggressiveness in association with Region showed a significant effect ( $P < 1.00e^{-4}$ ),  
 299 with means that ranged from 5.8 (MN) to 4.0 (CA) (Fig. 3, Table S2). MCGs also showed a significant  
 300 effect ( $P < 0.001$ ), with means that ranged from 6.0 ('MCG 44') to 4.6 ('MCG 49'; Table S3). We  
 301 additionally found a significant effect for MLHs ( $P < 0.001$ ), with means that ranged from 6.0 ('MLH  
 302 78') to 4.3 ('MLH 140') (Table S4).

### 303 Correlation of Multilocus Haplotypes and Mycelial Compatibility Groups

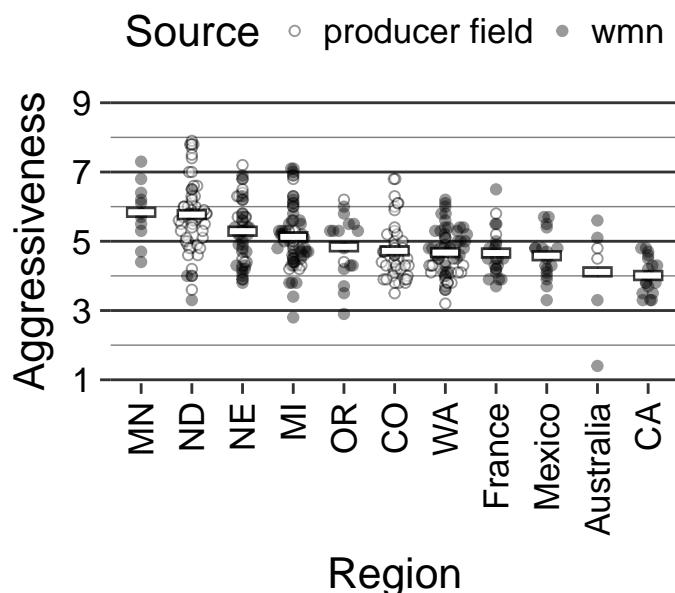
304 In our analysis, we found 165 MLHs with 70 singletons and 87 MCGs with 43 singletons (Fig. 1A,B)  
 305 where the eight most abundant MCGs represented  $> 51\%$  of the data over 11 Regions, and all years  
 306 except for 2012. Our network-based approach to correlating MLHs with MCGs revealed a large and  
 307 complex network (Fig. 1, Table 3). Community analysis showed 51 communities, 15 of which consisted



**Figure 1.** Associations between Mycelial Compatibility Groups and Multilocus Haplotypes. **A)** Barplot of Mycelial Compatibility Group (MCG) abundance in descending order. Singletons (46) were truncated, leaving 41 MCGs. White bars represent sample counts and grey bars represent counts of unique multilocus haplotypes (MLH). The transparency of the bars represent the evenness of the distribution of the MLHs within a given MCG. A dashed box surrounds the eight most common MCGs representing > 51% of the data. **B)** Full graph-representation of the relationship between MCGs (open circles) and MLHs (filled circles). Details in Fig. S3. **C)** A subset of **B** representing the 8 most common MCGs and their associated MLHs (dashed box in **A**). Filled nodes (circles) represent MLHs and open nodes represent MCGs. Node area scaled to the number of samples represented (range: 1–73). Numbers inside nodes are the MLH/MCG label (if n > 1). Edges (arrows) point from MLH to MCG where the weight (thickness) of the edge represents the number of shared isolates (range: 1–19). Edges extending from MLHs displayed to other MCGs are not shown.



**Figure 2.** Biplot showing five most influential explanatory variables (arrows) overlaid on the first two eigenvectors of distance based redundancy analysis of *Sclerotinia sclerotiorum* isolates. The length of the arrows are directly proportional to the strength of the correlation between explanatory and molecular variables. Open circles represent the 318 clone-corrected haplotypes in ordination space.



**Figure 3.** Strip plot of aggressiveness by population arranged in descending order of mean aggressiveness for all populations with  $N > 5$ . White bars represent mean value. Circles represent individual isolates where filled circles are isolates from white mold screening nurseries (wmn) and open circles are isolates from producer fields.

of a single MLH unconnected with any other community indicating that just 9.09% of the 165 MLHs are unable to cross with any other MLH in this data set (Fig. S3). The three communities with the most members contained eight of the 10 most abundant MCGs. Comparing these communities with Bruvo's genetic distance showed an average distance of 0.451 among communities and an average distance of 0.437 within communities, which were not significantly different. When we assessed the number of times two different MLHs that are in the same MCG, considering these as potential heterothallic pairings that could result in sexual recombination, we found an average of 14.3 potential heterothallic pairings per MLH. Representing just four isolates, 'MLH 75' had 57 neighbors that shared the same mycelial compatibility group (Fig. 1, S3). Overall, there was no clear pattern to the association between MLH and MCGs.

**Table 3.** The five most abundant Multilocus Haplotypes (MLH) with the probability of second encounter ( $P_{sex}$ ), Mycelial Compatibility Groups (MCG), and Regions with sample sizes in parentheses.

MLH	$P_{sex}$	MCG	Region
25	0.016824	5	ND (15), CO (2), MI (2)
		13	ND (3)
		60	ND (2), WA (1)
		1	NE (1)
		4	MI (1)
163	0.049932	45	CO (5), ND (2), NE (1)
		5	MI (7)
65	0.000071	44	NE (10)
		5	MI (1)
140	0.000155	8	CO (5)
		5	MI (3)
		20	MI (2)
66	0.000016	9	NE (4), CO (2), MI (2)

### Structure of Shared Multilocus Haplotypes

The most abundant MLH was represented by 27 isolates (Table 3) from five Regions (NE, MI, WA, CO, and ND). Within Regions, haplotypes were relatively evenly distributed with moderate to high diversity (Table 2). Of the 165 MLHs, 76 (46%) were found in at least two Regions, except those found in WI (2), ID (1), and Mexico (18) (Fig. 4).

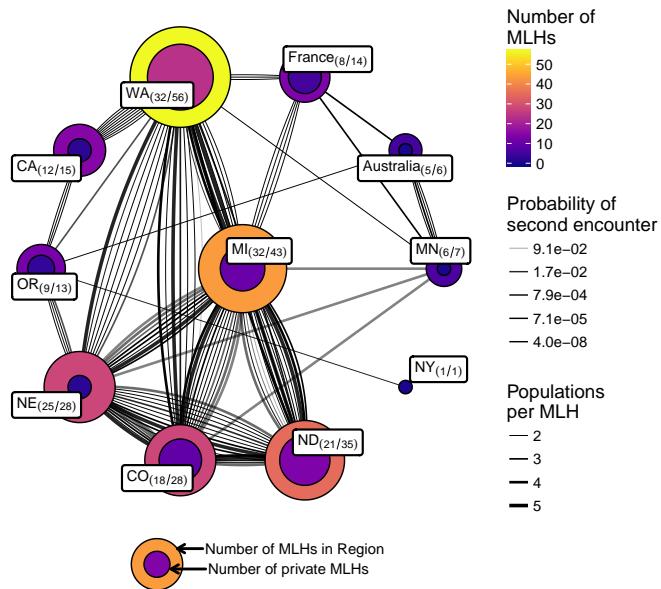
We had performed an analysis on a network where the connections represented shared MLHs across populations, weighted by  $1 - P_{sex}$  (Fig. 4, Table 3). Community analysis of the MLHs shared between populations revealed 4 communities with a modularity of 0.17: A coastal community (CA, OR, WA, and NY), a midwest community (CO, ND, NE, MI), and an international community (Australia, France, MN). Although analysis with 16 loci resulted in the removal of the NY node because it no longer shared a haplotype with OR, the same overall community structure was present with a modularity of 0.2 (Fig. S4). Relative to the US, the international community appears to be driven by MLH 4, which is shared between all three populations and has a  $P_{sex}$  value of  $2.87e^{-5}$ , in contrast to the abundant MLH 25, which has a  $P_{sex}$  value of 0.0168.

### Population Differentiation

#### Analysis of Molecular Variance

The analysis of molecular variance (AMOVA) for clone-corrected samples over the hierarchy of Region, Source, and Year showed significant variation between Regions and Years, but no significant variation between wmn and producer fields (Table 4). In contrast, when we compared the three cultivars, Beryl, Bansi, and G122, we found no significant differentiation (See section on 'Host Differentiation' in the wmn-differentiation.md<sup>2</sup> file in the supplemental files (Kamvar et al., 2017)).

