

# 1 Population structure and phenotypic 2 variation of *Sclerotinia sclerotiorum* from 3 dry bean (*Phaseolus vulgaris*) in the United 4 States

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

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

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

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

10 Corresponding author:

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

12 Email address: everhart@unl.edu

## 13 ABSTRACT

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

## 31 INTRODUCTION

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

39 On dry bean, *S. sclerotiorum* is the causal agent of white mold, a devastating disease that can be yield-  
40 limiting in temperate climates (Steadman, 1983). All above-ground tissues (flowers, stems, leaves, pods)  
41 are susceptible to infection, first appearing as wet lesions with white mycelial tufts, and then bleaching  
42 as the tissue senesces (Steadman, 1983; Bolton et al., 2006). For many years, white mold has been the  
43 most serious dry bean disease in the Northwestern United States (Otto-Hanson et al., 2011; Knodel et al.,  
44 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  
46 al., 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 ~~Mexico~~Australia,  
50 France, and ~~Australia~~Mexico (Steadman et al., 2003, 2004, 2005, 2006; Otto-Hanson & Steadman, 2007,  
51 2008; McCoy & Steadman, 2009). These white mold screening nurseries use no chemical or cultural  
52 treatments against *S. sclerotiorum* and employ standardized protocols for screening new cultivars for  
53 resistance to white mold. ~~It has previously been~~ (Steadman et al., 2003; Otto-Hanson et al., 2011). These  
54 protocols included three established cultivars used for comparison in the trials: Beryl (great northern  
55 bean, susceptible), Bunsí (a.k.a. Ex Rico, navy bean, low susceptibility), and G122 (cranberry bean,  
56 partial resistance) (Tu & Beversdorf, 1982; Steadman et al., 2005; Otto-Hanson et al., 2011). It was  
57 previously shown that aggressiveness (the severity of disease symptoms on the host) is significantly  
58 different across white mold screening nursery sites in separate geographic regions (Otto-Hanson et  
59 al., 2011). The genetic structure and mode of reproduction in these populations, however, is currently  
60 unknown.

61 Understanding genetic relationships and reproduction behavior of *S. sclerotiorum* populations is  
62 beneficial for breeders seeking to develop new resistant cultivars for worldwide deployment (Milgroom,  
63 1996; McDonald & Linde, 2002). In particular, genetically diverse populations with high rates of sexual  
64 reproduction are more likely to overcome host resistance. Most populations of *S. sclerotiorum* are  
65 predominantly clonal with low genetic diversity and have a large degree of population fragmentation  
66 (Kohli et al., 1995; Cubeta et al., 1997; Kohli & Kohn, 1998; Carbone & Kohn, 2001; Ekins et al., 2011;  
67 Attanayake et al., 2012). Some studies, however have found populations that show signatures of sexual  
68 reproduction (Atallah et al., 2004; Sexton & Howlett, 2004; Attanayake et al., 2013; Aldrich-Wolfe et al.,  
69 2015).

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

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

## 95 **MATERIALS AND METHODS**

### 96 **Isolate collection**

97 Several (156) of the isolates used for this study were collected as reported in previous studies using  
98 the same methods (Otto-Hanson et al., 2011). Broadly, isolates were collected from two sources: white

99 mold screening nurseries (wmn) or producer fields. White mold screening nurseries were 5m x 10m in  
100 size and maintained without application of fungicides to observe natural incidence of white mold. The  
101 early nursery plots were incorporated with a basal dressing of N:P:K = 1:3:2 and side dressing of 0:3:2  
102 during the growing season (Steadman et al., 2003).

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

107 Sclerotia of *S. sclerotiorum* were collected over several years from grower fields and/or wmn in 11  
108 states of the ~~United States of America, Mexico, France, and Australia~~ Australia, France, Mexico, and the  
109 United States (Table ??). After collection, sclerotia were stored in Petri plates lined with filter paper,  
110 then stored at 20 °F or -4 °C. Sclerotia were surface-sterilized with 50% Clorox bleach (at least 6%  
111 ~~NaOCL~~NaOCl, The Clorox Company, Oakland, CA) solution for 3 min, and double rinsed with ddH<sub>2</sub>O  
112 for 3 min. The sterilized sclerotia were then placed on water agar plates (16g of Bacto agar per liter of  
113 ddH<sub>2</sub>O, BD Diagnostic Systems, Sparks, MD), with four to five sclerotia of each isolate separated on  
114 each plate and stored on the counter top at room temperature for 5 to 6 days. An 8-mm plug from a 5- or  
115 6-day-old culture was transferred from the advancing margin of the mycelia onto a plate of Difco potato  
116 dextrose agar (PDA at 39 g/liter of ddH<sub>2</sub>O) (Otto-Hanson et al., 2011). In combination with the 156  
117 isolates described previously, we collected 210 isolates for a total of 366 isolates (Otto-Hanson et al.,  
118 2011).

### 119 **Mycelial ~~Compatibility~~compatibility**

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

### 127 **Aggressiveness**

128 Aggressiveness of each isolate was assessed using a straw test as described in Otto-Hanson et al.  
129 (2011) that rated necrotic lesion size (Petzoldt & Dickson, 1996; Teran et al., 2006). Briefly, the straw  
130 test uses 21-day-old G122 plants as the host in a greenhouse setting. Clear drinking straws cut to 2.5 cm  
131 and heat sealed were used to place two mycelial plugs of inoculum on the host plant after removing plant  
132 growth beyond 2.5 cm above the fourth node. Measurements of the necrotic lesion were taken 8 days later  
133 using the Modified Petzoldt and Dickson scale of 1–9, where 1 is no disease and 9 is plant death (Petzoldt  
134 & Dickson, 1996; Teran et al., 2006).

### 135 **Microsatellite genotyping**

136 Prior to DNA extraction, isolates were grown on PDA and plugs were subsequently transferred to  
137 Potato Dextrose Broth (PDB) where they were grown until there was significant mycelial growth, but  
138 before the mycelial mat became solidified (4–5 days). Each mycelial mat was collected in a filtered  
139 Büchner funnel, agar plugs removed, lyophilized and pulverized manually in Whirl-pak® HDPE sampling  
140 bags (Sigma-Aldrich, St. Louis, MO). Lyophilized mycelia was then stored in microcentrifuge tubes at  
141 -20 °C until needed for DNA extraction. DNA from 25mg of pulverized mycelia was purified using a  
142 phenol-chloroform extraction method followed by alcohol precipitation and evaporation, suspending the  
143 DNA in 200µl TE (Sambrook et al., 1989). Suspended DNA was stored at 4 °C until genotyping.

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

151 PeakScanner version 1.0 (Life Technologies Corporation, Carlsbad, CA) and recorded manually in a  
152 spreadsheet.

### 153 **Data processing and ~~Analysis~~analysis**

154 All data processing and analyses were performed in a [Rocker “verse” project container running](#)  
155 R version 3.4.1-2 (Boettiger & Eddelbuettel, 2017; R Core Team, 2017) and are openly available  
156 and reproducible at <https://github.com/everhartlab/sclerotinia-366/>. Of the 16  
157 microsatellite loci genotyped, five included compound repeats, which made it challenging to accu-  
158 rately/confidently bin alleles into fragment sizes expected for each locus based on the described repeat  
159 motif. Loci with compound repeats were removed for the reported statistics. To ensure the integrity of the  
160 results we additionally processed these loci and included them in concurrent analyses. We assessed the  
161 power of our 11 markers by generating a genotype accumulation curve in the R package *poppr* version  
162 2.5.0, looking for evidence of saturation, which would indicate that loci were sufficiently sampled to  
163 adequately represent the full set of haplotypes (Arnaud-Hanod et al., 2007; Kamvar et al., 2015). To avoid  
164 including isolates potentially collected from the same plant (which increases the probability of collecting  
165 sclerotia from the same point of infection more than once), data were clone-corrected on a hierarchy  
166 of Region/Source/Host/Year—meaning that duplicated genotypes were reduced to a single observation  
167 when they were collected in the same year from the same host cultivar located in the same source field  
168 (wmn or producer)—for subsequent analysis. We assessed haplotype diversity by calculating Stoddart and  
169 Taylor’s index ( $G$ ) (Stoddart & Taylor, 1988), Shannon’s index ( $H$ ) (Shannon, 1948), Simpson’s index ( $\lambda$ )  
170 (Simpson, 1949), evenness ( $E_s$ ), and the expected number of multilocus haplotypes ( $eMLH$ ) (Hurlbert,  
171 1971; Heck et al., 1975; Pielou, 1975; Grünwald et al., 2003). To assess the potential for random mating,  
172 we tested for linkage disequilibrium with the index of association,  $I_A$  and its standardized version,  $\bar{r}_d$   
173 using 999 permutations (Brown et al., 1980; Smith et al., 1993; Agapow & Burt, 2001). Both haplotype  
174 diversity and linkage disequilibrium were calculated in *poppr* (Kamvar et al., 2014).

### 175 **Assessing ~~Importance~~importance of ~~Variables~~variables**

#### 176 **~~Distance-based Redundancy Analysis~~redundancy analysis**

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

#### 192 **~~Aggressiveness Assessment~~assessment**

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

#### 200 **~~Correlating Multilocus Haplotypes~~multilocus haplotypes with ~~Mycelial Compatibility Groups~~mycelial** 201 **~~compatibility groups~~**

202 We wanted to assess if there was correlation between MLHs and MCGs. This was performed using a  
203 network-based approach where both MLHs and MCGs were considered nodes and the number of isolates

204 in which they were found together was the strength of the connection between an MLH and and MCG  
205 node. The network-based approach allowed us to assess the associations between MLHs and MCGs.  
206 To construct the network, a contingency table was created with MLHs and MCGs and converted to a  
207 directed and weighted edgelist where each edge represented a connection from an MCG to an MLH,  
208 weighted by the number of samples shared in the connection. This was then converted to a bipartite  
209 graph where top nodes represented MLHs and bottom nodes represented MCGs. To identify clusters of  
210 MLHs and MCGs within the network, we used the cluster\_walktrap community detection algorithm as  
211 implemented in the `cluster_walktrap()` function in *igraph* version 1.1.2 (Csardi & Nepusz, 2006;  
212 Pons & Latapy, 2006). This algorithm attempts to define clusters of nodes by starting at a random ~~nodes~~  
213 node and performing short, random “walks” along the edges between nodes, assuming that these walks  
214 would stay within clusters. For this analysis, we set the number of steps within a walk to four and allowed  
215 the algorithm to use the edge weights in determining the path. All of the resulting communities that had  
216 fewer than 10 members were then consolidated into one. Community definitions were used to assess the  
217 average genetic distance (as defined by Bruvo’s distance) within members of the community (Bruvo et al.,  
218 2004).

