Population Biology

*e-X*tra*

Multiple Introductions Without Admixture of *Colletotrichum truncatum* Associated with Soybean Anthracnose in Brazil

Flávia Rogério,[1,†] Pierre Gladieux,[2] Nelson Sidnei Massola Jr.,[1] and Maisa Ciampi-Guillardi[1]

[1]Department of Plant Pathology and Nematology, University of São Paulo, Piracicaba, SP, Brazil; and [2]UMR BGPI, Univ Montpellier, INRA, CIRAD, Montpellier SupAgro, Montpellier, France
Accepted for publication 15 November 2018.

ABSTRACT

Knowledge of the population structure, genetic diversity, and reproductive mode of plant pathogens can help to implement effective disease management strategies. Anthracnose is one of the most prominent diseases in soybean and is mainly associated with the species *Colletotrichum truncatum*. However, the genetic structure of *C. truncatum* populations associated with soybean remains unknown. We collected *C. truncatum* isolates from 10 sites representing two Brazilian states (Mato Grosso and Goiás) and used 13 highly polymorphic microsatellite markers to investigate the population genetic structure of the pathogen. Analyses revealed high gene and haplotypic diversity within populations, as well low genetic differentiation and sharing of multilocus haplotypes among populations and regions. Bayesian and multivariate analysis revealed the presence of three distinct genetic clusters with at least two coexisting in all locations, and all of them coexisting in eight locations. We found limited evidence for admixture between clusters, with only two isolates showing nonzero membership with a second cluster. Analyses of linkage disequilibrium rejected the hypothesis of random mating in all clusters, but values of the index of association were low and not consistent with long-term lack of sexual reproduction. Our findings suggest that Brazilian *C. truncatum* populations resulted from at least three founder events that led to three genetic clusters that spread throughout the country, raising questions with respect to the factors allowing their maintenance in syntopy without evidence of admixture between them.

Soybean is one of the main agricultural commodities with high global economic relevance. The extensive area of soybean sown in Brazil, including the incorporation of new growing areas in the northern and central-west regions, characterized by genetically uniform crops and nontillage system, affect the occurrence and the intensity of diseases. Among these, one of the most prominent is anthracnose, as soybeans are susceptible to infection at all stages of development. Soybean anthracnose is a complex system, with many unresolved questions about the identity of the causal agent, the fungicide management and the need to develop a specific integrated control program (Dias et al. 2016). In Brazil, little information on control of anthracnose is available. Increasing of disease reports in northern and central-west regions indicate that chemical control program for fungal diseases in soybean has not been effective against anthracnose. The disease can be an important limiting factor for soybean production in areas with favorable weather conditions, such as high humidity and temperature, where reports of significant yield reductions are frequent (Dias et al. 2016).

Soybean (*Glycine max* (Linnaeus) Merrill) originates from East Asia, most likely from China, and would have spread to the rest of the world from this area (Hymowitz 1970). Initially used as forage plant, it was introduced into North America in 1765, achieving a prominent role as a grain crop in the 1920s (Hymowitz 1990). In Brazil, the crop was introduced in 1882 in Bahia state and began to be widely cultivated in Rio Grande do Sul state. Production was boosted in the southern states with the beginning of wheat/soybean succession, and it was expanded to the Central region (Brazilian savanna) in the 1960s and 1970s (EMBRAPA 1996). The soybean cultivars sown in those decades were the outcome of breeding programs for the development of better-adapted cultivars to Brazilian conditions, based on cultivars introduced from the South of the United States (Arantes and Miranda 1993; Hirimoto and Vello 1986; Wysmierski and Vello 2013).

Anthracnose is mainly associated with the species *Colletotrichum truncatum* (Hyde et al. 2009), which can also infect important plant species belonging to Fabaceae and Solanaceae families (Cannon et al. 2012; Damm et al. 2009; Weidemann et al. 1988). The pathogen might be a native of Asia since the first report of a *Colletotrichum* species causing soybean anthracnose came from Korea (Tiffany 1950). However, a report of soybean anthracnose in Korea in 1917 was attributed to *C. glycines* (Hemmi 1920 cited in Damm et al. 2019). Lehman and Wolf (1926) studied soybean anthracnose in North Carolina (United States) and noticed an asexual morph associated with the disease symptoms, which was identified as *C. glycines*, based on similarity with the pathogen reported in Korea in 1917. Apparently, this strain is identical to strain CBS 195.32, isolated by Lehman and deposited in the CBS collection in 1932, which was later identified as *C. truncatum* by Damm et al. (2009) based on the previous classification done by Andrus and Moore (1935). In Brazil, soybean anthracnose was first observed in Rio Grande do Sul in 1961 and has been reported more frequently in the Brazilian savanna currently (Araújo et al. 1988).