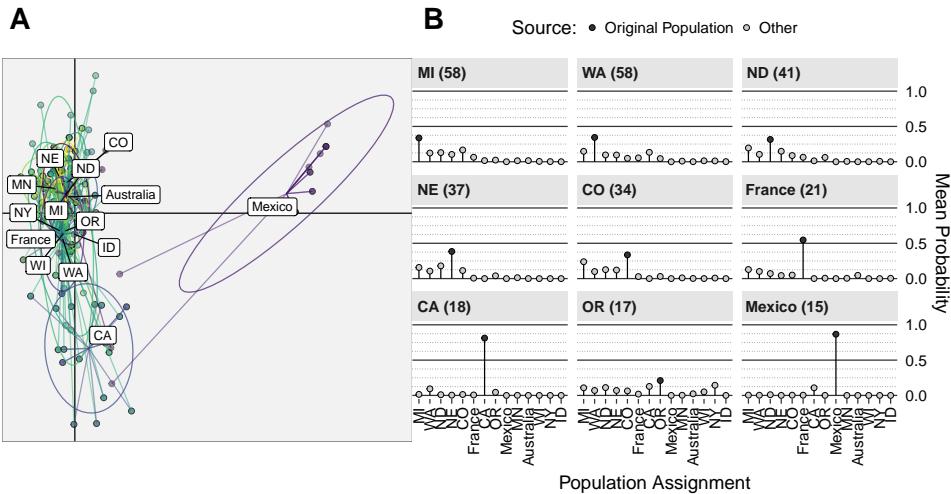
<sup>2</sup>Direct link: <https://github.com/everhartlab/sclerotinia-366/blob/master/results/wmn-differentiation.md#host-differentiation>



**Figure 4.** Network of populations (nodes/circles) and their shared multilocus haplotypes (MLH) (edges/lines) genotyped over 11 loci. Each node is labeled with **name (number of MLHs shared/number of MLHs total)**. The shade and area of the nodes are proportional to the number of unique MLHs within the node and the inner nodes are proportional to the number of private MLHs to the region (bottom legend). Each edge represents a single MLH where its thickness represents the number of populations that share the MLH and the shade represents the value of  $P_{sex}$ , or the probability of encountering that MLH from two independent meiotic events.

**Table 4.** Comparison of populations in the white mold screening nurseries (wmn) and producer fields using an analysis of molecular variance on Bruvo's genetic distance showing no apparent differentiation between wmn and other sources. The hierarchy was constructed as Source/Region where source is defined as belonging to a wmn or producer field. Bold  $\Phi$  values indicate significant difference ( $P < 0.05$ ). S.S. = Sum of Squares, d.f. = degrees of freedom.

Hierarchy	d.f.	S.S.	% variation	$\Phi$ statistic	P
Between Region	13	10.19	8.45	<b>0.0845</b>	0.031
Between Source within Region	8	2.74	-2.29	-0.0250	0.497
Between Year within Source	22	9.37	16.28	<b>0.173</b>	0.001
Within Year	274	47.30	77.56	<b>0.224</b>	0.001



**Figure 5.** Discriminant Analysis of Principal Components on regions showing that Mexico is differentiated from other populations. **A)** Scatter plot of first two components from DAPC. Points represent observed individuals connected to the population centroids with ellipses representing a 66% confidence interval for a normal distribution. The center of each component is represented as black grid lines. **B)** Mean population assignment probability from the DAPC for all populations with  $N > 10$  (facets). Populations represented along the horizontal axis and probability of assignment on the vertical. Numbers next to source populations indicate population size. All values sum to one.

338      **Discriminant Analysis of Principal Components**

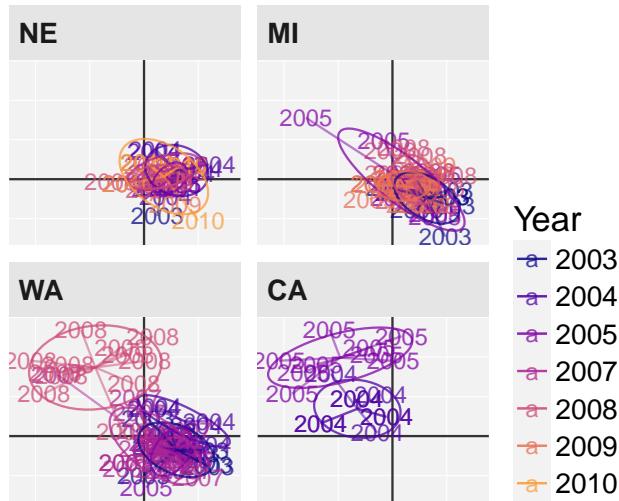
339      Discriminant Analysis of Principal Components (DAPC) was performed by grouping Region with  
 340      the first 21 principal components, representing 88.1% of the total variance. The first discriminant axis  
 341      (representing 63.9% of the discriminatory power) separated the centroid for the Mexico isolates from the  
 342      rest of the data, indicating strong differentiation (Fig. 5b). The second discriminant axis, representing  
 343      10.8% of the discriminatory power, separated the centroid for the CA isolates. The mean population  
 344      assignment probabilities for all populations with  $n > 10$  showed that only isolates from Mexico, CA, and  
 345      France had  $> 50\%$  probabilities of being reassigned to their source populations (Fig. 5a).

346      DAPC grouping by cultivar used the first 20 principal components, representing 89% of the total  
 347      variance. The first two discriminant axes (representing 100% of the discriminatory power) failed to  
 348      separate any of the cultivars where the mean posterior assignment probabilities were 34% (G122), 35.9%  
 349      (Beryl), and 30.1% (Bunsi). DAPC grouping by Region and Year used the first 15 principal components,  
 350      representing 80.3% of the total variance. The North Central USA populations (NE, MI, CO, ND) did not  
 351      appear to have any variation across time in contrast to WA, which showed a shift in population structure  
 352      in the last year of sampling, 2008 (Fig. 6). Further analysis of this population revealed that all 12 isolates  
 353      in WA circa 2008 originated in a wmn; nine haplotypes were shared with CA, and three were shared with  
 354      France (Fig. 4, S4).

355      **DISCUSSION**

356      In this study, we characterized the diversity of *Sclerotinia sclerotiorum* from dry bean fields across  
 357      the United States. Our results suggest that, broadly, populations from white mold screening nurseries  
 358      reflect the populations of the surrounding regions, indicating that resistance screening may be successful  
 359      within regions. We found significant population differentiation by geographic region and year, mainly  
 360      differentiated into three broad North American groups based on shared haplotypes and posterior groupings,  
 361      a Coastal Region, Midwestern Region, and Mexico. To date, with 366 isolates, this is the largest single  
 362      population genetic study of *S. sclerotiorum* assessing population structure within managed and unmanaged  
 363      agricultural environments. These findings indicate that the white mold screening nurseries can be effective  
 364      at screening for potential resistant lines within growing regions.

365      We found that the best predictors of genetic structure are Region and Year, supporting the hypothesis



**Figure 6.** Scatter plot of Discriminant Analysis of Principal Components on Regions and Years showing non-differentiated temporal variation NE and MI and temporal variation in WA and CA. Points (text labels) represent observed individuals connected to the population centroids with ellipses representing a 66% confidence interval for a normal distribution. The center of each component is represented as black grid lines. A more detailed view is shown in Fig. S5.

that *S. sclerotiorum* populations are spatially structured (Carbone & Kohn, 2001). Borrowing a technique often used in the ecological literature, we used dbRDA to elucidate the effect of all variables (MCG, Region, Source, Year, Host, and Aggressiveness) (Legendre & Anderson, 1999). From the initial results, it appeared that the most important factors for predicting genetic structure were MCG, region, and year. When we inspected the biplot of the initial results, we saw that the most important predictors were 'MCG 44', 'MCG 5', and 'MCG 9'. We believe that this was driven by the fact that these particular MCGs have uneven MLH distributions, meaning that they are heavily associated with one particular MLH (Fig. 1). We note these results with caution because of the apparent multicollinearity between MCG and Region, which is a violation of the analysis (Legendre & Anderson, 1999). While the results indicated that Mexico and the year 2005 were the two most important variables, it's worth noting that all Mexico isolates were collected in 2005 (Fig. 2). The results also show that the Vista cultivar explains some of the variance, but this represents six isolates in MI, and thus we cannot draw broad conclusions from this axis. Aggressiveness and source field had little to no effect on prediction of genetic diversity. These results are in agreement with studies that examined differentiation based on Host (Aldrich-Wolfe et al., 2015) and Aggressiveness (Atallah et al., 2004; Attanayake et al., 2012, 2013) reporting little or no correlation of genetic diversity to these variables. This indicates that a) breeders should keep in mind regional differences when assessing resistance and b) it is possible that we have not yet measured biologically relevant variables that can predict genetic differentiation, which could include variables such as soil community composition.

While aggressiveness was not shown to predict genetic structure, it is an important factor in breeding efforts, and we observed significant differences in aggressiveness based on Region (Fig. 3, Table S2). These results show a similar pattern to what was found previously in Otto-Hanson et al. (2011) with the exception of North Dakota, which increased in mean aggressiveness from 5 to 5.77. This increase was due in part to new data from producer field isolates collected after the previous study. These straw tests were performed by a different person for these later isolates, which could suggest a more lenient or strict scoring system. However, when we examined the within-region differences, we found no significant effect by individual. Many of the ND isolates fell within the 6–7 range, which denotes a physical boundary (disease symptoms around the second node) between intermediate and susceptible (Otto-Hanson et al., 2011). Thus, we observed a shift in aggressiveness without a significant shift in genotypic structure, which may indicate that aggressiveness may be controlled by environmental factors as opposed to genetic profile.