## 219 **Genetic Diversity**diversity

### 220 **Population Differentiation**differentiation

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

### 237 **Analysis of Shared Multilocus Haplotypes**shared multilocus haplotypes

238 We wanted to evaluate patterns of connectivity between shared multilocus haplotypes across geo-  
239 graphic regions. We first tabulated the multilocus haplotypes shared between at least two populations  
240 (defined as states or countries) with the *poppr* function `mlg.crosspop()` (Kamvar et al., 2014). From  
241 these data, we constructed a graph with populations as nodes and shared haplotypes as edges (connections)  
242 between nodes using the R packages *igraph* (Csardi & Nepusz, 2006), *dplyr* version 0.7.4 (Wickham et al.,  
243 2017), and *purrr* version 0.2.~~3~~.4 (Henry & Wickham, 2017). Each node was weighted by the fraction of  
244 shared MLHs. Each edge represented a single MLH, but because a single MLH could be present in more  
245 than one population, that MLH would have a number of edges equivalent to the total number of possible  
246 connections, calculated as  $(n*(n-1))/2$  edges where  $n$  represents the number of populations crossed. Edges  
247 were weighted by  $1 - P_{sex}$ , where  $P_{sex}$  is the probability of encountering the same haplotype via two  
248 independent meiotic events (Parks & Werth, 1993; Arnaud-Hanod et al., 2007). This weighting scheme  
249 would thus strengthen the connection of edges that represented genotypes with a low probability of being  
250 produced via sexual reproduction. We then identified communities (among the Regions) in the graph  
251 using the `cluster_optimal()` function from *igraph* (Csardi & Nepusz, 2006). The graph was plotted  
252 using the R packages *ggplot2* version 2.2.1 (Wickham, 2009) and *ggraph* 1.0.0 (Pedersen, 2017). To  
253 ensure that we captured the same community signal, we additionally performed this analysis including  
254 the five polymorphic markers described above.



## RESULTS

A total of 366 isolates were collected from 2003 to 2012 (except 2006 and 2011) from diseased dry bean plants in ~~Australia, Mexico, 11 states in the United States~~ as well as Australia, France, and Mexico (Table ??). ~~With the 11 states in the USA (Table ??). We~~ loci used in the analyses (Table ??), we observed a total of 165 MLHs (215 with 16 loci). ~~These 11 loci are located on 7 chromosomes in the *S. sclerotiorum* genome with a minimum distance of 55Kbp between two loci on the same chromosome.~~ Over 50% of the isolates came from four states, MI (62), ND (60), WA (59), NE (47). Four regions had fewer than 10 isolates, Australia (6), WI (2), NY (1), ID (1). We observed 87 MCGs, the most abundant of which ('MCG 5') was represented by 73 isolates over 37 MLHs (Fig. ??A,C).

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

After clone-correction on the hierarchy of Region/Source/Host/Year, ~~we were left with a total of 48 isolates were removed from the data set, resulting in 318 isolates representing 165 MLHs that were used in subsequent analyses (Table ??).~~ The results ~~show that most populations exhibited relatively high amounts showed that, in terms of genotypic diversity, with the exception of ( $H$ ,  $G$ , and  $\lambda$ ), WA was the most diverse population with both  $G$  (54.3) and  $e^H$  (55.3) being close to the observed number of MLHs (56). This indicated that there are few duplicated genotypes in WA (Table ??). A more useful metric to compare populations, however, is  $E_5$ , which scales from 0 to 1, where 1 indicates all unique genotypes (Grünwald et al., 2003). Evaluating by  $E_5$  shows that both MI and NE, which exhibited lower than average  $E_5$  values. Mexico had the lowest average value of gene diversity, exhibit lower than average values, indicating that there are over-represented genotypes in the populations (table ??). When we look at Mexico, we observed that it had relatively high values of  $E_5$  and genotypic diversity, but low richness, as measured by ~~heMLG and the lowest~~. Moreover, Mexico had the lowest value for ~~eMLHh~~, suggesting low overall genetic which is a measure of allelic diversity. Nearly all populations showed evidence for of linkage (Table ??), which serves as evidence for clonal reproduction or other forms of non-random mating. The only exceptions ~~are were~~ CA ( $P = 0.043$ ) and Australia ( $P = 0.052$ ). 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	di-(GT)	4	0.45	0.62
6-2	483–495	hexa-(TTTTTC)(TTTTTG)(TTTTTC)	3	0.64	0.95
7-2	158–174	di-(GA)	7	0.73	0.76
8-3	244–270	di-(CA)	7	0.74	0.79
9-2	360–382	di-(CA)(CT)	9	0.35	0.41
12-2	214–222	di-(CA)	5	0.58	0.78
17-3	342–363	tri-(TTA)	7	0.55	0.53
20-3	280–282	di-(GT)GG(GT)	2	0.05	0.42
55-4	153–216	tetra-(TACA)	10	0.72	0.66
110-4	370–386	tetra-(TATG)	5	0.76	0.91
114-4	339–416	tetra-(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 asterisk 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	-

## 286 Variable Assessment

### 287 Variable Contributions

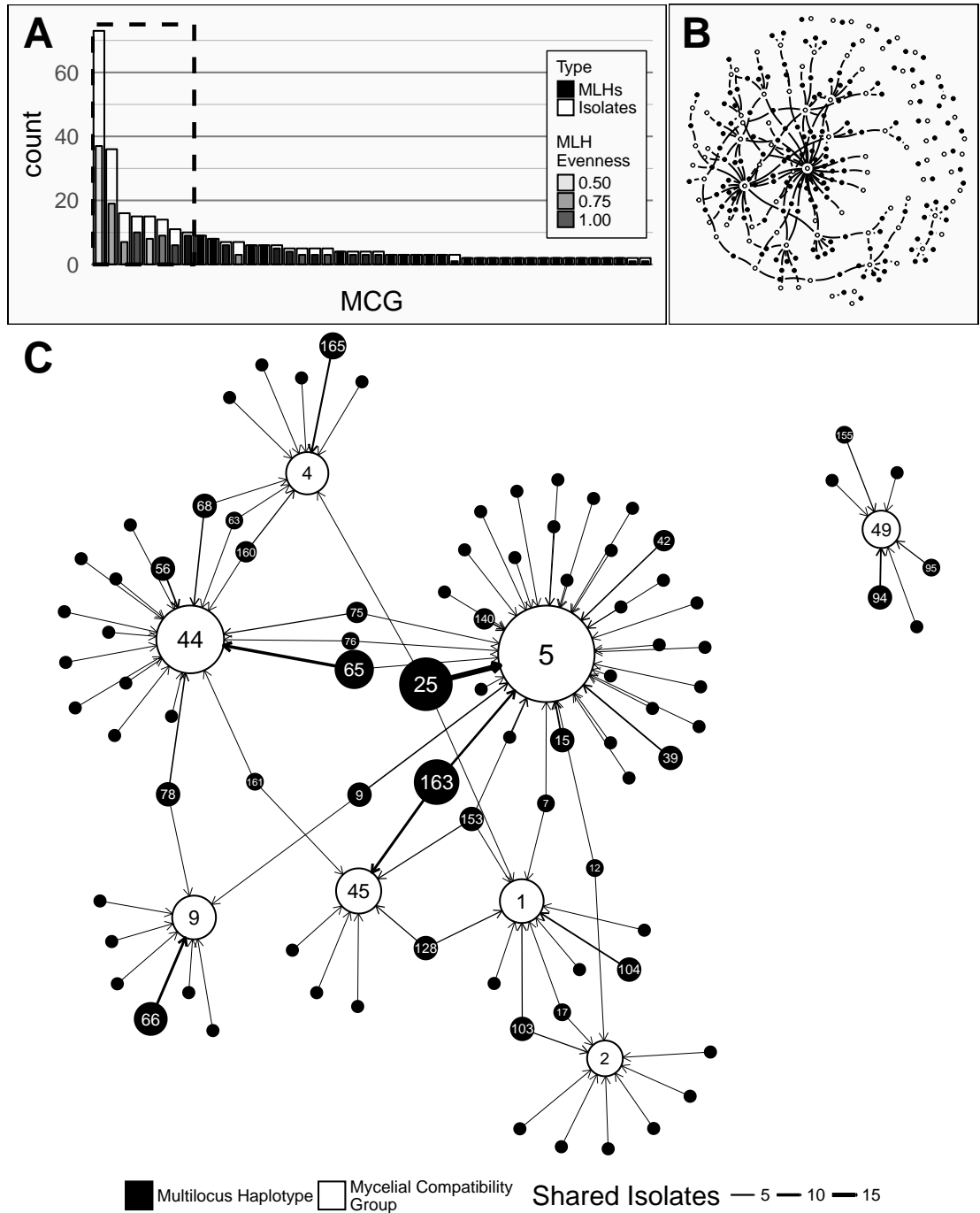
288 The forward-backward selection process of the dbrDA models on clone-corrected data revealed Year,  
 289 Region, Host, and MCG to be the optimal variables for the reduced model, accounting for 45% of the  
 290 total variation. ANOVA showed that the reduced model was significant with an adjusted  $R^2$  of 0.0675 ( $P$   
 291 = 0.001). Assessment of the marginal effects showed that all variables significantly explained  
 292 genetic variation ( $P \leq 0.007$ ). We found that there was multicollinearity when MCG was combined  
 293 with any other variable, so repeated the analysis, dropping MCG from the list of potential predictors.  
 294 From these results, Year, Region, Host, and Aggressiveness were found to be optimal, accounting for  
 295 17.6% of the total variation. ANOVA revealed significant effects with an adjusted  $R^2$  of 0.0325 ( $P$   
 296 = 0.001). While the marginal effect assessment revealed that Year, Region, and Host significantly explained  
 297 variation at  $P = 0.001$ , and Aggressiveness significantly explained variation at  $P = 0.039$ . Much of the  
 298 variation appeared to be driven by isolates from Mexico and 2005 (Fig. ??). Variance partitioning of  
 299 the independent variables without MCG indicated aggressiveness to be the least influential factor with  
 300 0.1% contributing to explaining the variation of molecular data, whereas the combination of variables  
 301 accounted for 3.3%.