Assuming that the center of origin of pathogenic species should correspond to the center of genetic diversity of the host species (Stukenbrock and McDonald 2008), *C. truncatum* is hypothesized to have originated in China and established in Brazil as an invasive species. Pathogens can be introduced into new areas by several sources, for instance through the spread of agriculture or globalization of travel and trade, allowing extensive movements of crop species and plant products, which could lead to unintentional transport of fungal pathogens far from their native location (Desprez-Loustau et al. 2007; Gladieux et al. 2008;

Yarwood 1970). Fungal invaders may adapt to the new environment after their introduction, and rapid and drastic evolutionary changes may be facilitated by huge population sizes and mixed mating systems, causing yield losses in agrosystems (Gladieux et al. 2015).

Knowledge of the population structure, genetic diversity and sexual recombination of plant pathogens can support definition of effective disease management strategies, for instance by keeping pathogen population sizes small in order to limit gene diversity (McDonald and Linde 2002). Quantifying the amount of gene diversity in pathogen populations can indicate founder populations, which is predicted to have the highest gene diversity levels (Hallatschek and Nelson 2008). Furthermore, population genetics is a useful approach to infer introduction routes of invasive pathogens and dispersal of fungal propagules (Baroncelli et al. 2015; Ciampi-Guillardi et al. 2014; Leo et al. 2015). The inference of historical patterns of pathogen migration among and within regions can be useful for management and attenuation of negative effects caused by invasive pathogens, for understanding current epidemics, for developing predictive models, and for characterizing the genetic basis of pathogen adaptation (Ali et al. 2014; Grünwald et al. 2012; McDonald and Linde 2002; Milgroom and Fry 1997).

Although asexual reproduction predominates in the majority of plant-pathogenic fungi, many species undergo regular sexual cycles (Milgroom 1996). The genus *Colletotrichum* has been predominantly observed in the asexual state. As reported for other *Colletotrichum* species, the identity of the sexual stage is still unclear for *C. truncatum* (Damm et al. 2009; Hyde et al. 2009). Reproduction and mating systems affect how gene diversity is distributed within and among individuals in a population (McDonald and Linde 2002). If sexual recombination occurs sporadically, and even in a low proportion it still could have an important impact on the population genetic structure by increasing the genotypic diversity (Milgroom 1996). High levels of genetic variability in *C. truncatum* from soybean and other hosts were observed in previous studies using various molecular markers (Ford et al. 2004; Katoch et al. 2017; Ranathunge et al. 2009; Rogério et al. 2016; Sant'anna et al. 2010; Sharma 2009; Vasconcelos et al. 1994). Evidences of recombination coupled with high levels of genetic diversity were shown in

C. truncatum populations infecting chili from India and China (Diao et al. 2015; Sharma et al. 2014).

There is no information regarding the population structure of *C. truncatum* associated with soybean anthracnose in Brazil. Previous studies, mainly conducted in Asian countries, focused on genetic differences among isolates obtained from hosts such as chili and pepper, using simple sequence repeat (SSR) or other markers (inter-simple sequence repeats, random amplification of polymorphic DNA [RAPD], and nuclear genes), which revealed high genetic diversity and possible recombination (Diao et al. 2015; Mahmodi et al. 2014; Ranathunge et al. 2009; Sharma et al. 2014). In Brazil, pioneer works in this pathosystem using RAPD markers also evidenced relevant levels of polymorphism in soybean isolates (Sant'anna et al. 2010; Vasconcelos et al. 1994).

A wide distribution of predominant *C. truncatum* haplotypes associated with soybean anthracnose was revealed by a previous study, suggesting the existence of a highly efficient mechanism of pathogen dispersal over long distances in Brazil (Rogério et al. 2016). Low differentiation among *C. truncatum* populations infecting *Capsicum* was found in India, which was attributed to extensive movement of infected material among sampled regions, indicating that neither topography nor distance acts a barrier in differentiating them (Katoch et al. 2017).

The objective of this study was to investigate the population structure of *C. truncatum* in the most important soybean production areas in Brazil using microsatellite markers. Based on the assumption of efficient mechanisms of propagules dissemination of *C. truncatum* over long distances and the levels of diversity and recombination previously reported for this species, we hypothesized that the pathogen should exhibit nondifferentiated populations across soybean fields and that pathogen populations should have a recombining population structure.