397 The primary interest of this study was to assess if isolates sampled from white mold screening  
398 nurseries represent isolates from producer fields within the region (Steadman et al., 2003; Otto-Hanson  
399 et al., 2011). According to our AMOVA results, we have evidence for differentiation at the Region and  
400 Year, but little to no differentiation between wmn isolates and production field isolates (Table 4). This  
401 lack of differentiation, however, may reflect the breeder practice of inoculating screening plots with  
402 sclerotia collected from sources within the region. When we analyze the AMOVA results in light of the  
403 DAPC results (Fig. 5), it becomes clear that the regional patterns of differentiation are largely driven by  
404 isolates from Mexico and CA. Isolates from these Regions had a higher posterior probability ( $> 0.75$ ) of  
405 being reassigned to their own populations than any other (Fig. 5A). All other populations in comparison  
406 (except France) has reassignment probabilities of  $< 0.5$ , which is reflected in the failure of the first two  
407 discriminant functions to separate these populations (Fig. 5B).

408 Despite the evidence that Mexico and CA contributed to much of the population differentiation,  
409 Regions like WA still had a large amount of internal variation. The two distinct clusters for the WA Region  
410 showed that the 2008 population appeared differentiated and, under further investigation, we found that  
411 all the haplotypes from this year were shared between CA and France (Fig. 4, 6, S5). All of the isolates  
412 from WA in 2003–2005, and 2008 came from the same wmn; within the wmn, those in 2003–2005 came  
413 a Northeastern field location cropped with dry bean since 2002, and those in 2008 from a Southeastern  
414 field that was previously cropped with brassica, sundgrass, peas, beans, and potatoes (Miklas, Phil Pers.  
415 comm.). Both of these fields were inoculated with sclerotia in 2002, the Northeastern field with sclerotia  
416 provided by a commercial bean producer and the Southeastern field with sclerotia from peas (although  
417 this was thought to be unsuccessful). Despite this information, it is still unclear what has contributed  
418 to the differentiation of the 2008 population from WA or why it shares haplotypes with CA and France.  
419 When we assessed aggressiveness between the two fields across years with an ANOVA model, we found  
420 that there was a slight effect based on field ( $P = 0.0127$ ). While the evidence may suggest host as being a  
421 factor, previous studies have shown no significant differentiation across host species (Aldrich-Wolfe et  
422 al., 2015). It was of interest to compare our data with that of Aldrich-Wolfe et al. (2015), but we found  
423 that, due to differences in data generation, we were unable to confidently perform a comparison (See  
424 supplemental file compare-aldrich-wolfe.md<sup>3</sup> (Kamvar et al., 2017)).

425 With the exception of the WA Region, populations that were sampled across several years appeared to  
426 be relatively stable over time with overlapping distributions in the DAPC (i.e. NE and MI, Fig. 6). DAPC  
427 is based on the principal components of allele counts (Jombart et al., 2010). Unlike Bruvo's distance,  
428 this does not take into account the magnitude of the difference between alleles, which could inflate the  
429 distance measure in the presence of private alleles (Bruvo et al., 2004). While we found no evidence  
430 of private alleles in the Mexico and CA isolates, we did find that the alleles driving the first axis in Fig.  
431 5A (alleles 174, 256, and 372 in loci 7-2, 8-3, and 9-2, respectively) were overrepresented in Mexico  
432 (where  $>75\%$  of the alleles came from the region). However, all three of these alleles, i) conform to the  
433 expected stepwise mutation model (Bruvo et al., 2004) and ii) are at or near the extremes of the total range  
434 (except for allele 372 at locus 9-2). Moreover, the fact that we find three alleles at three independent loci  
435 segregating for Mexico suggests that the pattern separating these populations from the others was not an  
436 artifact. We believe that the differences in populations observed from Mexico may be due to differences  
437 in climate that allow greater diversification via sexual outcrossing.

438 Many of the isolates in our study were from temperate climates and the only isolates representing a  
439 sub-tropical climate were from Mexico. It has been proposed within the *S. sclerotiorum* literature that  
440 isolates from sub-tropical and tropical climates are differentiated or more variable than populations from  
441 temperate climates (Carbone & Kohn, 2001; Attanayake et al., 2013; Lehner & Mizubuti, 2017). This  
442 has been attributed to the notion that the fungus has the chance to undergo more reproductive cycles in  
443 the warmer climate (Carbone & Kohn, 2001; Attanayake et al., 2013). The strongest evidence to date  
444 supporting this hypothesis is from Attanayake et al. (2013), showing that populations in sub-tropical  
445 regions of China have been found to be more variable, sexually reproducing, and unrelated to populations  
446 in temperate regions of the USA. This result however, may be driven more by geography and agricultural  
447 practice as opposed to climate.

448 The results from our shared haplotype analysis show several populations with at least one haplotype  
449 between them, except for Mexico and two states that had fewer than three samples each (Fig. 4). Our

450 <sup>3</sup>Direct link: <https://github.com/everhartlab/sclerotinia-366/blob/master/results/compare-aldrich-wolfe.md>

450 network-based approach by treating the haplotypes as edges and weighting each edge with the inverse  
451 of  $P_{sex}$  treated the edges as springs connecting the populations with the strength proportional to the  
452 probability of obtaining the same haplotype as a clone. This allowed us to use a graph walking algorithm  
453 to see how close the populations were simply based off of the proportion of clones they shared. The  
454 most abundant haplotype was shared across four populations, but its high value of  $P_{sex}$  meant that it did  
455 not contribute significantly to the overall structure. The graph walking algorithm was able to divide the  
456 network into three groups, but had a modularity of 0.17, which indicates that the groups are only weakly  
457 differentiated.

458 The widespread nature of multilocus haplotypes in both wmn and production fields with relatively  
459 small values of  $P_{sex}$  may indicate the spread of inoculum between regions. While seedborne transmission  
460 is thought to be of insignificant epidemiological importance (Strausbaugh & Forster, 2003), it has since  
461 been shown that *S. sclerotiorum* infections can be transmitted through seed (Botelho et al., 2013). Thus,  
462 we hypothesize that shared haplotypes between populations may arise due to transmission events of seed  
463 or sclerotinia. This could explain the fact that we see shared genotypes with low  $P_{sex}$  values shared between  
464 Australia, France, and the United States. While we speculate that these transmission events are rare due to  
465 the genetic structuring by Region, these results suggest that seedborne infections may indeed reflect a  
466 source of inoculum. This may, in turn increase the risk of introducing new sources of genetic variation  
467 through potential outcrossing events.

468 When we tested for sexual reproduction, we were unable to find evidence for it in any region except  
469 for Australia and CA. While the Australia population had a non-significant value of  $\bar{r}_d$ —which would  
470 suggest that we cannot reject the null hypothesis of random mating—the sample size was insufficient  
471 from which to draw conclusions (Milgroom, 1996; Agapow & Burt, 2001). The low value of  $\bar{r}_d$  in the CA  
472 population may represent sexual reproduction, but we can see in Fig. 6 that there is differentiation by year.  
473 Thus, this could also be an artifact of sampling two different populations, which is known to reduce the  
474 value of  $\bar{r}_d$  (Prugnolle & de Meeûs, 2010).

475 The previous study of the white mold screening nursery populations used MCGs to assess genotypic  
476 diversity (Otto-Hanson et al., 2011). Historically, MCGs have been used as a proxy for clonal lineages,  
477 and thus, of interest in this study was testing the association between multilocus haplotypes (MLHs) and  
478 mycelial compatibility groups (MCGs) (Kohn et al., 1990; Leslie, 1993; Kohn, 1995; Carbone et al., 1999;  
479 Schafer & Kohn, 2006; Otto-Hanson et al., 2011). Our results, however, do not support this assumption.  
480 It can be seen in Fig. 1A that the most abundant MCG contains several MLHs, but the diversity of those  
481 MLHs are low as indicated by the evenness (transparency), which indicates that there is one dominant  
482 MLH ('MLH 25'). What is not shown in Fig. 1A is the MLHs that are shared between MCGs. This is  
483 illustrated in both Table 3 and Fig. 1B,C. It could be argued, however that 'MLH 25', with its high value  
484 of  $P_{sex}$  represents different true MLHs across the five MCGs it occupies, but this does not account for the  
485 overall structure of Fig. S3 where, for example, 'MLH 75' ( $P_{sex} = 1.81e^{-4}$ ) is compatible with 57 other  
486 haplotypes through three MCG when the population structure of *S. sclerotiorum* is known to be clonal.

487 Over the past few years, researchers have noticed inconsistencies among the relationship between  
488 MCGs and MLHs (Carbone et al., 1999; Attanayake et al., 2012; Aldrich-Wolfe et al., 2015; Lehner et al.,  
489 2015). Either several MCGs belong to one MLH, which could be explained by insufficient sampling of  
490 loci; several MLHs belong to one MCG, which could be explained by clonal expansion; or a mixture of  
491 both. Some studies have shown a correlation between MCG and MLH (Carbone et al., 1999; Aldrich-  
492 Wolfe et al., 2015; Lehner et al., 2015), whereas other studies have shown no apparent correlation, even  
493 on small spatial scales (Atallah et al., 2004; Attanayake et al., 2012, 2013).