### 302 Aggressiveness

303 Aggressiveness of the isolates ranged from 1.4 to 7.9 with a mean of 5.02 and median  
 304 of 4.85. The group mean averages were 4.88, 5.13, and 5.19 for Region, MCG, and MLH, respectively. A  
 305 strip plot showing the distribution of severity across these three variables simultaneously can be seen in  
 306 Fig. ?. Our assessment of aggressiveness in association with Region showed a significant effect ( $P <$   
 307  $1.00e^{-4}$ ), with means that ranged from 5.8 (MN) to 4.0 (CA) (Fig. ??, Table ??). MCGs also showed a  
 308 significant effect ( $P < 1.00e^{-4}$ ), with means that ranged from 6.0 ('MCG 44') to 4.6 ('MCG 49';  
 309 Table ??). We additionally found a significant effect for MLHs ( $P = 7.44e^{-4} < 0.001$ ), with means that  
 310 ranged from 6.0 ('MLH 78') to 4.3 ('MLH 140') (Table ??).

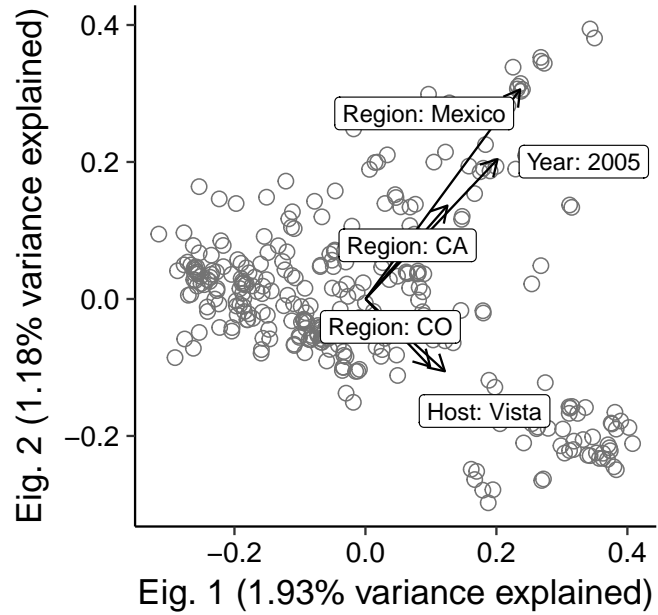
### 311 Correlation of Multilocus Haplotypes and Mycelial Compatibility 312 Groups

313 In our analysis, we found 165 MLHs with 70 singletons and 87 MCGs with 43 singletons (Fig. ??A,B)  
 314 where the eight most abundant MCGs represented > 51% of the data over 11 Regions, and all years except

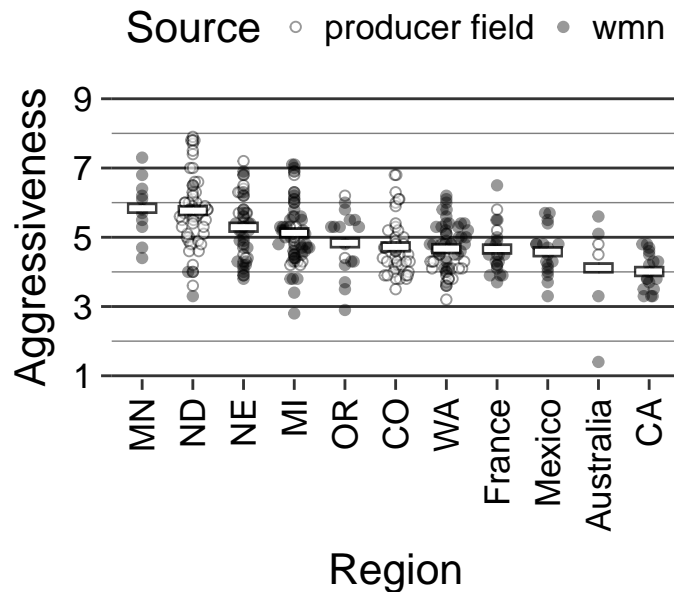


**Figure 1.** Associations between Mycelial Compatibility Groups and Multilocus Haplotypes. **A)** Barplot of Mycelial Compatibility Group (MCG) abundance in descending order. Singletons (46) were truncated, leaving 41 MCGs. White bars represent sample counts and grey bars represent counts of unique 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. ?? **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.

315 for 2012. Our network-based approach to correlating MLHs with MCGs revealed a large and complex  
 316 network (Fig. ??, Table ??). Community analysis showed 51 communities, 15 of which consisted of  
 317 a single MLH unconnected with any other community indicating that just 9.09% of the 165 MLHs are  
 318 unable to cross with any other MLH in this data set (Fig. ??). The three communities with the most  
 319 members contained eight of the 10 most abundant MCGs. Comparing these communities with Bruvo's  
 320 genetic distance showed an average distance of 0.451 among communities and an average distance of  
 321 0.437 within communities, which were not significantly different. When we assessed the number of times  
 322 two different MLHs that are in the same MCG, considering these as potential heterothallic pairings that  
 323 could result in sexual recombination, we found an average of 14.3 potential heterothallic pairings per MLH.  
 324 Representing just four isolates, 'MLH 75' had 57 neighbors that shared the same **mycelial compatibility**  
 325 **group** (Fig. MCG (Fig. ??, ??). Overall, there was no clear pattern to the association between MLH and  
 326 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)	

### 327 **Structure of Shared Multilocus Haplotypes** shared multilocus haplotypes

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

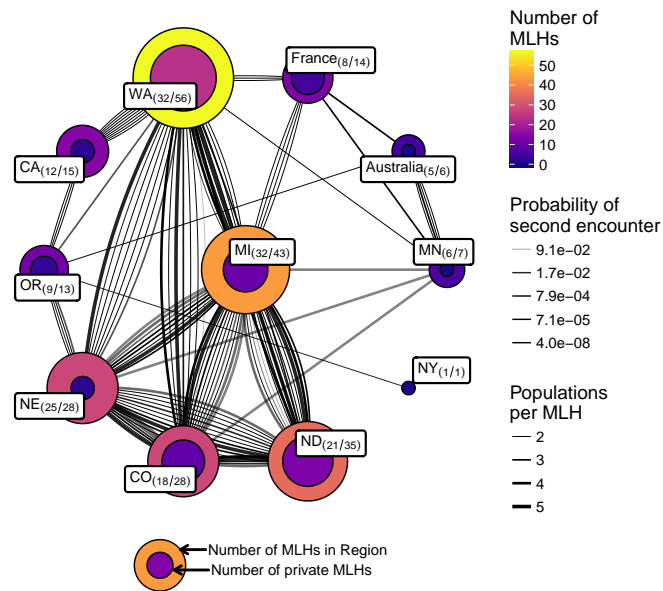
332 We had performed an analysis on a network where the connections represented shared MLHs across  
 333 populations, weighted by  $1 - P_{sex}$  (Fig. ??, Table ??). Community analysis of the MLHs shared between  
 334 populations revealed 4 communities with a modularity of 0.17: A coastal community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

### 341 **Population Differentiation** differentiation

#### 342 **Analysis of Molecular Variance** molecular variance

343 The ~~analysis of molecular variance (AMOVA)~~ AMOVA for clone-corrected samples over the hierar-  
 344 chy of Region, Source, and Year showed significant variation between Regions and Years, but no significant  
 345 variation between wmn and producer fields (Table ??). In contrast, when we compared the three cultivars,  
 346 Beryl, Bunsil, and G122, we found no significant differentiation (**Supplementary Information**) See section  
 347 on 'Host Differentiation' in the wmn-differentiation.md<sup>2</sup> file in the supplemental files (Kamvar et al.,

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

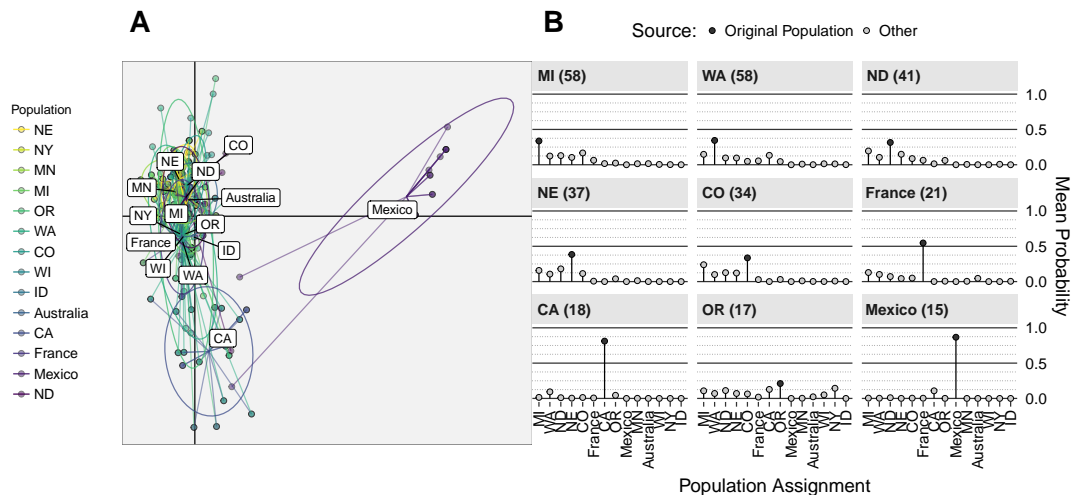


**Figure 4.** Network of populations (nodes/circles) and their shared multilocus haplotypes (MLH) (edges/lines) genotyped over 11 loci. Each node is labeled with **name (number of MLHs shared/number of MLHs total)**. The shade and area of the nodes are proportional to the number of unique MLHs within 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.