MATERIALS AND METHODS

Population sampling. We collected symptomatic plant samples in 10 naturally infected commercial soybean fields in Mato Grosso (MT) and Goiás (GO) states (Fig. 1). These regions are

Fig. 1. Sampling location of *Colletotrichum truncatum* populations collected from 10 soybean fields in Mato Grosso and Goiás States, Brazil.

located in the Brazilian savanna and represent the most important soybean production areas in Brazil. All samples were collected in 2017 but MT2 sample was collected in 2016. In Goiás, samples were collected in Montevidiu and Rio Verde counties (49 km apart). In Mato Grosso, samples were collected in Lucas do Rio Verde and Sinop counties (148 km apart). Stems, leaves, and pods showing typical symptoms of anthracnose were collected from each field. To minimize the chance of sampling recently propagated clones of the pathogen, it was maintained a minimal distance of 10 m between the collected plants.

Isolation of *C. truncatum*. Symptomatic stems, leaves, and pods were surface disinfected with 70% ethanol for 30 s, and then 1% sodium hypochlorite for 1 min, and rinsed in sterile distilled water. Small pieces of leaves and stems were plated onto potato dextrose agar (PDA) medium and incubated at 25°C for 3 days. Pods were kept in Petri dishes onto moistened, sterilized filter paper and checked every day under the microscope to monitor the production of acervuli and conidia. Hyphal tips and acervuli emerging from plated tissues or pods were transferred to PDA medium and cultivated at 25°C. All fungal isolates were single-spored prior to use. Monosporic cultures were preserved on filter paper (Alfenas and Mafia 2007) and in sterile distilled water (Castellani 1939), and kept at 4°C in the Plant Pathology Department of 'Luiz de Queiroz' College of Agriculture, University of São Paulo, Brazil. Each group of fungal isolates collected from the same field was considered as part of the same population.

DNA extraction and SSR genotyping. Fungal isolates were grown for 7 days on PDA at 25°C. Microscopic observations were carried out to certify that the morphological features of the isolates corresponded to *C. truncatum*. Mycelium plugs were collected from the media, ground in sterilized sand, and fungal genomic DNA was extracted using Wizard Genomic DNA Purification Kit (Promega) according to the manufacturer's recommendations. SSR genotyping was carried out using a set of 13 loci (SSR1, SSR5, SSR10, SSR17, SSR23, SSR29, SSR34, SSR42, SSR44, SSR53, SSR55, SSR56, and SSR59) developed by Ranathunge et al. (2009). Loci were individually amplified by PCR as previously described by the authors. Forward primers were labeled with the fluorescent dyes 6-FAM, VIC, PET, and NED. Amplifications were performed in a 15 µl volume containing 25 ng of template DNA, 0.8 mM of fluorescently labeled forward and common reverse primers, 20 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl, 0.15 mM of each dNTP, and 1 U of Taq DNA polymerase in a Techne thermal cycler (TC-512; Techne Inc., Burlington, NJ) using the following program: 95°C for 2 min, then 30 cycles of 95°C for 1 min, 60°C for 30 s, 72°C for 1 min, and a final extension step of 72°C for 10 min. PCR products were visualized on 2.0% agarose gel and amplicon sizes were initially estimated against known molecular size standards (1 kb plus ladder, Invitrogen). Amplified microsatellite fragments were electrophoretically separated on an automated DNA sequencer ABI 3500 (Applied Biosystems, Foster City, CA) using GS-600 Liz (Applied Biosystems) as an internal size standard. Raw data were collected, scored, and exported as a fragment size for simple sequence repeats (SSRs) using Genemarker v.1.191 (SoftGenetics). To avoid any error bias, the genotype of each individual was confirmed by two independent repeats, revealing that the amplification patterns were highly reproducible. Statistical binning of the alleles into fragment size categories consistent with the repeat unit increments was performed with FlexiBin software (B. Amos, Cambridge University, U.K.).

Genetic and haplotypic diversity. The GenAlex v.6.5 software (Peakall and Smouse 2006) was used to identify multilocus microsatellite haplotypes (MLHs). Only one allele was amplified per locus, as expected for a haploid organism. GenAlEx was also used to generate a clone-corrected data set by detecting isolates with identical MLHs within each population and retaining only one representative sample per MLH in the data set. Unbiased gene diversity (Nei 1978) was calculated for each population using

the *poppr* package for R (Kamvar et al. 2014). Allelic richness (A_R) was estimated using a rarefaction method (El Mousadik and Petit 1996), which enables comparisons among populations of unequal sample sizes by calculating the expected number of alleles in a subsample of *n* genes corresponding to the sample size of the smallest population, implemented in ADZE (Szpiech et al. 2008). We estimated haplotypic diversity as the number of distinct haplotypes/ sample size in each fungal population. Clonal fraction was estimated per population as 1 − haplotypic diversity.