494 One long-held assumption was that MCGs (as determined via barrage reaction) represent vegetative  
495 compatibility groups (VCGs) (Kohn et al., 1990; Schafer & Kohn, 2006; Lehner et al., 2015), which are  
496 known to have a genetic component (Saupe, 2000; Hall et al., 2010; Strom & Bushley, 2016). While  
497 our protocol for assessing MCGs utilized Diana Sermons Medium (Cubeta et al., 2001) as compared to  
498 Patterson's Medium or Potato Dextrose Agar (Schafer & Kohn, 2006) for the MCG reactions, the patterns  
499 we observe are not dissimilar from what have previously been reported in the literature. It has been  
500 demonstrated in several Ascomycetes—including *Neurospora crassa* (Micali & Smith, 2003), *Sclerotinia*  
501 *homoeocarpa* (Jo et al., 2008), *Verticillium dahliae* (Papaioannou & Typas, 2014), and *S. sclerotiorum* (Ford  
502 et al., 1995)—that barrage reactions are independent from stable anastomosis. Thus, the inconsistencies  
503 in this study and other studies indicate that researchers studying *S. sclerotiorum* should not rely on MCG  
504 data derived from barrage reactions as an indicator for genetic diversity.

505 **Limitations**

506 One of the main limitations of this study is the focus on *P. vulgaris* as a host. It has been shown  
507 that *S. sclerotiorum* in the midwestern United States does not have a particular preference for host  
508 (Aldrich-Wolfe et al., 2015). If the distribution of *S. sclerotiorum* is even across agricultural hosts in the  
509 USA, then our sample may yet be representative of the genetic pool present in other crops and weedy  
510 species. Additionally, while we found no significant association between genotype and aggressiveness, it  
511 is important to note that the straw test is only one measure of aggressiveness. Additional phenotypes for  
512 aggressiveness should be evaluated for future studies.

513 Another limitation was the microsatellite markers used for this particular study (Sirjusingh & Kohn,  
514 2001). The haplotype accumulation curve showed no indication of a plateau, indicating that if we had  
515 sampled more loci, we would have resolved more multilocus haplotypes. While 16 loci showed us  
516 similar results and began to show a plateau for the haplotype accumulation curve, we were unable to  
517 use these results due to our uncertainty in the allele calls for these five extra loci. With the availability  
518 of an optically-mapped genome (Derbyshire et al., 2017), future studies describing the genetic diversity  
519 of *S. sclerotiorum* should employ techniques such as Genotyping-By-Sequencing (Davey et al., 2011),  
520 Sequence Capture (Grover et al., 2012), or Whole Genome Sequencing.

521 **Conclusions**

522 This study represents the largest genetic analysis of *S. sclerotiorum* from the USA to date, giving us a  
523 unique insight to continent-wide population structure and relationships between phenotypic and genotypic  
524 variables. Populations in wmn appear to show no significant differentiation when compared to their  
525 production field counterparts, suggesting that the wmn populations of *S. sclerotiorum* may be considered  
526 representative of the surrounding regions. While we found no direct relationship between haplotype and  
527 severity, it is evident that there is a gradient of severity by region, further supporting the need for screening  
528 in multiple locations. Based on our analysis of the relationships between MCG and MLH, we found  
529 no clear evidence that the two are directly related, suggesting that MCG does not necessarily represent  
530 vegetative compatibility groups and thus should not be used as a proxy for identifying clones.

531 **Data Availability**

532 All scripts, data, and resources used to generate the results presented in this publication (including  
533 Supplementary Information) are fully reproducible and available at The Open Science Framework  
534 <https://osf.io/ejb5y> (Kamvar et al., 2017).

535 **Funding**

536 Funding for this research was provided by the North Central Soybean Research Program (#639K623)  
537 to SEE, a Layman Award (#2446) to SEE, USDA-ARS National Sclerotinia Initiative (#58-5442-2-209)  
538 to JRS/SEE, and start-up funds from the University of Nebraska-Lincoln to SEE. The funders had no role  
539 in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

540 **Acknowledgements**

541 The authors would like to thank Rebecca Higgins for technical support in generating the data for the  
542 MCG assessment, aggressiveness ratings, and genotyping; and for providing valuable insights into the  
543 historical context of the data collection and curation.

544 We would also like to thank Denita Hadziabdic and two other anonymous reviewers for their valuable  
545 comments and insights that improved the quality of the manuscript.

546 **Conflicts of Interest**

547 The authors declare no conflict of interest.

548 **Author Contributions**

549 • **Zhian N. Kamvar** analyzed the data, contributed analysis tools, wrote the paper, prepared figures  
550 and tables, edited and reviewed drafts of the paper.

551 • **Bimal Sajeewa Amaradasa** analyzed the data, contributed analysis tools, wrote drafts of the paper,  
552 edited and reviewed drafts of the paper.

- 553     • **Rachana Jhala** Carried out experiments (MCG assessment, aggressiveness ratings, genotyping),  
554        edited and reviewed drafts of the paper.
- 555     • **Serena McCoy** Carried out experiments (MCG assessment, aggressiveness ratings, genotyping),  
556        edited and reviewed drafts of the paper.
- 557     • **James R. Steadman** Conceived and designed experiments, organized network of white mold  
558        screening nurseries, provided *S. sclerotiorum* isolates, edited and reviewed drafts of the paper.
- 559     • **Sydney E. Everhart** supervised data analysis, analyzed the data, contributed analysis tools, wrote  
560        the paper, edited and reviewed drafts of the paper.

561 **SUPPLEMENTARY INFORMATION**

**Table S1.** Description of *Sclerotinia sclerotiorum* isolates used in this study. N = Number of Isolates. Key abbreviations: wmn = white mold screening nursery, producer = producer field, unk = unknown cultivar.

Country	State	Field Code	Year	Host	N
USA	CA	wmn	2004, 2005	Beryl, Bansi, G122	18
USA	CO	producer	2007, 2010	Pinto, Yellow	41
		wmn	2003	GH	1
USA	ID	producer	2003	GH	1
USA	MI	wmn	2003, 2004, 2005, 2008, 2009	11A, 37, 38, B07104, Beryl, Bansi, Cornell, G122, Orion, PO7863, WM31	43
		producer	2003, 2008, 2009	BL, Black, Fuji, GH, Merlot, SR06233, unk, Vista, Zorro	19
USA	MN	wmn	2003, 2004	Beryl, Bansi, G122	11
USA	ND	producer	2007, 2010	unk	53
		wmn	2005	Beryl, Bansi, G122	7
USA	NE	wmn	2004, 2005, 2008, 2010	Beryl, Bansi, G122, PO7683, unk	27
		producer	2003, 2007, 2009, 2010	Beryl, Emerson, GH, Orion, Pinto, Weihing	20
USA	NY	producer	2003	GH	1
USA	OR	wmn	2003, 2004	Beryl, Bansi, G122	15
		producer	2003	G122, GH	2
USA	WA	wmn	2003, 2004, 2005, 2008	11A, 37, 38, Beryl, Bansi, Cornell, G122, Orion, PO7 104, PO7863, WM31	36
		producer	2003, 2007	GH, Merlot, Pinto, Redkid	23
USA	WI	producer	2003	GH	2
Mexico	-	wmn	2005	Beryl, Bansi, G122	18
France	-	wmn	2004, 2005	Beryl, Bansi, G122	18
		producer	2012	unk	4
Australia	-	wmn	2004	Beryl, Bansi, G122	4
		producer	2004	Beryl	2

**Table S2.** Mean aggressiveness ratings for Regions with more than five samples; groupings according to 95% family-wise confidence interval.

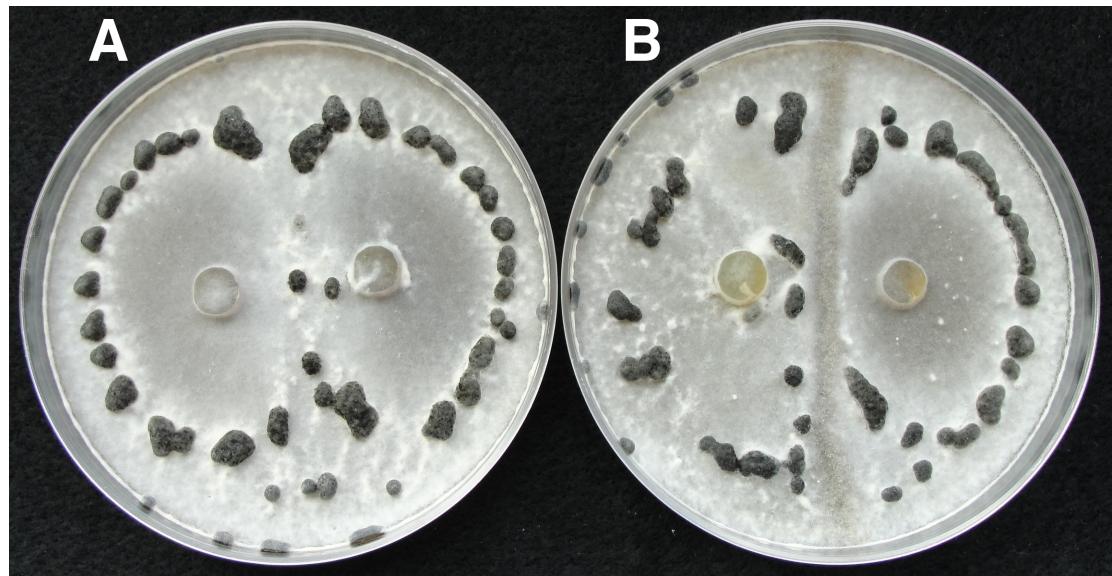
Region	Mean Aggressiveness	Group
MN	5.84	a
ND	5.77	a
NE	5.29	ab
MI	5.13	abc
OR	4.84	abcd
CO	4.72	bcd
WA	4.67	cd
France	4.66	cd
Mexico	4.58	cd
Australia	4.12	cd
CA	4.01	d

**Table S3.** Mean aggressiveness ratings for the 10 most abundant MCG; groupings according to 95% family-wise confidence interval.