348 [2017](#)).

**Table 4.** Comparison of populations in the white mold screening nurseries (wmn) and producer fields using an analysis of molecular variance ([AMOVA](#)) 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>	<b>0.0290.031</b>
Between Source within Region	8	2.74	-2.29	-0.0250	<b>0.4860.497</b>
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 (DAPC) 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.

349 **Discriminant Analysis of Principal Components**

350 Discriminant Analysis of Principal Components (DAPC) was performed grouping by DAPC was  
 351 performed by grouping Region with the first 21 principal components, representing 88.1% of the total  
 352 variance. The first discriminant axis (representing 36.163.9% of the discriminatory power) separated the  
 353 centroid for the Mexico isolates from the rest of the data, indicating strong differentiation (Fig. ??b). The  
 354 second discriminant axis, representing 25.310.8% of the discriminatory power, separated the centroid for  
 355 the CA isolates. The mean population assignment probabilities for all populations with  $n > 10$  showed  
 356 that only isolates from Mexico, CA, and France had  $> 50\%$  probabilities of being reassigned to their  
 357 source populations (Fig. ??a).

358 DAPC grouping by cultivar used the first 20 principal components, representing 89% of the total  
 359 variance. The first two discriminant axes (representing 38.5100% of the discriminatory power) failed to  
 360 separate any of the cultivars where the mean posterior assignment probabilities were 34% (G122), 35.9%  
 361 (Beryl), and 30.1% (Bunsi). DAPC grouping by Region and Year used the first 15 principal components,  
 362 representing 80.3% of the total variance. The North Central USA populations (NE, MI, CO, ND) did not  
 363 appear to have any variation across time in contrast to WA, which showed a shift in population structure in  
 364 the last year of sampling, 2008 (Fig. ??). Further analysis of this population revealed that all 12 isolates  
 365 in WA circa 2008 originated in a wmn; nine haplotypes were shared with CA, and three were shared with  
 366 France (Fig. ??, ??).

367 **DISCUSSION**

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



**Figure 6.** Scatter plot of Discriminant Analysis of Principal Components (DAPC) 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. ??.

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

397 While aggressiveness was not shown to predict genetic structure, it is an important factor in breeding  
 398 efforts, and we observed significant differences in aggressiveness based on Region (Fig. ??, Table ??).  
 399 These results show a similar pattern to what was found previously in Otto-Hanson et al. (2011) with the  
 400 exception of North Dakota, which increased in mean aggressiveness from 5 to 5.77. This increase was  
 401 due in part to new data from producer field isolates collected after the previous study. These straw tests  
 402 were performed by a different person for these later isolates, which could suggest a more lenient or strict  
 403 scoring system. However, when we examined the within-region differences, we found no significant effect  
 404 by individual. Many of the ND isolates fell within the 6–7 range, which denotes a physical boundary  
 405 (disease symptoms around the second node) between intermediate and susceptible (Otto-Hanson et al.,  
 406 2011). Thus, we observed a shift in aggressiveness without a significant shift in genotypic structure,  
 407 which may indicate that aggressiveness may be controlled by environmental factors as opposed to genetic



408 profile.

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

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

437 With the exception of the WA Region, populations that were sampled across several years appeared to  
438 be relatively stable over time with overlapping distributions in the DAPC (i.e. NE and MI, Fig. ??). DAPC  
439 is based on the principal components of ~~Euclidean distance between genotypes-allele counts~~ (Jombart et  
440 al., 2010). Unlike Bruvo's distance, ~~Euclidean distance~~ this does not take into account the magnitude  
441 of the difference between alleles, which could inflate the distance measure in the presence of private  
442 alleles (Bruvo et al., 2004). ~~We-While we~~ found no evidence of private alleles in the Mexico and CA  
443 isolates, ~~suggesting that this-we did find that the alleles driving the first axis in Fig. ??A (alleles 174,~~  
444 ~~256, and 372 in loci 7-2, 8-3, and 9-2, respectively) were overrepresented in Mexico (where >75% of~~  
445 ~~the alleles came from the region). However, all three of these alleles, i) conform to the expected stepwise~~  
446 ~~mutation model (Bruvo et al., 2004) and ii) are at or near the extremes of the total range (except for allele~~  
447 ~~372 at locus 9-2). Moreover, the fact that we find three alleles at three independent loci segregating the~~  
448 ~~Mexican genotypes suggests that the~~ pattern separating these populations from the others was not an  
449 artifact. We believe that the differences ~~we observe in populations observed~~ from Mexico may ~~reflect-be~~  
450 ~~due to~~ differences in climate that allow greater diversification via sexual outcrossing.

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

---

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

461 The results from our shared haplotype analysis ~~show~~ showed several populations with at least one  
462 haplotype between them, except for Mexico and two states that had fewer than three samples each (Fig.  
463 ~~??~~). Our network-based approach by treating the haplotypes as edges and weighting each edge with the  
464 inverse of  $P_{sex}$  treated the edges as springs connecting the populations with the strength proportional to the  
465 probability of obtaining the same haplotype as a clone. This allowed us to use a graph walking algorithm  
466 to see how close the populations were, simply based off of the proportion of clones they shared. The  
467 most abundant haplotype was shared across four populations, but its high value of  $P_{sex}$  meant that it did  
468 not contribute significantly to the overall structure. The graph walking algorithm was able to divide the  
469 network into three groups, but had a modularity of 0.17, which indicates that the groups are only weakly  
470 differentiated.

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

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

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

501 Over the past few years, researchers have noticed inconsistencies among the relationship between  
502 MCGs and MLHs (Carbone et al., 1999; Attanayake et al., 2012; Aldrich-Wolfe et al., 2015; Lehner et al.,  
503 2015). Either several MCGs belong to one MLH, which could be explained by insufficient sampling of  
504 loci; several MLHs belong to one MCG, which could be explained by clonal expansion; or a mixture of  
505 both. Some studies have shown a correlation between MCG and MLH (Carbone et al., 1999; Aldrich-  
506 Wolfe et al., 2015; Lehner et al., 2015), whereas other studies have shown no apparent correlation, even  
507 on small spatial scales (Atallah et al., 2004; Attanayake et al., 2012, 2013).

508 One long-held assumption was that MCGs (as determined via barrage reaction) represent vegetative  
509 compatibility groups (VCGs) (Kohn et al., 1990; Schafer & Kohn, 2006; Lehner et al., 2015), which are  
510 known to have a genetic component (Saupe, 2000; Hall et al., 2010; Strom & Bushley, 2016). While  
511 our protocol for assessing MCGs utilized Diana Sermons Medium (Cubeta et al., 2001) as compared to  
512 Patterson's Medium or Potato Dextrose Agar (Schafer & Kohn, 2006) for the MCG reactions, the patterns  
513 we observe are not dissimilar from what have previously been reported in the literature. It has been  
514 demonstrated in several Ascomycetes—including *Neurospora crassa* (Micali & Smith, 2003), *Sclerotinia*  
515 ~~homoeearpa~~ homoeocarpa (Jo et al., 2008), *Verticillium dahliae* (Papaioannou & Typas, 2014), and *S.*

516 *sclerotiorum* (Ford et al., 1995)—that barrage reactions are independent from stable anastomosis. Thus,  
517 the inconsistencies in this study and other studies indicate that researchers studying *S. sclerotiorum* should  
518 not rely on MCG data derived from barrage reactions as an indicator for genetic diversity.

### 519 **Limitations**

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

527 Another limitation was the microsatellite markers used for this particular study (Sirjusingh & Kohn,  
528 2001). The haplotype accumulation curve showed no indication of a plateau, indicating that if we had  
529 sampled more loci, we would have resolved more multilocus haplotypes. While 16 loci showed us  
530 similar results and began to show a plateau for the haplotype accumulation curve, we were unable to  
531 use these results due to our uncertainty in the allele calls for these five extra loci. With the availability  
532 of an optically-mapped genome (Derbyshire et al., 2017), future studies describing the genetic diversity  
533 of *S. sclerotiorum* should employ techniques such as Genotyping-By-Sequencing (Davey et al., 2011),  
534 Sequence Capture (Grover et al., 2012), or Whole Genome Sequencing.

### 535 **Conclusions**

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

### 545 **Data Availability**

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

### 549 **Funding**

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

### 554 **Acknowledgements**

555 The authors would like to thank Rebecca [Higgins-Higgins](#) for technical support in generating the data  
556 for the MCG assessment, aggressiveness ratings, and genotyping; and for providing valuable insights into  
557 the historical context of the data collection and curation.

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

### 560 **Conflicts of Interest**

561 The authors declare no conflict of interest.

### 562 **Author Contributions**

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

- 565 • **Bimal B. Sajeewa Amaradasa** analyzed the data, contributed analysis tools, wrote drafts of the  
566 paper, edited and reviewed drafts of the paper.
- 567 • **Rachana Jhala** Carried out experiments (MCG assessment, aggressiveness ratings, genotyping),  
568 edited and reviewed drafts of the paper.
- 569 • **Serena McCoy** Carried out experiments (MCG assessment, aggressiveness ratings, genotyping),  
570 edited and reviewed drafts of the paper.
- 571 • **James R. Steadman** Conceived and designed experiments, organized network of white mold  
572 screening nurseries, provided *S. sclerotiorum* isolates, edited and reviewed drafts of the paper.
- 573 • **Sydney E. Everhart** supervised data analysis, analyzed the data, contributed analysis tools, wrote  
574 the paper, edited and reviewed drafts of the paper.