Population differentiation. The distribution of genetic variation within and among *C. truncatum* populations was estimated by analysis of molecular variance (AMOVA) (Excoffier et al. 1992) based on weighted average *F*-statistic over loci (Weir and Cockerham 1984) and pairwise comparisons of F_{ST}. Analyses were performed in Arlequin v.3.5 software (Excoffier and Lischer 2010), with significance testing by using 1,000 permutations. Genetic differentiation between population pairs was considered significant when *P* value was less than the 5% nominal level.

We used three methods to examine population subdivision. First, we used the model-based Bayesian clustering approach implemented in the STRUCTURE v.2.3 software (Pritchard et al. 2000) to access the number of genetic clusters and estimate admixture. The software uses multilocus haplotype data to define a set of clusters that maximizes linkage equilibrium and to probabilistically assign haplotypes to clusters, without considering information on their geographic origin (Falush et al. 2003). STRUCTURE was run using the admixture ancestry model-based clustering method, allowing mixed ancestry among individuals from different populations (*K*) with correlated allele frequencies among populations. Ten independent runs were performed using 500,000 Markov chain steps after a burn-in period of 100,000 steps, assuming 1 to 10 subpopulations. Second, to confirm the pattern of population subdivision inferred using STRUCTURE, we used an alternative, non-model-based, clustering method: the discriminant analysis of principal components (DAPC). This multivariate method is based on a discriminant analysis on data transformed after a principal component analysis, which seeks to maximize the intergroup component of variation (Jombart et al. 2010). DAPC was carried out using the Adegenet package 1.3-1 (Jombart and Ahmed 2011) in R environment. We retained the first 30 principal components. Ten separate runs of *K*-means were performed and the mean Bayesian information criterion (BIC) value at each value of *K* from 1 to 10 was plotted (Jombart et al. 2010). Third, to visualize ancestry relationships among multilocus haplotypes and populations, while taking into account the possibility of recombination, we built a phylogenetic network based on a matrix of Euclidean distances obtained using the neighbor-net method in the SplitsTree software (http://www. splitstree.org/). Splitstree allows visualizing conflicting phylogenetic signals, indicated by the presence of reticulations in the network and caused by recombination or incomplete lineage sorting.

Reproductive mode. To assess the predominant reproduction strategy of the pathogen, the nonrandom association of alleles between pairs of loci was measured by using the index of association (I_A), a multilocus linkage disequilibrium estimate based on the variance of pairwise distances between individuals, using clone-corrected dataset in the R package *poppr*. Because the magnitude of I_A depends on the number of loci studied, we also estimated r_D, a measure that is independent of the number of loci (Agapow and Burt 2001). An r_D value close to zero indicates random mating, while significant gametic disequilibrium, and thus higher r_D values, is expected in asexual or inbreeding populations. To examine whether the observed values deviated significantly from the null hypothesis of linkage equilibrium among loci, the observed variance was compared with the expected variance under random mating, which was obtained after reshuffling the alleles within each clone-corrected population data set over 1,000 permutations.

Summary of microsatellite variability. Ten populations were collected from individual fields, and a total of 237 *C. truncatum* fungal strains were isolated across the sampled soybean fields (Table 1). The sample size ranged from 3 to 53 isolates per population. All SSR markers were polymorphic and the dataset showed only 0.9% of missing data. A total of 108 alleles were found across populations and the number of alleles per locus ranged from 4 (SSR56 locus) to 13 (SSR59 locus), with an average of 8.3 alleles per locus. The haplotype accumulation curve tended to plateau as the number of loci increased, confirming the importance of using all 13 loci for discriminating MLHs (Fig. 2). A total of 129 unique MLHs were detected within the 237 isolates, representing an overall clonal fraction of 46% (Table 2). For each population, the number of distinct MLHs ranged from 3 to 35 and was highly dependent on population sample size ($R^2 = 0.95$) (Supplementary Fig. S1). Population GO3, due to its small sample size ($N = 3$), was excluded from genetic diversity analyses. Averaged estimates of allelic richness (A_R) and Nei's gene diversity (H) were 8.24 and 0.63, respectively. The GO1 population displayed the highest allelic richness and gene diversity values (4.4 and 0.76, respectively), while the MT3 population showed the smallest ones (2.48 and 0.21) (Table 2).