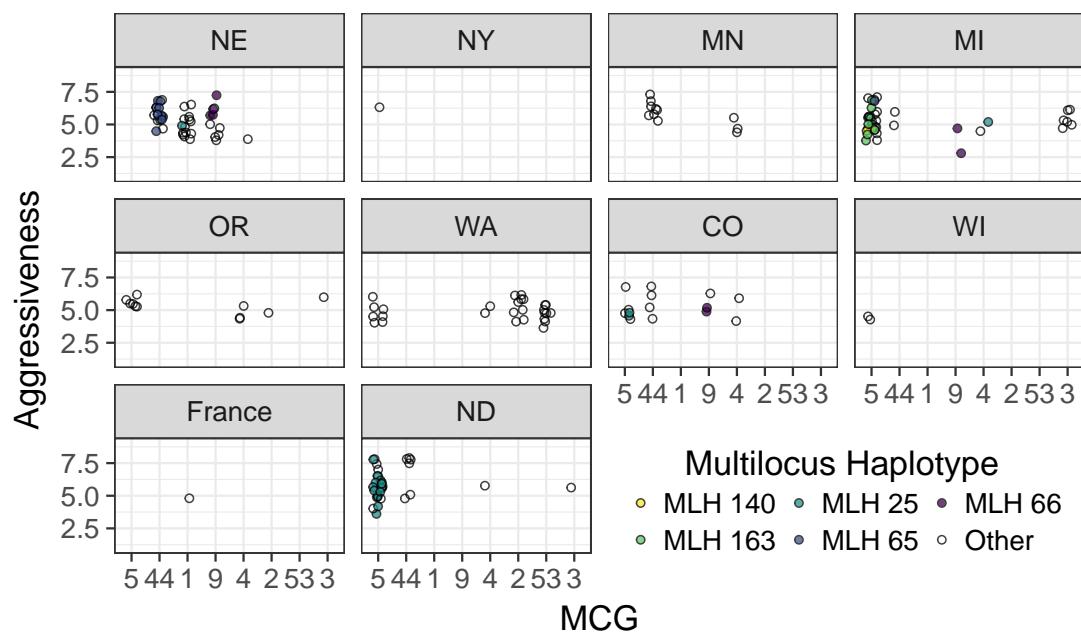
MCG	Mean Aggressiveness	Group
44	6.03	a
3	5.50	ab
5	5.40	b
2	5.25	b
9	5.11	b
1	4.95	b
45	4.88	b
4	4.87	b
53	4.69	b
49	4.60	b

**Table S4.** Mean aggressiveness ratings for the 10 MLH most abundant; groupings according to 95% family-wise confidence interval.

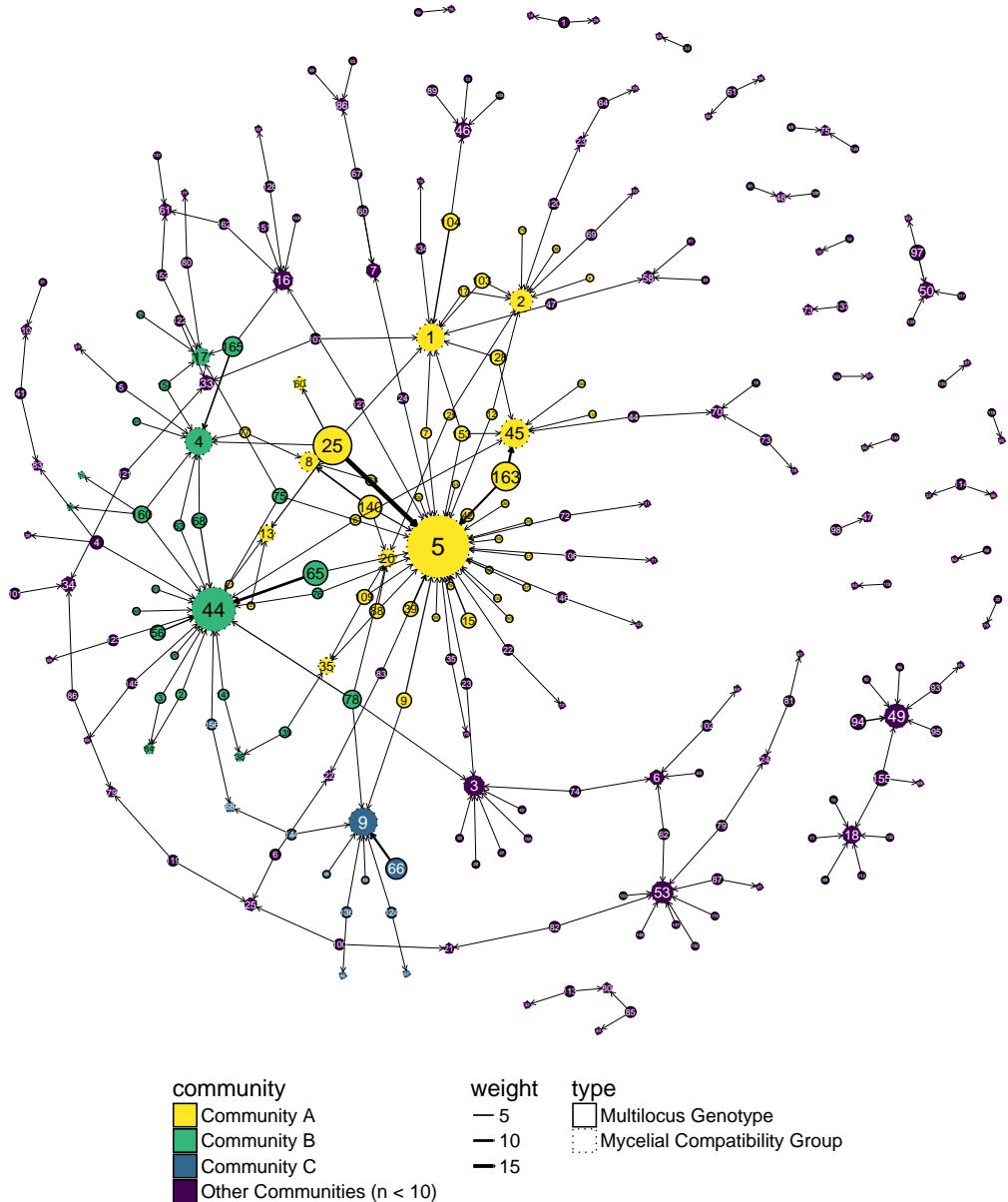
MLH	Mean Aggressiveness	Group
78	6.07	a
65	5.94	a
9	5.67	ab
25	5.41	ab
66	5.30	ab
104	5.22	ab
160	4.80	ab
163	4.80	ab
165	4.34	b
140	4.31	b



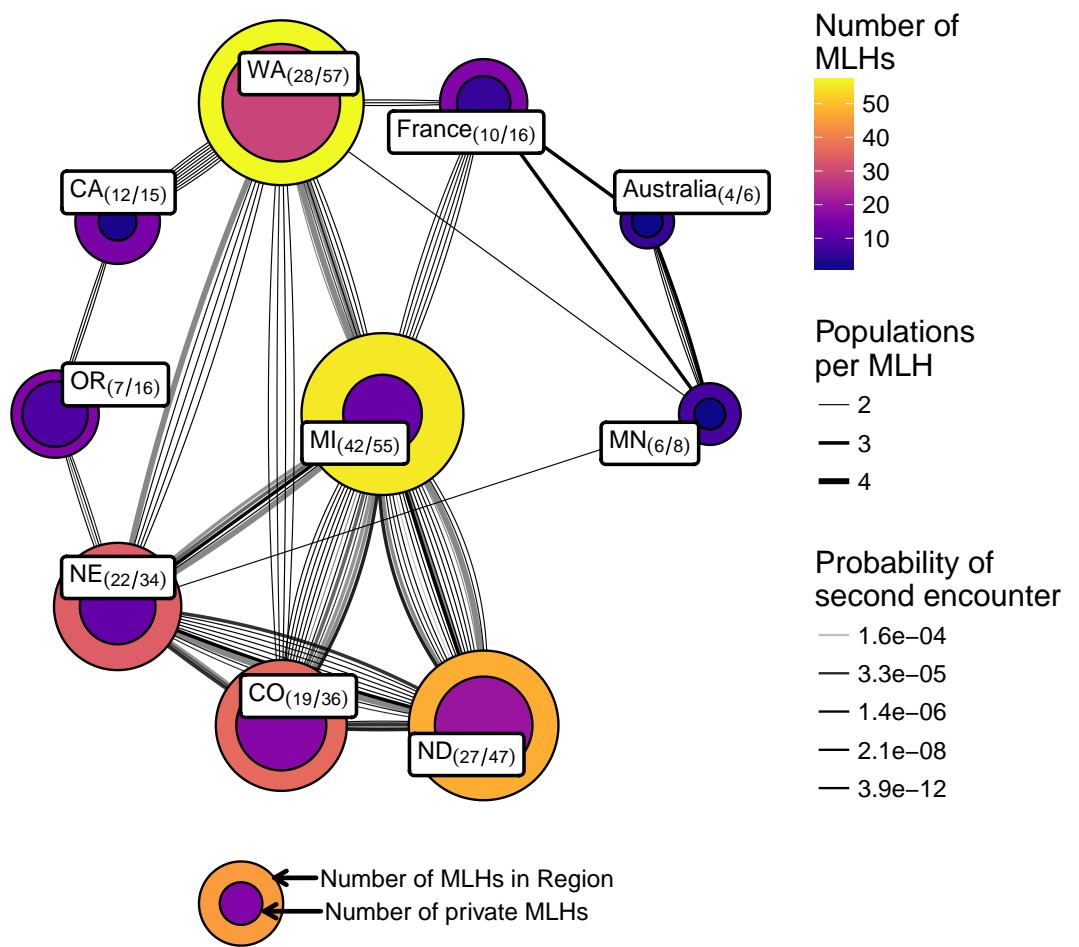
**Figure S1.** Example of MCG test plates showing (A) a compatible reaction with mycelia from two strains overgrowing each other and (B) an incompatible reaction with a barrage line of dead tissue forming between the two strains. Photo Credit: Rebecca Higgins.