## 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, Bunsu, 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, Bunsu, 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, Bunsu, G122	11
USA	ND	producer	2007, 2010	unk	53
		wmn	2005	Beryl, Bunsu, G122	7
USA	NE	wmn	2004, 2005, 2008, 2010	Beryl, Bunsu, 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, Bunsu, G122	15
		producer	2003	G122, GH	2
USA	WA	wmn	2003, 2004, 2005, 2008	11A, 37, 38, Beryl, Bunsu, Cornell, G122, Orion, PO7104, PO7863, WM31	36
		producer	2003, 2007	GH, Merlot, Pinto, Redkid	23
USA	WI	producer	2003	GH	2
Mexico	-	wmn	2005	Beryl, Bunsu, G122	18
France	-	wmn	2004, 2005	Beryl, Bunsu, G122	18
		producer	2012	unk	4
Australia	-	wmn	2004	Beryl, Bunsu, 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

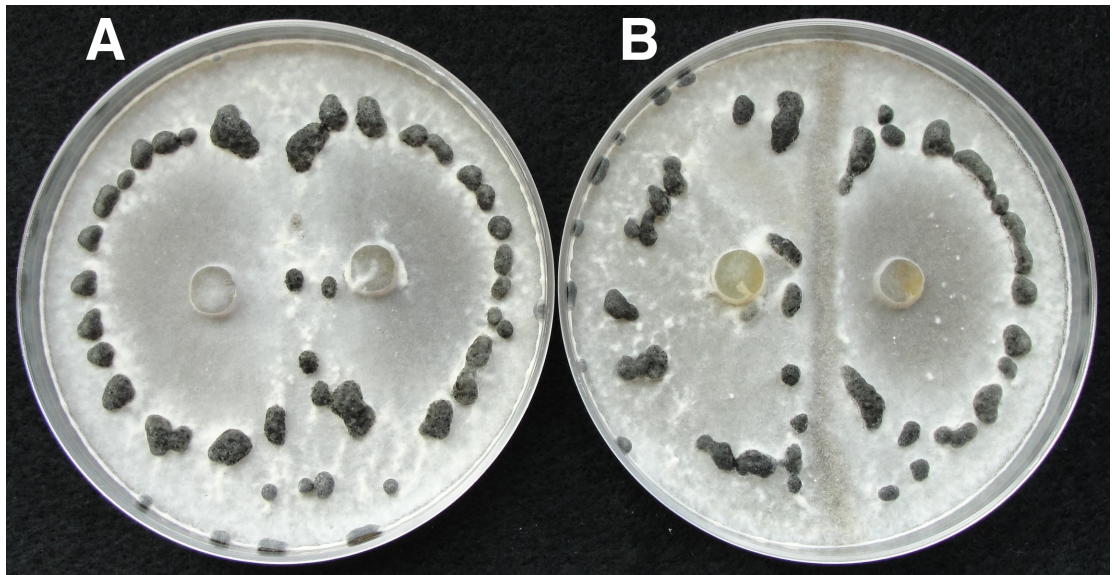


**Table S3.** Mean aggressiveness ratings for the 10 most abundant MCG; groupings according to 95% family-wise confidence interval.

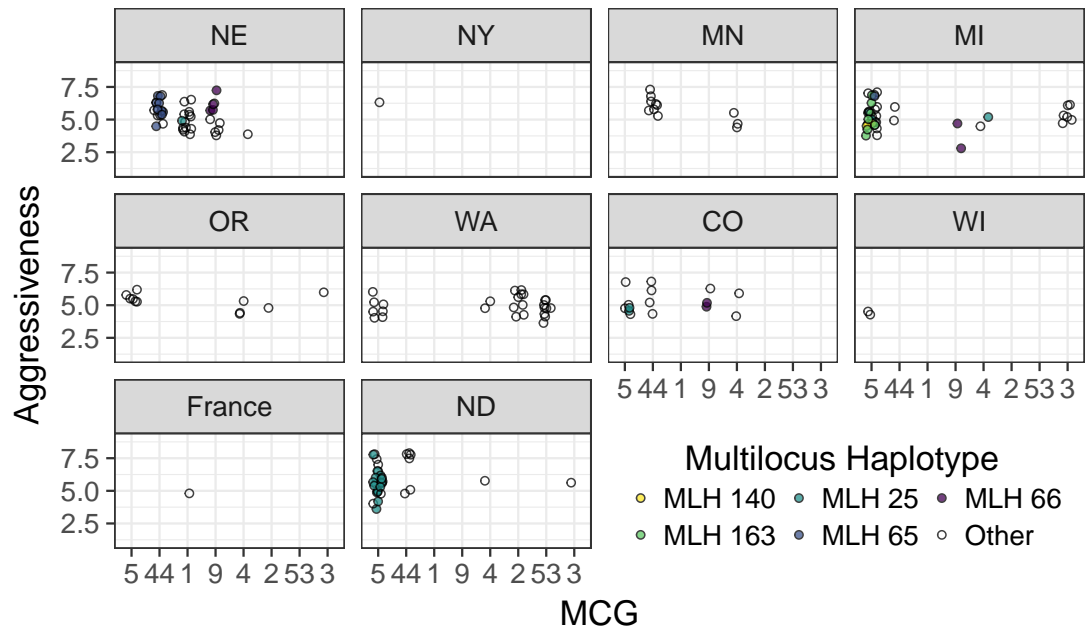
MCG	Mean Aggressiveness	Group
44	6.03	a
3	5.50	ab
5	5.40	b
2	5.25	b
9	5.11	b
1	4.95	b
45	4.88	b
4	4.87	b
53	4.69	b
49	4.60	b

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

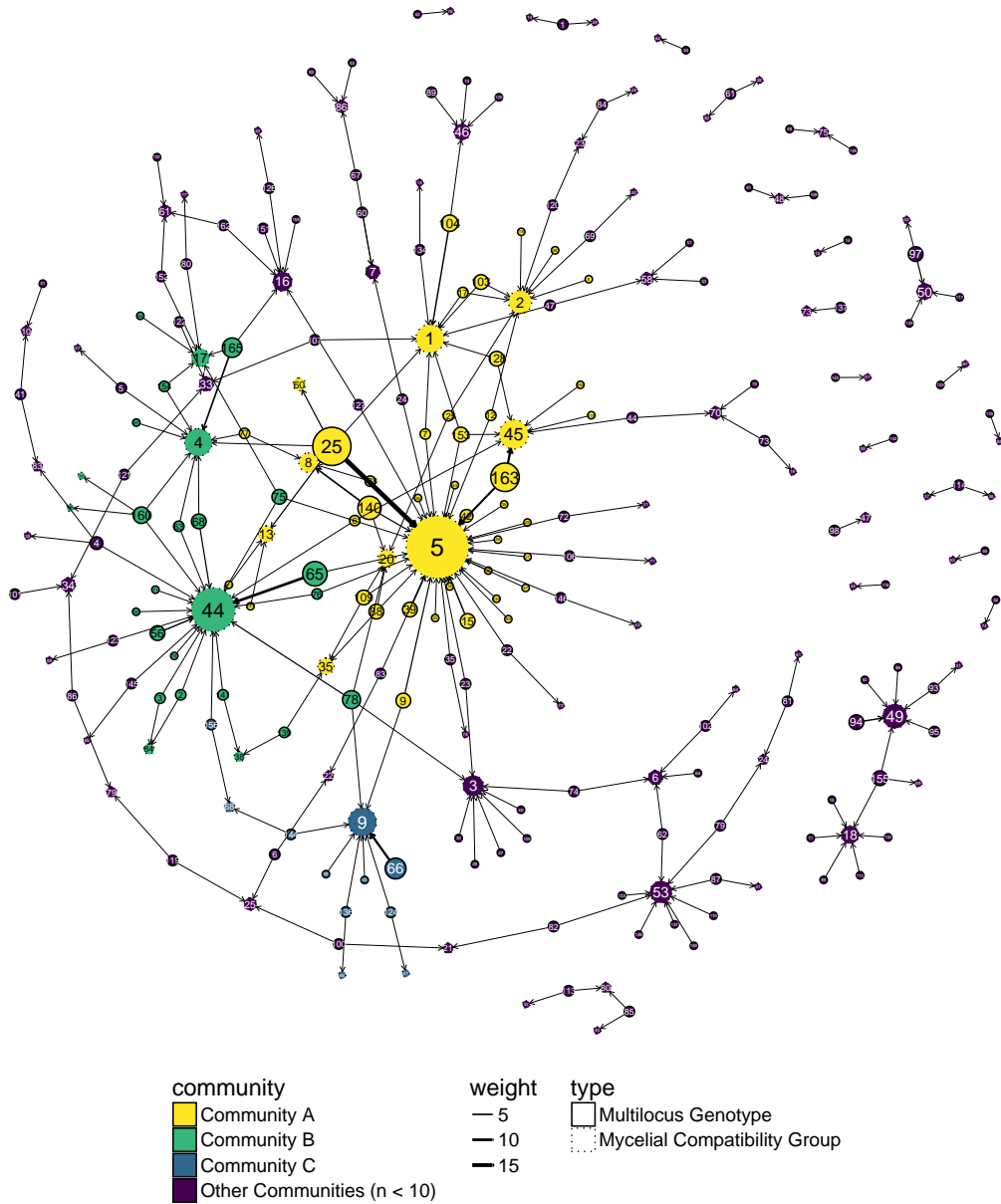
MLH	Mean Aggressiveness	Group
78	6.07	a
65	5.94	a
9	5.67	ab
25	5.41	ab
66	5.30	ab
104	5.22	ab
160	4.80	ab
163	4.80	ab
165	4.34	b
140	4.31	b