Population differentiation. Hierarchical AMOVA was performed using a clone-corrected data set in which a single representative of each MLH was kept in each population. When defining samples from different regions as two distinct groups, related to their state of origin, low to moderate differentiation was found ($F_{ST} = 0.08$; $P = 0.05$), with 91.51% of genetic diversity distributed within populations and only 3.69% between states. The populations from Mato Grosso were significantly differentiated from other populations in 18 out of 35 pairwise comparisons. The highest differentiation was between MT1 and MT3 ($F_{ST} = 0.334$, $P = 0.0$) and the smallest significant value was between MT5 and GO4 ($F_{ST} = 0.07$, $P = 0.018$, Supplementary Table S1). Population MT3 was significantly differentiated from all populations; F_{ST} estimates between this population and others tended to be the highest ones. The populations sampled in Goiás were not significantly differentiated from each other, and they showed significant differentiation only in comparisons with populations from Mato Grosso.

Patterns of shared MLHs between populations were consistent with estimates of population differentiation. Of 129 MLHs detected, 25 were shared within the same population, within the same state and even between different states. Results showed five cases of haplotypes sharing among the sampled states, always involving some GO populations (sampled in Goiás in 2017) and the MT2 population (sampled in Mato Grosso in 2016) (Supplementary Table S2). The MT3 population shared haplotypes only within other populations from Mato Grosso state.

We used the clustering method implemented in STRUCTURE, assuming a model with admixture and correlated allele frequencies, to investigate the number of genetic clusters represented in our dataset. Patterns of clustering and likelihood values were highly similar across repeats at each K value, indicating convergence of the Markov chain Monte Carlo method toward single clustering solutions. The model with $K = 3$ had both the highest likelihood [$Pr(K)$] and Evanno's ΔK, suggesting this model captures the most salient features of population subdivision in this system (Fig. 3B). At $K = 2$ and $K = 3$, differentiation between clusters was clear, with only limited admixture among groups (only two haplotypes with maximum membership proportion in a single cluster below 98%, Figure 3A). The three clusters coexisted in eight locations; in the remaining two locations, only two clusters were found. $K > 4$ did not reveal any further biologically relevant structure and mostly produced genotypes with intermediate membership probabilities, thus only the membership coefficients for runs with $2 \leq K \leq 6$ are presented. A similar pattern of clustering was obtained with the nonparametric DAPC analysis and the network obtained using the neighbor-net method (Figs. 3C and 4).

When AMOVA was performed by grouping haplotypes according to the inferred clusters, 51% of genetic diversity was distributed among clusters and only 10.3% among geographical regions within clusters. However, when we partitioned the variation according to geographical origin, nested within the inferred clusters, there was no difference between regions and 78.26% of the diversity was
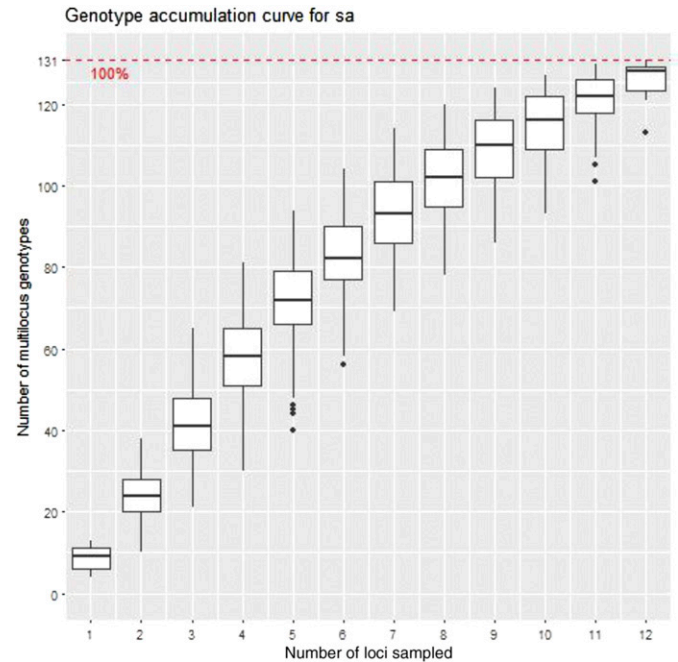

Fig. 2. Haplotype accumulation curve for 13 microsatellite loci in 10 *Colletotrichum truncatum* populations. Number of observed multilocus haplotypes are denoted by the vertical axis.