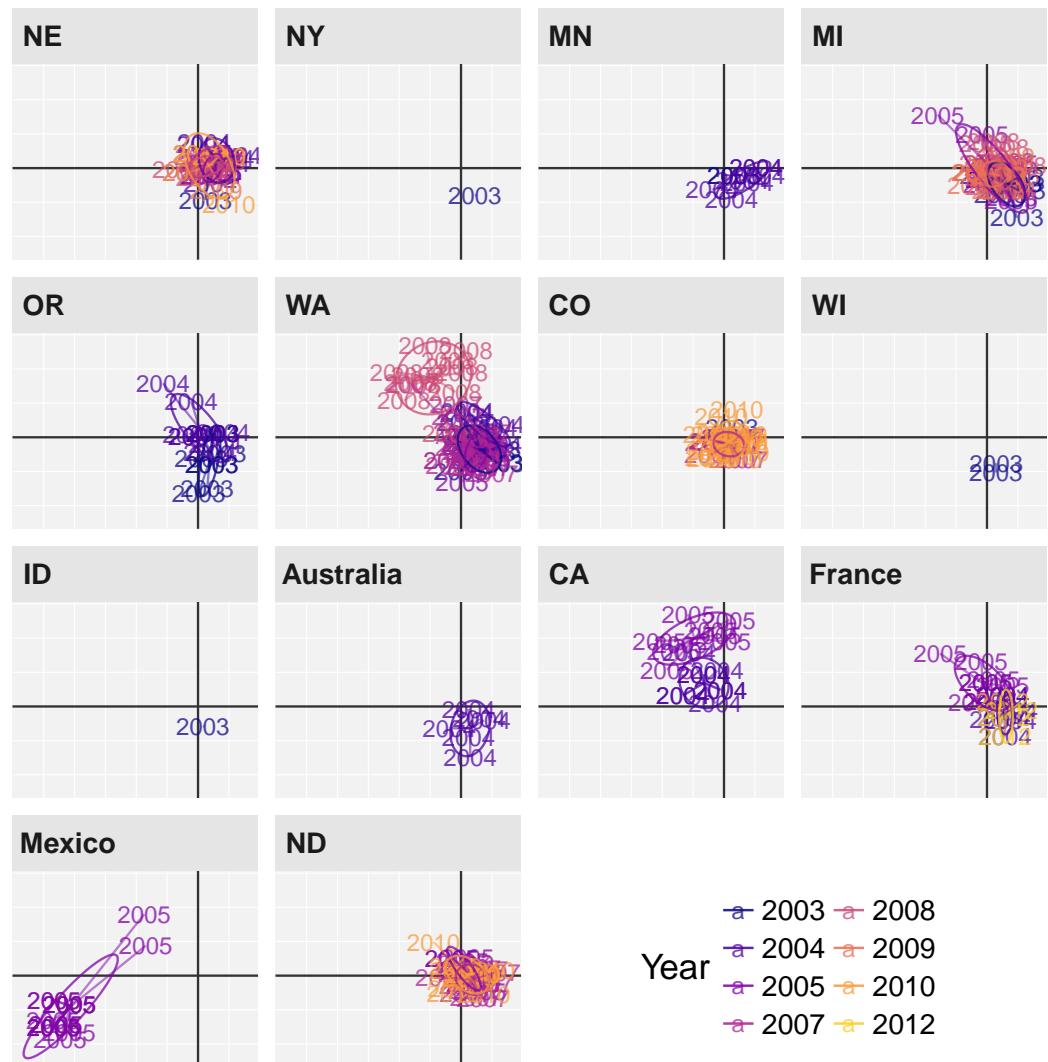
**Figure S2.** Strip plot of aggressiveness for the eight most abundant MCGs partitioned by region. Filled circles indicate one of the five most abundant MLHs and open circles indicate a MLH of lesser abundance.



**Figure S3.** Graph showing complex associations between Mycelial Compatibility Groups (MCG) (dotted nodes) and Multilocus Haplotypes (MLH) (full nodes) where the number in each node represents the MLH/MCG assignment. Node size reflect the number of samples represented by each node (circle). Edges (arrows) point from MLH to MCG where the weight (thickness) of the edge represents the number of samples shared. Node color represents the community assignment based on the walktrap algorithm with a maximum of four steps (Pons & Latapy, 2006). An interactive version of this network can be recreated using the code in the “Interactive visualizations” section of the mlg-mcg.md file in the supplementary information (Direct Link: <https://github.com/everhartlab/sclerotinia-366/blob/master/results/mlg-mcg.md#interactive-visualizations>) (Kamvar et al., 2017).



**Figure S4.** Network of populations (nodes/circles) and their shared multilocus haplotypes (MLH) (edges/lines) haplotyped over 16 loci. Each node is labeled with **name (number of MLHs shared/number of MLHs total)**. The shade and area of the nodes are proportional to the number of unique MLHs within the node and the inner nodes are proportional to the number of private MLHs to the region (bottom legend). Each edge represents a single MLH where its thickness represents the number of populations that share the MLH and the shade represents the value of  $P_{sex}$ , or the probability of encountering that MLH from two independent meiotic events.



**Figure S5.** Scatter plot of Discriminant Analysis of Principal Components on Regions and Years showing temporal variation across all Regions. Points (text labels) represent observed individuals connected to the population centroids with ellipses representing a 66% confidence interval for a normal distribution. The center of each component is represented as black grid lines.

## 562 REFERENCES

- 563 Agapow, P., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, 1, 101–102.
- 564 Aldrich-Wolfe, L., Travers, S., & Nelson, B. D. (2015). Genetic variation of *Sclerotinia sclerotiorum* from multiple crops in the north central United States. *PLOS ONE*, 10(9), e0139188. <https://doi.org/10.1371/journal.pone.0139188>
- 565 Arnaud-Hanod, S., Duarte, C. M., Alberto, F., & Serrão, E. A. (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology*, 16(24), 5115–5139. <https://doi.org/10.1111/j.1365-294X.2007.03535.x>
- 566 Atallah, Z. K., Larget, B., Chen, X., & Johnson, D. A. (2004). High genetic diversity, phenotypic uniformity, and evidence of outcrossing in *Sclerotinia sclerotiorum* in the Columbia Basin of Washington State. *Phytopathology*, 94, 737–742.
- 567 Attanayake, R. N., Carter, P. A., Jiang, D., del Río-Mendoza, L., & Chen, W. (2013). *Sclerotinia sclerotiorum* populations infecting canola from China and the United States are genetically and phenotypically distinct. *Phytopathology*, 103(7), 750–761. <https://doi.org/10.1094/phyto-07-12-0159-r>
- 568 Attanayake, R., Porter, L., Johnson, D., & Chen, W. (2012). Genetic and phenotypic diversity and random association of DNA markers of isolates of the fungal plant pathogen *Sclerotinia sclerotiorum* from soil on a fine geographic scale. *Soil Biology and Biochemistry*, 55, 28–36. <https://doi.org/10.1016/j.soilbio.2012.06.002>
- 569 Boettiger, C., & Eddelbuettel, D. (2017). An introduction to rocker: Docker containers for R. *CoRR*, abs/1710.03675. Retrieved from <http://arxiv.org/abs/1710.03675>
- 570 Boland, G., & Hall, R. (1994). Index of plant hosts of *Sclerotinia sclerotiorum*. *Canadian Journal of Plant Pathology*, 16(2), 93–108. <https://doi.org/10.1080/07060669409500766>
- 571 Bolton, M. D., Thomma, B. P. H. J., & Nelson, B. D. (2006). *Sclerotinia sclerotiorum* (Lib.) de Bary: Biology and molecular traits of a cosmopolitan pathogen. *Molecular Plant Pathology*, 7(1), 1–16. <https://doi.org/10.1111/j.1364-3703.2005.00316.x>
- 572 Botelho, L. d., Zancan, W. L. A., Cruz Machado, J. da, & Barrocas, E. N. (2013). Performance of common bean seeds infected by the fungus *Sclerotinia sclerotiorum*. *Journal of Seed Science*, 35(2), 153–160. <https://doi.org/10.1590/s2317-15372013000200003>
- 573 Brown, A. H. D., Feldman, M. W., & Nevo, E. (1980). Multilocus structure of natural populations of *Hordeum Spontaneum*. *Genetics*, 96(2), 523–536. Retrieved from <http://www.genetics.org/content/96/2/523>
- 574 Bruvo, R., Michiels, N. K., D’Souza, T. G., & Schulenburg, H. (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*, 13(7), 2101–2106.
- 575 Carbone, I., & Kohn, L. M. (2001). Multilocus nested haplotype networks extended with DNA fingerprints show common origin and fine-scale, ongoing genetic divergence in a wild microbial metapopulation. *Molecular Ecology*, 10(10), 2409–2422. <https://doi.org/10.1046/j.0962-1083.2001.01380.x>
- 576 Carbone, I., Anderson, J. B., & Kohn, L. M. (1999). Patterns of descent in clonal lineages and their multilocus fingerprints are resolved with combined gene genealogies. *Evolution*, 53(1), 11–21. <https://doi.org/10.1111/j.1558-5646.1999.tb05329.x>
- 577 Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695. Retrieved from <http://igraph.org>
- 578 Cubeta, M. A., Cody, B. R., Kohli, Y., & Kohn, L. M. (1997). Clonality in *Sclerotinia sclerotiorum* on infected cabbage in eastern North Carolina. *Phytopathology*, 87, 1000–1004.
- 579 Cubeta, M., Sermons, D., & Cody, B. (2001). Mycelial interactions of *Sclerotinia minor*. *Phytopathology*, 91(6S), S19. <https://doi.org/10.1094/phyto.2001.91.6.s1>
- 580 Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12(7), 499–510. <https://doi.org/10.1038/nrg3012>
- 581 Derbyshire, M., Denton-Giles, M., Hegedus, D., Seifbarghy, S., Rollins, J., van Kan, J., Seidl, M. F., Faino, L., Mbengue, M., Navaud, O., Raffaele, S., Hammond-Kosack, K., Heard, S., & Oliver, R. (2017). The complete genome sequence of the phytopathogenic fungus *Sclerotinia sclerotiorum* reveals