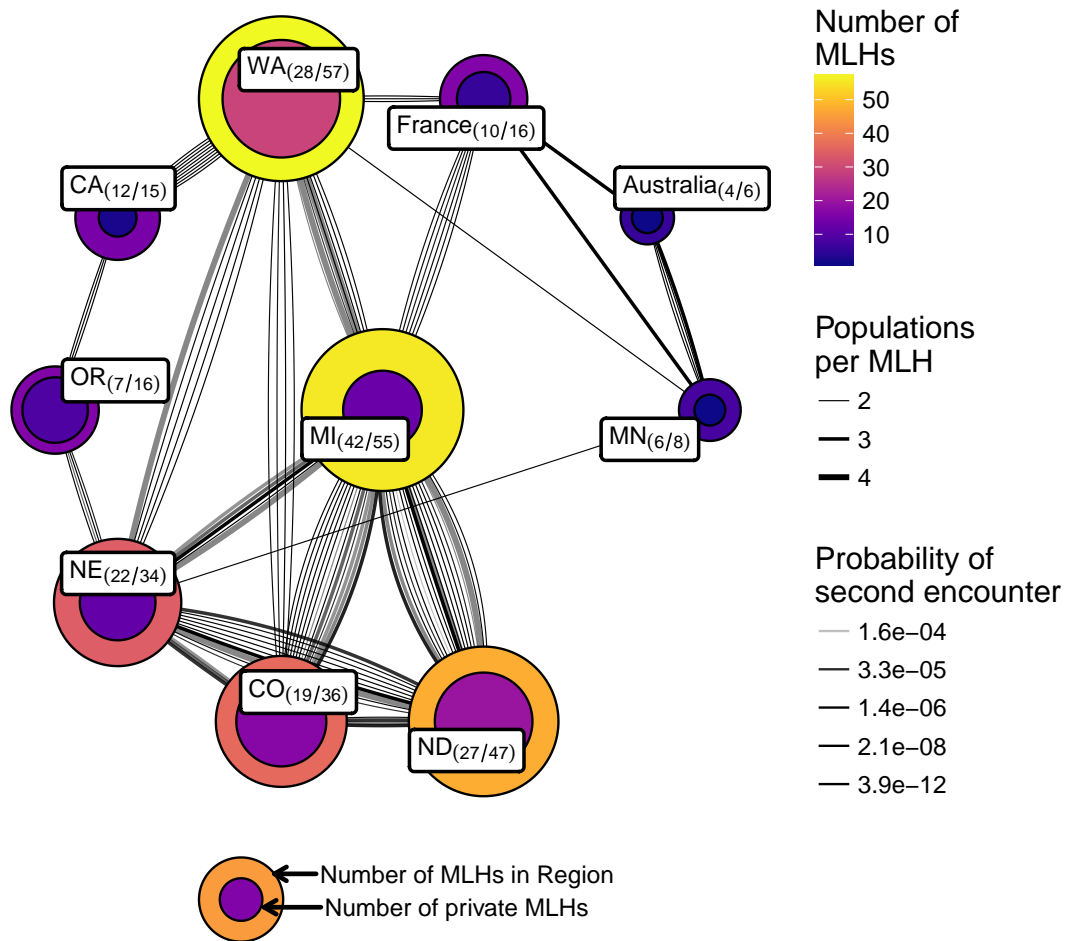
**Figure S1.** [Example of MCG test plates showing \(A\) a compatible reaction with mycelia from two strains overgrowing each other and \(B\) an incompatible reaction with a barrage line of dead tissue forming between the two strains.](#) Photo Credit: Rebecca Higgins.



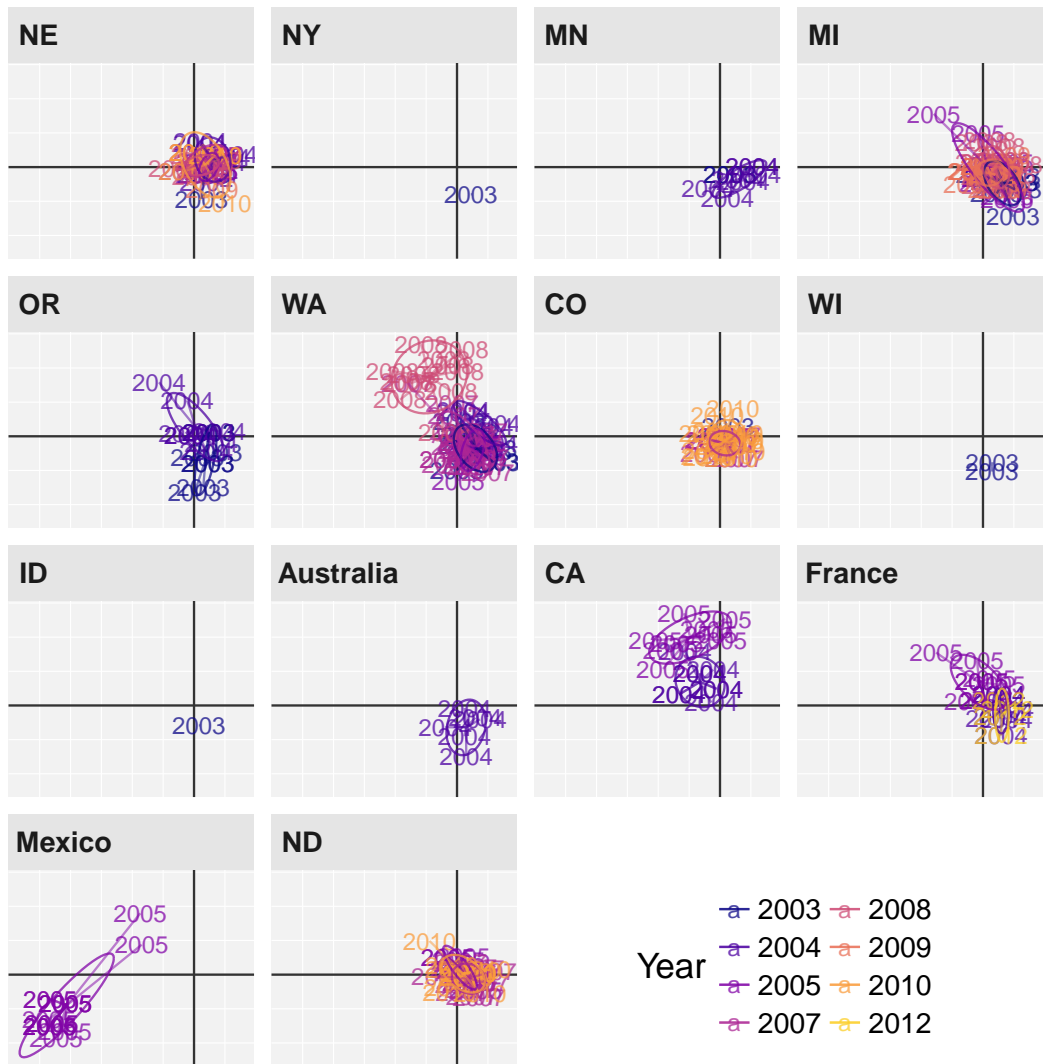
**Figure S2.** [Strip plot of aggressiveness for the eight most abundant MCGs partitioned by region. Filled circles indicate one of the five most abundant MLHs and open circles indicate a MLH of lesser abundance.](#)



**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 \(Direct Link: https://github.com/everhartlab/sclerotinia-366/blob/master/results/mlg-mcg.md#interactive-visualizations\)](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.



576 **REFERENCES**

- 577 Agapow, P., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology*  
578 *Notes, 1*, 101–102.
- 579 Aldrich-Wolfe, L., Travers, S., & Nelson, B. D. (2015). Genetic variation of *Sclerotinia sclerotiorum*  
580 from multiple crops in the north central United States. *PLOS ONE, 10*(9), e0139188. <https://doi.org/10.1371/journal.pone.0139188>
- 582 Arnaud-Hanod, S., Duarte, C. M., Alberto, F., & Serrão, E. A. (2007). Standardizing methods to  
583 address clonality in population studies. *Molecular Ecology, 16*(24), 5115–5139. <https://doi.org/10.1111/j.1365-294X.2007.03535.x>
- 585 Atallah, Z. K., Larget, B., Chen, X., & Johnson, D. A. (2004). High genetic diversity, phenotypic  
586 uniformity, and evidence of outcrossing in *Sclerotinia sclerotiorum* in the Columbia Basin of Washington  
587 State. *Phytopathology, 94*, 737–742.
- 588 Attanayake, R. N., Carter, P. A., Jiang, D., del Río-Mendoza, L., & Chen, W. (2013). *Scler-*  
589 *rotinia sclerotiorum* populations infecting canola from China and the United States are genetically  
590 and phenotypically distinct. *Phytopathology, 103*(7), 750–761. [https://doi.org/10.1094/](https://doi.org/10.1094/phyto-07-12-0159-r)  
591 [phyto-07-12-0159-r](https://doi.org/10.1094/phyto-07-12-0159-r)
- 592 Attanayake, R., Porter, L., Johnson, D., & Chen, W. (2012). Genetic and phenotypic diversity and  
593 random association of DNA markers of isolates of the fungal plant pathogen *Sclerotinia sclerotiorum*  
594 from soil on a fine geographic scale. *Soil Biology and Biochemistry, 55*, 28–36. <https://doi.org/10.1016/j.soilbio.2012.06.002>
- 596 [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](https://arxiv.org/abs/1710.03675)
- 598 Boland, G., & Hall, R. (1994). Index of plant hosts of *Sclerotinia sclerotiorum*. *Canadian Journal of*  
599 *Plant Pathology, 16*(2), 93–108. <https://doi.org/10.1080/07060669409500766>
- 600 Bolton, M. D., Thomma, B. P. H. J., & Nelson, B. D. (2006). *Sclerotinia sclerotiorum* (Lib.) de  
601 Bary: Biology and molecular traits of a cosmopolitan pathogen. *Molecular Plant Pathology, 7*(1), 1–16.  
602 <https://doi.org/10.1111/j.1364-3703.2005.00316.x>
- 603 Botelho, L. d., Zancan, W. L. A., Cruz Machado, J. da, & Barrocas, E. N. (2013). Performance of  
604 common bean seeds infected by the fungus *Sclerotinia sclerotiorum*. *Journal of Seed Science, 35*(2),  
605 153–160. <https://doi.org/10.1590/s2317-15372013000200003>
- 606 Brown, A. H. D., Feldman, M. W., & Nevo, E. (1980). Multilocus structure of natural populations of  
607 *Hordeum Spontaneum*. *Genetics, 96*(2), 523–536. Retrieved from [http://www.genetics.org/](http://www.genetics.org/content/96/2/523)  
608 [content/96/2/523](http://www.genetics.org/content/96/2/523)
- 609 Bruvo, R., Michiels, N. K., D'Souza, T. G., & Schulenburg, H. (2004). A simple method for the  
610 calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology, 13*(7),  
611 2101–2106.
- 612 Carbone, I., & Kohn, L. M. (2001). Multilocus nested haplotype networks extended with DNA  
613 fingerprints show common origin and fine-scale, ongoing genetic divergence in a wild microbial metapop-  
614 ulation. *Molecular Ecology, 10*(10), 2409–2422. [https://doi.org/10.1046/j.0962-1083.](https://doi.org/10.1046/j.0962-1083.2001.01380.x)  
615 [2001.01380.x](https://doi.org/10.1046/j.0962-1083.2001.01380.x)
- 616 Carbone, I., Anderson, J. B., & Kohn, L. M. (1999). Patterns of descent in clonal lineages and  
617 their multilocus fingerprints are resolved with combined gene genealogies. *Evolution, 53*(1), 11–21.  
618 <https://doi.org/10.1111/j.1558-5646.1999.tb05329.x>
- 619 Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research.  
620 *InterJournal, Complex Systems, 1695*. Retrieved from <http://igraph.org>
- 621 Cubeta, M. A., Cody, B. R., Kohli, Y., & Kohn, L. M. (1997). Clonality in *Sclerotinia sclerotiorum* on  
622 infected cabbage in eastern North Carolina. *Phytopathology, 87*, 1000–1004.
- 623 Cubeta, M., Sermons, D., & Cody, B. (2001). Mycelial interactions of *Sclerotinia minor*. *Phytopathol-*  
624 *ogy, 91*(6S), S19. <https://doi.org/10.1094/phyto.2001.91.6.s1>
- 625 Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011).  
626 Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature*  
627 *Reviews Genetics, 12*(7), 499–510. <https://doi.org/10.1038/nrg3012>
- 628 Derbyshire, M., Denton-Giles, M., Hegedus, D., Seifbarghy, S., Rollins, J., van Kan, J., Seidl, M.  
629 F., Faino, L., Mbengue, M., Navaud, O., Raffaele, S., Hammond-Kosack, K., Heard, S., & Oliver, R.  
630 (2017). The complete genome sequence of the phytopathogenic fungus *Sclerotinia sclerotiorum* reveals