TABLE 1. Populations of soybean-infecting *Colletotrichum truncatum* evaluated in this study[a]

State	County	Population	N	Geographical coordinates	Cultivar	Sampling year
Goiás	Montividiu	GO1	15	17°25′45.3′′S - 51°10′14.2′′W	Bonus	2017
	Montividiu	GO2	13	17°24′43.1′′S - 50°57′29.0′′W	7739 Monsoy	2017
	Montividiu	GO3*	3	17°25′50.9′′S - 51°01′10.9′′W	7739 Monsoy	2017
	Rio Verde	GO4	34	17°27′36.2′′S - 51°07′10.3′′W	PP7200 Macro seed	2017
	Rio Verde	GO5	53	17°45′51.5′′S - 51°02′06.9′′W	Nidera 5909	2017
Mato Grosso	Lucas do Rio Verde	MT1	13	13°10′10.1′′S - 56°04′08.0′′W	Y-70 Pioneer	2017
	Sinop	MT2	20	13°18′46.7′′S - 56°02′33.4′′W	8766 Monsoy	2016
	Lucas do Rio Verde	MT3	28	13°24′39.4′′S - 56°04′04.0′′W	8372 Monsoy	2017
	Sinop	MT4	24	11°55′22.8′′S - 55°37′00.0′′W	7709 Nidera	2017
	Sinop	MT5	34	11°58′58.5′′S - 55°30′28.6′′W	7709 Nidera	2017

[a] Population GO3 was excluded from genetic diversity analyses due to its small sample size.

found among clusters within regions. The highest pairwise differentiation was observed among clusters within regions (Supplementary Table S3), which indicates that the genetic variation was mostly distributed between clusters and not between geographic regions within clusters.

Genetic diversity was estimated for both samples (Table 2) and inferred genetic clusters (Table 3). In the clone-corrected dataset of 166 MLHs, 131 unique haplotypes were detected across the three clusters, not associated with the geographic origin of isolates. The first cluster (C1), composed of 81 haplotypes, encompassed the highest proportion of repeated haplotypes (clonal fraction = 0.37) and was the least diverse ($H = 0.24$; $N_a = 3.85$; $A_R = 2.51$). The second cluster (C2) showed 26 isolates with intermediate level of diversity (clonal fraction = 0.19; $H = 0.34$; $N_a = 2.54$; $A_R = 2.54$). The 59 isolates assigned to the third cluster (C3) were all unique

MLHs, making this cluster the most diverse one ($H = 0.70$; $N_a = 5.46$; $A_R = 6.65$).

Reproductive mode. Levels of multilocus linkage disequilibrium r_D were low and consistent with recombination within clusters (r_D ranging from 0.03 to 0.07), but the null hypothesis ($r_D = 0$) for a random association of alleles between pairs of loci was rejected for all clusters (Table 3). Multilocus linkage disequilibrium was higher within populations (r_D ranging from 0.28 to 0.60; Table 2) than within clusters, consistent with the coexistence of multiple clusters within populations (Fig. 3).

DISCUSSION

This study represents the first investigation of *C. truncatum* population structure associated with soybean anthracnose in Brazil.

TABLE 2. Estimates of genetic and haplotypic diversity among *Colletotrichum truncatum* populations associated with soybean anthracnose in Brazil

Population	N[a]	MLHs[b]	Genotypic diversity[c]	Clonal fraction[d]	H[e]	N_a[f]	A_R[g]	r_D[h] (P)[i]
GO1	15	13	0.87	0.13	0.76	4.8 (0.3)	4.40	0.33
GO2	13	11	0.85	0.15	0.62	3.9 (0.3)	3.77	0.60
GO3	3	3	1.00	0.00	N/A[j]	N/A	N/A	N/A
GO4	34	18	0.53	0.47	0.45	4.7 (0.5)	3.85	0.47
GO5	53	35	0.66	0.34	0.61	5.7 (0.4)	2.13	0.54
MT1	13	13	1.00	0.00	0.68	3.8 (0.2)	3.58	0.28
MT2	20	16	0.80	0.20	0.59	4.6 (0.3)	3.84	0.44
MT3	28	14	0.50	0.50	0.21	2.9 (0.4)	2.48	0.40
MT4	24	19	0.79	0.21	0.66	5.2 (0.4)	4.12	0.45
MT5	34	24	0.71	0.29	0.64	5.2 (0.4)	3.92	0.31
Total	237	129	0.54	0.46	0.63	4.4 (0.1)	–	–