- 617 insights into the genome architecture of broad host range pathogens. *Genome Biology and Evolution*,  
 618 9(3), 593–618. <https://doi.org/10.1093/gbe/evx030>
- 619 Ekins, M. G., Hayden, H. L., Aitken, E. A. B., & Goulter, K. C. (2011). Population structure of  
 620 *Sclerotinia sclerotiorum* on sunflower in Australia. *Australasian Plant Pathology*, 40, 99–108.
- 621 Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from  
 622 metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data.  
 623 *Genetics*, 131(2), 479–91.
- 624 Ford, E., Miller, R., Gray, H., & Sherwood, J. (1995). Heterokaryon formation and vegetative  
 625 compatibility in *Sclerotinia sclerotiorum*. *Mycological Research*, 99(2), 241–247.
- 626 Grover, C. E., Salmon, A., & Wendel, J. F. (2012). Targeted sequence capture as a powerful tool for  
 627 evolutionary analysis. *American Journal of Botany*, 99(2), 312–319. <https://doi.org/10.3732/ajb.1100323>
- 628 Grünwald, N. J., Goodwin, S. B., Milgroom, M. G., & Fry, W. E. (2003). Analysis of genotypic  
 629 diversity data for populations of microorganisms. *Phytopathology*, 93(6), 738–746. <https://doi.org/10.1094/phyto.2003.93.6.738>
- 630 Hall, C., Welch, J., Kowbel, D. J., & Glass, N. L. (2010). Evolution and diversity of a fungal  
 631 self/nonself recognition locus. *PLoS ONE*, 5(11), e14055. <https://doi.org/10.1371/journal.pone.0014055>
- 632 Heck, K. L., van Belle, G., & Simberloff, D. (1975). Explicit calculation of the rarefaction diversity  
 633 measurement and the determination of sufficient sample size. *Ecology*, 56(6), 1459–1461. <https://doi.org/10.2307/1934716>
- 634 Henry, L., & Wickham, H. (2017). *purrr: Functional programming tools*. Retrieved from <https://CRAN.R-project.org/package=purrr>
- 635 Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters.  
 636 *Ecology*, 52(4), 577–586. <https://doi.org/10.2307/1934145>
- 637 Jo, Y.-K., Chang, S. W., Rees, J., & Jung, G. (2008). Reassessment of vegetative compatibility of  
 638 *Sclerotinia homoeocarpa* using nitrate-nonutilizing mutants. *Phytopathology*, 98(1), 108–114. <https://doi.org/10.1094/phyto-98-1-0108>
- 639 Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*,  
 640 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- 641 Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components:  
 642 A new method for the analysis of genetically structured populations. *BMC Genetics*, 11:94. <https://doi.org/10.1186/1471-2156-11-94>
- 643 Kamvar, Z. N., Amaradasa, B. S., Jhala, R., McCoy, S., Steadman, J., & Everhart, S. E. (2017,  
 644 November). Data and analysis for population structure and phenotypic variation of *Sclerotinia sclero-*  
 645 *tiorum* from dry bean (*Phaseolus vulgaris*) in the United States. Open Science Framework. <https://doi.org/10.17605/OSF.IO/EJB5Y>
- 646 Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide  
 647 population genetic data with emphasis on clonality. *Frontiers in Genetics*, 6. <https://doi.org/10.3389/fgene.2015.00208>
- 648 Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of  
 649 populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281. <https://doi.org/10.7717/peerj.281>
- 650 Knodel, J., Beauzay, P., Franzen, D., Kandel, H., Markell, S., Osorno, J., Pasche, J., & Zollinger, R.  
 651 (2012). 2012 dry bean grower survey of production, pest problems and pesticide use in Minnesota and  
 652 North Dakota. *North Dakota State University Extension*, E1640.
- 653 Knodel, J., Beauzay, P., Franzen, D., Kandel, H., Markell, S., Osorno, J., Pasche, J., & Zollinger, R.  
 654 (2015). 2015 dry bean grower survey of production, pest problems and pesticide use in Minnesota and  
 655 North Dakota. *North Dakota State University Extension*, E1802.
- 656 Knodel, J., Beauzay, P., Franzen, D., Kandel, H., Markell, S., Osorno, J., Pasche, J., & Zollinger, R.  
 657 (2016). 2016 dry bean grower survey of production, pest problems and pesticide use in Minnesota and  
 658 North Dakota. *North Dakota State University Extension*, E1841.
- 659 Kohli, Y., & Kohn, L. M. (1998). Random association among alleles in clonal populations of  
 660 *Sclerotinia sclerotiorum*. *Fungal Genetics and Biology*, 23, 139–149.
- 661 Kohli, Y., Brunner, L. J., Yoell, H., Milgroom, M. G., Anderson, J. B., Morrall, R. A. A., & Kohn, L.