631 insights into the genome architecture of broad host range pathogens. *Genome Biology and Evolution*,  
632 9(3), 593–618. <https://doi.org/10.1093/gbe/evx030>

633 Ekins, M. G., Hayden, H. L., Aitken, E. A. B., & Goulter, K. C. (2011). Population structure of  
634 *Sclerotinia sclerotiorum* on sunflower in Australia. *Australasian Plant Pathology*, 40, 99–108.

635 Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from  
636 metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data.  
637 *Genetics*, 131(2), 479–91.

638 Ford, E., Miller, R., Gray, H., & Sherwood, J. (1995). Heterokaryon formation and vegetative  
639 compatibility in *Sclerotinia sclerotiorum*. *Mycological Research*, 99(2), 241–247.

640 Grover, C. E., Salmon, A., & Wendel, J. F. (2012). Targeted sequence capture as a powerful tool for  
641 evolutionary analysis. *American Journal of Botany*, 99(2), 312–319. <https://doi.org/10.3732/ajb.1100323>

642 Grünwald, N. J., Goodwin, S. B., Milgroom, M. G., & Fry, W. E. (2003). Analysis of genotypic  
643 diversity data for populations of microorganisms. *Phytopathology*, 93(6), 738–746. <https://doi.org/10.1094/phyto.2003.93.6.738>

644 Hall, C., Welch, J., Kowbel, D. J., & Glass, N. L. (2010). Evolution and diversity of a fungal  
645 self/nonself recognition locus. *PLoS ONE*, 5(11), e14055. <https://doi.org/10.1371/journal.pone.0014055>

646 Heck, K. L., van Belle, G., & Simberloff, D. (1975). Explicit calculation of the rarefaction diversity  
647 measurement and the determination of sufficient sample size. *Ecology*, 56(6), 1459–1461. <https://doi.org/10.2307/1934716>

648 Henry, L., & Wickham, H. (2017). *purrr: Functional programming tools*. Retrieved from <https://CRAN.R-project.org/package=purrr>

649 Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters.  
650 *Ecology*, 52(4), 577–586. <https://doi.org/10.2307/1934145>

651 Jo, Y.-K., Chang, S. W., Rees, J., & Jung, G. (2008). Reassessment of vegetative compatibility of  
652 *Sclerotinia homoeocarpa* using nitrate-nonutilizing mutants. *Phytopathology*, 98(1), 108–114. <https://doi.org/10.1094/phyto-98-1-0108>

653 Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>

654 Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components:  
655 A new method for the analysis of genetically structured populations. *BMC Genetics*, 11:94. <https://doi.org/10.1186/1471-2156-11-94>

656 Kamvar, Z. N., Amaradasa, B. S., Jhala, R., McCoy, S., Steadman, J. R., & Everhart, S. E.  
657 (2017, November). Data and analysis for population structure and phenotypic variation of *Sclerotinia*  
658 *sclerotiorum* from dry bean (*Phaseolus vulgaris*) in the United States. Open Science Framework.  
659 Retrieved from [osf.io/k8wtm](https://osf.io/k8wtm)

660 Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide  
661 population genetic data with emphasis on clonality. *Frontiers in Genetics*, 6. <https://doi.org/10.3389/fgene.2015.00208>

662 Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of  
663 populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281. <https://doi.org/10.7717/peerj.281>

664 Knodel, J., Beauzay, P., Franzen, D., Kandel, H., Markell, S., Osorno, J., Pasche, J., & Zollinger, R.  
665 (2012). 2012 dry bean grower survey of production, pest problems and pesticide use in Minnesota and  
666 North Dakota. *North Dakota State University Extension*, E1640.

667 Knodel, J., Beauzay, P., Franzen, D., Kandel, H., Markell, S., Osorno, J., Pasche, J., & Zollinger, R.  
668 (2015). 2015 dry bean grower survey of production, pest problems and pesticide use in Minnesota and  
669 North Dakota. *North Dakota State University Extension*, E1802.

670 Knodel, J., Beauzay, P., Franzen, D., Kandel, H., Markell, S., Osorno, J., Pasche, J., & Zollinger, R.  
671 (2016). 2016 dry bean grower survey of production, pest problems and pesticide use in Minnesota and  
672 North Dakota. *North Dakota State University Extension*, E1841.

673 Kohli, Y., & Kohn, L. M. (1998). Random association among alleles in clonal populations of  
674 *Sclerotinia sclerotiorum*. *Fungal Genetics and Biology*, 23, 139–149.

675 Kohli, Y., Brunner, L. J., Yoell, H., Milgroom, M. G., Anderson, J. B., Morrall, R. A. A., & Kohn, L.

686 M. (1995). Clonal dispersal and spatial mixing in populations of the plant pathogenic fungus, *Sclerotinia*  
687 *sclerotiorum*. *Molecular Ecology*, 4, 69–77.

688 Kohn, L. M. (1995). The clonal dynamic in wild and agricultural plant-pathogen populations.  
689 *Canadian Journal of Botany*, 73(S1), 1231–1240. <https://doi.org/10.1139/b95-383>

690 Kohn, L. M., Carbone, I., & Anderson, J. B. (1990). Mycelial interactions in *Sclerotinia sclerotiorum*.  
691 *Experimental Mycology*, 14, 255–267.

692 Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: Testing multispecies  
693 responses in multifactorial ecological experiments. *Ecological Monographs*, 69, 1–24.

694 Lehner, M. S., & Mizubuti, E. S. G. (2017). Are *Sclerotinia sclerotiorum* populations from the tropics  
695 more variable than those from subtropical and temperate zones? *Tropical Plant Pathology*, 42(2), 61–69.  
696 <https://doi.org/10.1007/s40858-016-0125-1>

697 Lehner, M. S., Júnior, T. J. P., Júnior, B. T. H., Teixeira, H., Vieira, R. F., Carneiro, J. E. S., & Mizubuti,  
698 E. S. G. (2015). Low genetic variability in *Sclerotinia sclerotiorum* populations from common bean  
699 fields in Minas Gerais State, Brazil, at regional, local and micro-scales. *Plant Pathology*, 64(4), 921–931.  
700 <https://doi.org/10.1111/ppa.12322>

701 Lehner, M. S., Paula Júnior, T. J. de, Del Ponte, E. M., Mizubuti, E. S., & Pethybridge, S. J. (2017).  
702 Independently founded populations of *Sclerotinia sclerotiorum* from a tropical and a temperate region have  
703 similar genetic structure. *PloS One*, 12(3), e0173915. <https://doi.org/10.1371/journal.pone.0173915>

704

705 Leslie, J. (1993). Fungal vegetative compatibility. *Annual Review of Phytopathology*, 31, 127–150.  
706 Review. <https://doi.org/10.1146/annurev.py.31.090193.001015>

707 McCoy, S., & Steadman, J. R. (2009). Use of multi-site screening to identify partial resistance to  
708 white mold in common bean in 2008. *Bean Improvement Cooperative Annual Report*, 86–87. Retrieved  
709 from <https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND44207142>

710 McDonald, B. A., & Linde, C. (2002). Pathogen population genetics, evolutionary potential, and  
711 durable resistance. *Annual Review of Phytopathology*, 40(1), 349–379. <https://doi.org/10.1146/annurev.phyto.40.120501.101443>

712

713 Mendiburu, F. D., & Simon, R. (2015). Agricolae - ten years of an open source statistical tool for exper-  
714 iments in breeding, agriculture and biology. <https://doi.org/10.7287/peerj.preprints.1404v1>

715

716 Micali, C. O., & Smith, M. L. (2003). On the independence of barrage formation and heterokaryon  
717 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)

718

719 Milgroom, M. G. (1996). Recombination and the multilocus structure of fungal populations. *Annual*  
720 *Review of Phytopathology*, 34(1), 457–477.