[a] Sample size.
[b] Number of distinct multilocus haplotypes (MLHs).
[c] Number of distinct genotypes/sample size or fraction of nonrepeated genotypes in the sample.
[d] 1 − haplotypic diversity.
[e] Nei's unbiased gene diversity (Nei 1978).
[f] Number of distinct alleles averaged across loci and standard deviation in parenthesis.
[g] Allelic richness estimated by rarefaction to $N = 11$ (Szpiech et al. 2008).
[h] Multilocus linkage disequilibrium (Agapow and Burt 2001).
[i] $P = 0.001$ for all r_D values (associated probability estimated after 1,000 randomizations).
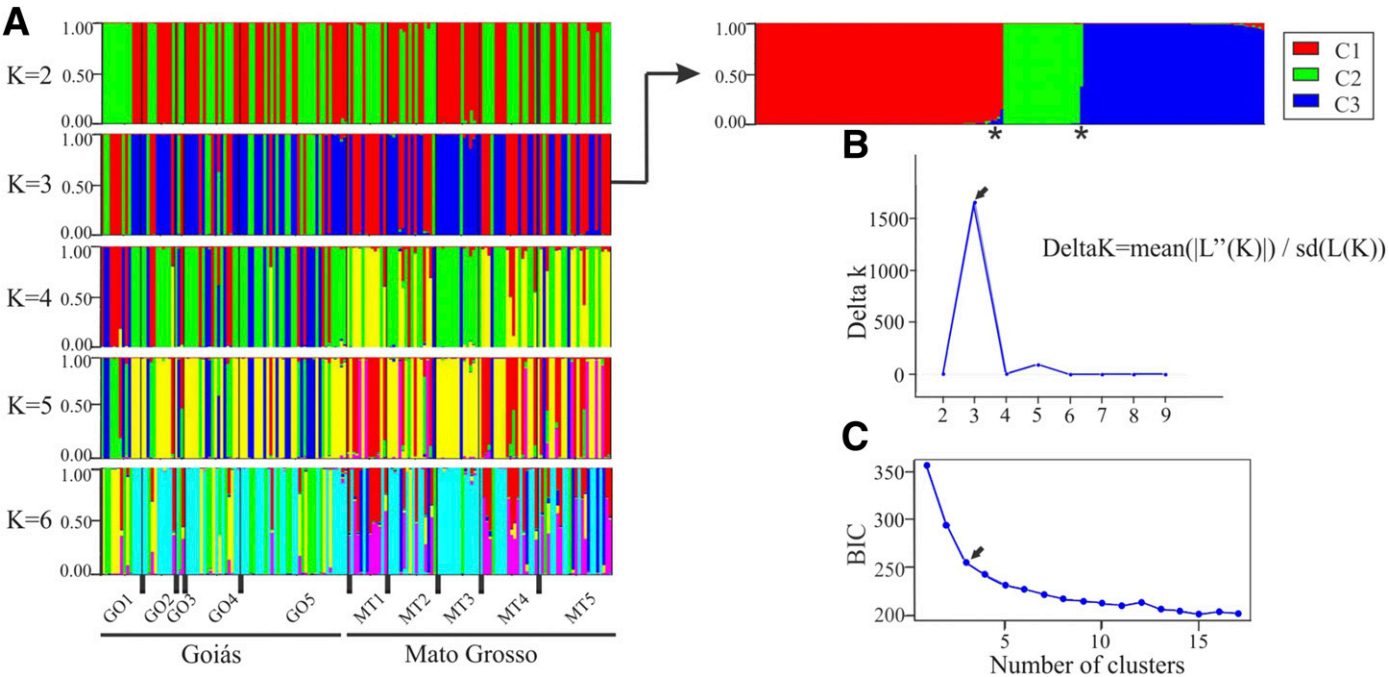[j] N/A indicates small sample size.

Fig. 3. Population subdivision of *Colletotrichum truncatum* associated with soybean anthracnose in Brazil. **A,** STRUCTURE barplots showing patterns of clustering in the data. Haplotypes with maximum membership proportion to a single cluster below 98% are denoted by asterisks. **B,** The most likely number of clusters (K) based on Evanno's ΔK method. **C,** Bayesian information criteria (BIC) indicating the most probable number of genetic groups by discriminant analysis of principal components analysis.

We used SSRs markers developed for the synonym species *C. capsici* (Damm et al. 2009) that have already proved to be highly efficient in the characterization of *C. truncatum* populations (Diao et al. 2015). This set of markers showed sufficient power to discriminate haplotypes and detect polymorphisms in this species. Soybean anthracnose is a complex system, with many unresolved issues. Our previous study, based on *C. truncatum* isolates obtained from different soybean-producing regions of Brazil, showed moderate genetic variability and sharing of haplotypes among areas (Rogério et al. 2016). The current study has focused on a population-level distribution of genetic diversity across pathogen populations sampled from central-west (Mato Grosso and Goiás states), the main soybean producing areas in Brazil. We tested the hypotheses that geographical populations of *C. truncatum* are genetically homogeneous (hence not subdivided or differentiated) and that these populations have a recombining population structure.