- 672 M. (1995). Clonal dispersal and spatial mixing in populations of the plant pathogenic fungus, *Sclerotinia*  
673 *sclerotiorum*. *Molecular Ecology*, 4, 69–77.
- 674 Kohn, L. M. (1995). The clonal dynamic in wild and agricultural plant-pathogen populations.  
675 *Canadian Journal of Botany*, 73(S1), 1231–1240. <https://doi.org/10.1139/b95-383>
- 676 Kohn, L. M., Carbone, I., & Anderson, J. B. (1990). Mycelial interactions in *Sclerotinia sclerotiorum*.  
677 *Experimental Mycology*, 14, 255–267.
- 678 Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: Testing multispecies  
679 responses in multifactorial ecological experiments. *Ecological Monographs*, 69, 1–24.
- 680 Lehner, M. S., & Mizubuti, E. S. G. (2017). Are *Sclerotinia sclerotiorum* populations from the tropics  
681 more variable than those from subtropical and temperate zones? *Tropical Plant Pathology*, 42(2), 61–69.  
682 <https://doi.org/10.1007/s40858-016-0125-1>
- 683 Lehner, M. S., Júnior, T. J. P., Júnior, B. T. H., Teixeira, H., Vieira, R. F., Carneiro, J. E. S., & Mizubuti,  
684 E. S. G. (2015). Low genetic variability in *Sclerotinia sclerotiorum* populations from common bean  
685 fields in Minas Gerais State, Brazil, at regional, local and micro-scales. *Plant Pathology*, 64(4), 921–931.  
686 <https://doi.org/10.1111/ppa.12322>
- 687 Lehner, M. S., Paula Júnior, T. J. de, Del Ponte, E. M., Mizubuti, E. S., & Pethybridge, S. J. (2017).  
688 Independently founded populations of *Sclerotinia sclerotiorum* from a tropical and a temperate region have  
689 similar genetic structure. *PloS One*, 12(3), e0173915. <https://doi.org/10.1371/journal.pone.0173915>
- 690 Leslie, J. (1993). Fungal vegetative compatibility. *Annual Review of Phytopathology*, 31, 127–150.  
691 Review. <https://doi.org/10.1146/annurev.py.31.090193.001015>
- 692 McCoy, S., & Steadman, J. R. (2009). Use of multi-site screening to identify partial resistance to  
693 white mold in common bean in 2008. *Bean Improvement Cooperative Annual Report*, 86–87. Retrieved  
694 from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND44207142>
- 695 McDonald, B. A., & Linde, C. (2002). Pathogen population genetics, evolutionary potential, and  
696 durable resistance. *Annual Review of Phytopathology*, 40(1), 349–379. <https://doi.org/10.1146/annurev.phyto.40.120501.101443>
- 697 Mendiburu, F. D., & Simon, R. (2015). Agricolae - ten years of an open source statistical tool for experiments  
698 in breeding, agriculture and biology. <https://doi.org/10.7287/peerj.preprints.1404v1>
- 699 Micali, C. O., & Smith, M. L. (2003). On the independence of barrage formation and heterokaryon  
700 incompatibility in *Neurospora crassa*. *Fungal Genetics and Biology*, 38(2), 209–219. [https://doi.org/10.1016/s1087-1845\(02\)00533-9](https://doi.org/10.1016/s1087-1845(02)00533-9)
- 701 Milgroom, M. G. (1996). Recombination and the multilocus structure of fungal populations. *Annual  
702 Review of Phytopathology*, 34(1), 457–477.
- 703 Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of  
704 individuals. *Genetics*, 89, 583–590.
- 705 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R.,  
706 O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoeecs, E., & Wagner, H. (2017). *Vegan: Community ecology package*. Retrieved from <https://CRAN.R-project.org/package=vegan>
- 707 Otto-Hanson, L., & Steadman, J. R. (2007). Identification of partial resistance to *Sclerotinia sclerotiorum* in common bean at multiple locations in 2006. *Bean Improvement Cooperative Annual Report*, 133–134. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43940892>
- 708 Otto-Hanson, L., & Steadman, J. R. (2008). Identification of partial resistance to *Sclerotinia sclerotiorum* in common bean at multiple locations in 2007. *Bean Improvement Cooperative Annual Report*, 214–215. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND44063230>
- 709 Otto-Hanson, L., Steadman, J. R., Higgins, R., & Eskridge, K. M. (2011). Variation in *Sclerotinia*  
710 *sclerotiorum* bean isolates from multisite resistance screening locations. *Plant Disease*, 95(11), 1370–  
711 1377. <https://doi.org/10.1094/pdis-11-10-0865>
- 712 Papaioannou, I. A., & Typas, M. A. (2014). Barrage formation is independent from heterokaryon  
713 incompatibility in *Verticillium dahliae*. *European Journal of Plant Pathology*, 141(1), 71–82. <https://doi.org/10.1007/s10651-013-0942-0>

- 726 //doi.org/10.1007/s10658-014-0525-3
- 727 Parks, J. C., & Werth, C. R. (1993). A study of spatial features of clones in a population of bracken  
728 fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, 80(5), 537. <https://doi.org/10.2307/2445369>
- 729 Pedersen, T. L. (2017). *ggraph: An implementation of grammar of graphics for graphs and networks*.  
730 Retrieved from <https://CRAN.R-project.org/package=ggraph>
- 731 Petzoldt, R., & Dickson, M. H. (1996). Straw test for resistance to white mold in beans. *Bean  
732 Improvement Cooperative Annual Report*, 39, 142–143. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND20562675>
- 733 Pielou, E. (1975). *Ecological Diversity*. New York: Wiley & Sons.
- 734 Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks.  
735 *Journal of Graph Algorithms and Applications*, 10(2), 191–218. <https://doi.org/10.7155/jgaa.00124>
- 736 Prugnolle, F., & de Meeûs, T. (2010). Apparent high recombination rates in clonal parasitic organisms  
737 due to inappropriate sampling design. *Heredity*, 104(2), 135–140. <https://doi.org/10.1038/hdy.2009.128>
- 738 R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R  
739 Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- 740 Ramasubramaniam, H., del Río Mendoza, L. E., & Bradley, C. A. (2008). Estimates of yield and  
741 economic losses associated with white mold of rain-fed dry bean in North Dakota. *Agronomy Journal*,  
742 100(2), 315. <https://doi.org/10.2134/agronj2007.0127>
- 743 Sambrook, J., Fritsch, E. F., Maniatis, T., & others. (1989). *Molecular Cloning: A Laboratory Manual*.  
744 Cold spring harbor laboratory press.
- 745 Saupe, S. J. (2000). Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes.  
746 *Microbiology and Molecular Biology Reviews*, 64(3), 489–502. <https://doi.org/10.1128/mmb.64.3.489-502.2000>
- 747 Schafer, M. R., & Kohn, L. M. (2006). An optimized method for mycelial compatibility testing in  
748 *Sclerotinia sclerotiorum*. *Mycologia*, 98(4), 593–597. <https://doi.org/10.1080/15572536.2006.11832662>
- 749 Sexton, A. C., & Howlett, B. J. (2004). Microsatellite markers reveal genetic differentiation among  
750 populations of *Sclerotinia sclerotiorum* from Australian canola fields. *Current Genetics*, 46(6), 357–365.  
751 <https://doi.org/10.1007/s00294-004-0543-3>
- 752 Sexton, A. C., Whitten, A. R., & Howlett, B. J. (2006). Population structure of *Sclerotinia sclerotiorum*  
753 in an Australian canola field at flowering and stem-infection stages of the disease cycle. *Genome*, 49(11),  
754 1408–1415. <https://doi.org/10.1139/g06-101>
- 755 Shannon, C. E. (1948). A mathematical theory of communication. *ACM SIGMOBILE Mobile  
756 Computing and Communications Review*, 5(1), 3–55.
- 757 Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163(4148), 688–688. <https://doi.org/10.1038/163688a0>
- 758 Sirjusingh, C., & Kohn, L. M. (2001). Characterisation of microsatellites in the fungal plant pathogen,  
759 *Sclerotinia sclerotiorum*. *Molecular Ecology Notes*, 1(4), 267–269. <https://doi.org/10.1046/j.1471-8278.2001.00102.x>
- 760 Smith, J. M., Smith, N. H., O'Rourke, M., & Spratt, B. G. (1993). How clonal are bacteria?  
761 *Proceedings of the National Academy of Sciences*, 90(10), 4384–4388. <https://doi.org/10.1073/pnas.90.10.4384>
- 762 Steadman, J. R. (1983). White mold - a serious yield-limiting disease of bean. *Plant Disease*, 67,  
763 346–350.
- 764 Steadman, J. R., Eskridge, K., & Powers, K. (2003). Identification of partial resistance to *Sclerotinia  
765 sclerotiorum* in common bean at multiple locations. *Bean Improvement Cooperative Annual  
766 Report*, 225–226. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43757287>
- 767 Steadman, J. R., Otto-Hanson, L., & Breathnach, J. (2006). Identification of partial resistance to  
768 *Sclerotinia sclerotiorum* in common bean at multiple locations in 2005. *Bean Improvement Cooperative  
769 Annual Report*, 223–224. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog>.

- 780       xhtml?id=IND43805570  
781       Steadman, J. R., Otto-Hanson, L., & Powers, K. (2004). Identification of partial resistance to *Sclerotinia sclerotiorum* in common bean at multiple locations. *Bean Improvement Cooperative Annual Report*, 47, 281–282. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43758354>
- 785       Steadman, J. R., Otto-Hanson, L., & Powers, K. (2005). Identification of partial resistance to  
786       *Sclerotinia sclerotiorum* in common bean at multiple locations in 2004. *Bean Improvement Cooperative*  
787       *Annual Report*, 124–125. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43759243>
- 789       Stoddart, J. A., & Taylor, J. F. (1988). Genotypic diversity: Estimation and prediction in samples.  
790       *Genetics*, 118(4), 705–11.
- 791       Strausbaugh, C., & Forster, R. (2003). Management of white mold of beans. *Pacific Northwest Extension, PNW568*. Retrieved from <http://http://www.extension.uidaho.edu/publishing/pdf/PNW/PNW0568.pdf>
- 794       Strom, N. B., & Bushley, K. E. (2016). Two genomes are better than one: History, genetics,  
795       and biotechnological applications of fungal heterokaryons. *Fungal Biology and Biotechnology*, 3(1).  
796       <https://doi.org/10.1186/s40694-016-0022-x>
- 797       Teran, H., Lema, M., Schwartz, H. F., Duncan, R., Gilbeitson, R., & Singh, S. P. (2006). Modified  
798       Petzoldt and Dickson scale for white mold rating of common bean. *Bean Improvement Cooperative An-*  
799       *nual Report*, 49, 115–116. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43805401>
- 801       Tu, J. C., & Beversdorf, W. D. (1982). Tolerance to white mold (*Sclerotinia sclerotiorum* (Lib.) De  
802       Bary) in Ex Rico 23, a cultivar of white bean (*Phaseolus vulgaris* L.). *Canadian Journal of Plant Science*,  
803       62(1), 65–69. <https://doi.org/10.4141/cjps82-010>
- 804       Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.  
805       Retrieved from <http://ggplot2.org>
- 806       Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). *dplyr: A grammar of data manipulation*.  
807       Retrieved from <https://CRAN.R-project.org/package=dplyr>