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

723 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R.,  
724 O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2017). *Ve-*  
725 *gan: Community ecology package*. Retrieved from [https://CRAN.R-project.org/package=](https://CRAN.R-project.org/package=vegan)  
726 *vegan*

727 Otto-Hanson, L., & Steadman, J. R. (2007). Identification of partial resistance to *Sclerotinia scle-*  
728 *rotiorum* in common bean at multiple locations in 2006. *Bean Improvement Cooperative Annual Re-*  
729 *port*, 133–134. Retrieved from [https://naldc.nal.usda.gov/naldc/catalog.xhtml?](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43940892)  
730 *id=IND43940892*

731 Otto-Hanson, L., & Steadman, J. R. (2008). Identification of partial resistance to *Sclerotinia scle-*  
732 *rotiorum* in common bean at multiple locations in 2007. *Bean Improvement Cooperative Annual Re-*  
733 *port*, 214–215. Retrieved from [https://naldc.nal.usda.gov/naldc/catalog.xhtml?](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND44063230)  
734 *id=IND44063230*

735 Otto-Hanson, L., Steadman, J. R., Higgins, R., & Eskridge, K. M. (2011). Variation in *Sclerotinia*  
736 *sclerotiorum* bean isolates from multisite resistance screening locations. *Plant Disease*, 95(11), 1370–  
737 1377. <https://doi.org/10.1094/pdis-11-10-0865>

738 Papaioannou, I. A., & Typas, M. A. (2014). Barrage formation is independent from heterokaryon  
739 incompatibility in *Verticillium dahliae*. *European Journal of Plant Pathology*, 141(1), 71–82. <https://doi.org/10.1007/s11360-013-0865-1>

740 //doi.org/10.1007/s10658-014-0525-3

741 Parks, J. C., & Werth, C. R. (1993). A study of spatial features of clones in a population of bracken  
 742 fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, 80(5), 537. <https://doi.org/10.2307/2445369>

743

744 Pedersen, T. L. (2017). *ggraph: An implementation of grammar of graphics for graphs and networks*.  
 745 Retrieved from <https://CRAN.R-project.org/package=ggraph>

746 Petzoldt, R., & Dickson, M. H. (1996). Straw test for resistance to white mold in beans. *Bean*  
 747 *Improvement Cooperative Annual Report*, 39, 142–143. Retrieved from [https://naldc.nal.usda.](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND20562675)  
 748 [gov/naldc/catalog.xhtml?id=IND20562675](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND20562675)

749 Pielou, E. (1975). *Ecological Diversity*. New York: Wiley & Sons.

750 Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks.  
 751 *Journal of Graph Algorithms and Applications*, 10(2), 191–218. [https://doi.org/10.7155/](https://doi.org/10.7155/jgaa.00124)  
 752 [jgaa.00124](https://doi.org/10.7155/jgaa.00124)

753 Prugnolle, F., & de Meeûs, T. (2010). Apparent high recombination rates in clonal parasitic organisms  
 754 due to inappropriate sampling design. *Heredity*, 104(2), 135–140. [https://doi.org/10.1038/](https://doi.org/10.1038/hdy.2009.128)  
 755 [hdy.2009.128](https://doi.org/10.1038/hdy.2009.128)

756 R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R  
 757 Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

758 [Ramasubramaniam, H., del Río Mendoza, L. E., & Bradley, C. A. \(2008\). Estimates of yield and](https://doi.org/10.2134/agronj2007.0127)  
 759 [economic losses associated with white mold of rain-fed dry bean in North Dakota. \*Agronomy Journal\*,](https://doi.org/10.2134/agronj2007.0127)  
 760 [100\(2\), 315. https://doi.org/10.2134/agronj2007.0127](https://doi.org/10.2134/agronj2007.0127)

761 Sambrook, J., Fritsch, E. F., Maniatis, T., & others. (1989). *Molecular Cloning: A Laboratory Manual*.  
 762 Cold spring harbor laboratory press.

763 Saupe, S. J. (2000). Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes.  
 764 *Microbiology and Molecular Biology Reviews*, 64(3), 489–502. [https://doi.org/10.1128/](https://doi.org/10.1128/mmbr.64.3.489-502.2000)  
 765 [mmbr.64.3.489-502.2000](https://doi.org/10.1128/mmbr.64.3.489-502.2000)

766 Schafer, M. R., & Kohn, L. M. (2006). An optimized method for mycelial compatibility testing in  
 767 *Sclerotinia sclerotiorum*. *Mycologia*, 98(4), 593–597. [https://doi.org/10.1080/15572536.](https://doi.org/10.1080/15572536.2006.11832662)  
 768 [2006.11832662](https://doi.org/10.1080/15572536.2006.11832662)

769 Sexton, A. C., & Howlett, B. J. (2004). Microsatellite markers reveal genetic differentiation among  
 770 populations of *Sclerotinia sclerotiorum* from Australian canola fields. *Current Genetics*, 46(6), 357–365.  
 771 <https://doi.org/10.1007/s00294-004-0543-3>

772 Sexton, A. C., Whitten, A. R., & Howlett, B. J. (2006). Population structure of *Sclerotinia sclerotiorum*  
 773 in an Australian canola field at flowering and stem-infection stages of the disease cycle. *Genome*, 49(11),  
 774 1408–1415. <https://doi.org/10.1139/g06-101>

775 Shannon, C. E. (1948). A mathematical theory of communication. *ACM SIGMOBILE Mobile*  
 776 *Computing and Communications Review*, 5(1), 3–55.

777 Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163(4148), 688–688. [https://doi.](https://doi.org/10.1038/163688a0)  
 778 [org/10.1038/163688a0](https://doi.org/10.1038/163688a0)

779 Sirjusingh, C., & Kohn, L. M. (2001). Characterisation of microsatellites in the fungal plant pathogen,  
 780 *Sclerotinia sclerotiorum*. *Molecular Ecology Notes*, 1(4), 267–269. [https://doi.org/10.1046/](https://doi.org/10.1046/j.1471-8278.2001.00102.x)  
 781 [j.1471-8278.2001.00102.x](https://doi.org/10.1046/j.1471-8278.2001.00102.x)

782 Smith, J. M., Smith, N. H., O'Rourke, M., & Spratt, B. G. (1993). How clonal are bacteria?  
 783 *Proceedings of the National Academy of Sciences*, 90(10), 4384–4388. [https://doi.org/10.](https://doi.org/10.1073/pnas.90.10.4384)  
 784 [1073/pnas.90.10.4384](https://doi.org/10.1073/pnas.90.10.4384)

785 Steadman, J. R. (1983). White mold - a serious yield-limiting disease of bean. *Plant Disease*, 67,  
 786 346–350.

787 Steadman, J. R., Eskridge, K., & Powers, K. (2003). Identification of partial resistance to *Scle-*  
 788 *rotinia sclerotiorum* in common bean at multiple locations. *Bean Improvement Cooperative Annual*  
 789 *Report*, 225–226. Retrieved from [https://naldc.nal.usda.gov/naldc/catalog.xhtml?](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43757287)  
 790 [id=IND43757287](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43757287)

791 Steadman, J. R., Otto-Hanson, L., & Breathnach, J. (2006). Identification of partial resistance to  
 792 *Sclerotinia sclerotiorum* in common bean at multiple locations in 2005. *Bean Improvement Cooperative*  
 793 *Annual Report*, 223–224. Retrieved from <https://naldc.nal.usda.gov/naldc/catalog>.

794 xhtml?id=IND43805570

795 Steadman, J. R., Otto-Hanson, L., & Powers, K. (2004). Identification of partial resistance to *Sclero-*  
796 *tinia sclerotiorum* in common bean at multiple locations. *Bean Improvement Cooperative Annual Report,*  
797 *47*, 281–282. Retrieved from [https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43758354)  
798 [IND43758354](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43758354)

799 Steadman, J. R., Otto-Hanson, L., & Powers, K. (2005). Identification of partial resistance to  
800 *Sclerotinia sclerotiorum* in common bean at multiple locations in 2004. *Bean Improvement Cooperative*  
801 *Annual Report*, 124–125. Retrieved from [https://naldc.nal.usda.gov/naldc/catalog.](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43759243)  
802 [xhtml?id=IND43759243](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43759243)

803 Stoddart, J. A., & Taylor, J. F. (1988). Genotypic diversity: Estimation and prediction in samples.  
804 *Genetics*, *118*(4), 705–11.

805 Strausbaugh, C., & Forster, R. (2003). Management of white mold of beans. *Pacific North-*  
806 *west Extension*, *PNW568*. Retrieved from [http://http://www.extension.uidaho.edu/](http://http://www.extension.uidaho.edu/publishing/pdf/PNW/PNW0568.pdf)  
807 [publishing/pdf/PNW/PNW0568.pdf](http://http://www.extension.uidaho.edu/publishing/pdf/PNW/PNW0568.pdf)

808 Strom, N. B., & Bushley, K. E. (2016). Two genomes are better than one: History, genetics,  
809 and biotechnological applications of fungal heterokaryons. *Fungal Biology and Biotechnology*, *3*(1).  
810 <https://doi.org/10.1186/s40694-016-0022-x>

811 Teran, H., Lema, M., Schwartz, H. F., Duncan, R., Gilbeitson, R., & Singh, S. P. (2006). Modified  
812 Petzoldt and Dickson scale for white mold rating of common bean. *Bean Improvement Cooperative An-*  
813 *nual Report*, *49*, 115–116. Retrieved from [https://naldc.nal.usda.gov/naldc/catalog.](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43805401)  
814 [xhtml?id=IND43805401](https://naldc.nal.usda.gov/naldc/catalog.xhtml?id=IND43805401)

815 [Tu, J. C., & Beversdorf, W. D. \(1982\). Tolerance to white mold \(\*Sclerotinia sclerotiorum\* \(Lib.\)](https://doi.org/10.4141/cjps82-010)  
816 [De Bary\) in Ex Rico 23, a cultivar of white bean \(\*Phaseolus vulgaris\* L.\). \*Canadian Journal of Plant\*  
817 \[Science\]\(https://doi.org/10.4141/cjps82-010\), \*62\*\(1\), 65–69. <https://doi.org/10.4141/cjps82-010>](https://doi.org/10.4141/cjps82-010)

818 Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.  
819 Retrieved from <http://ggplot2.org>

820 Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). *dplyr: A grammar of data manipulation*.  
821 Retrieved from <https://CRAN.R-project.org/package=dplyr>