The *C. truncatum* isolates obtained in this study showed relatively high levels of genetic variability, which would be
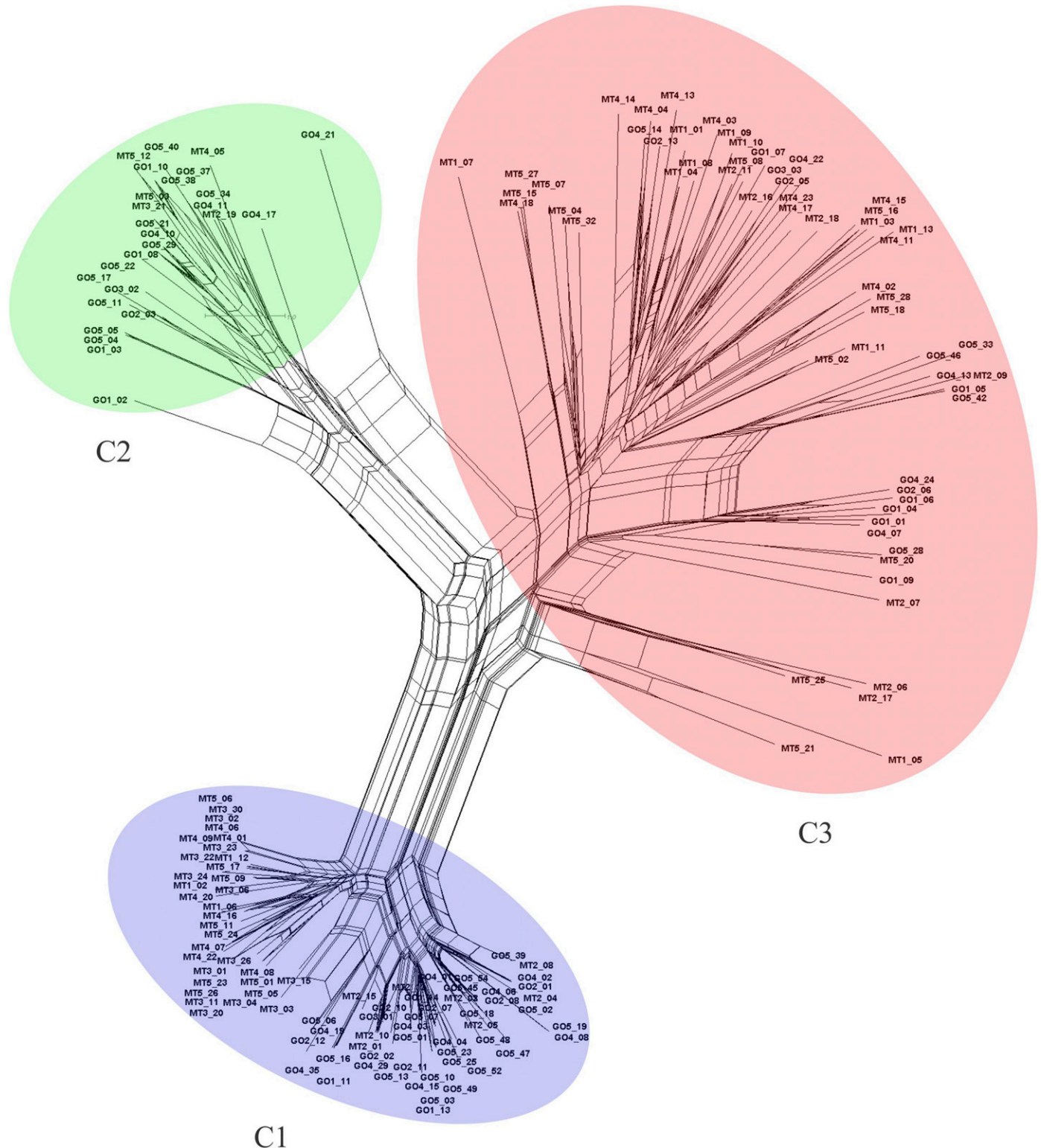

Fig. 4. Phylogenetic network using the neighbor-net methods of *Colletotrichum truncatum* isolates used in this study. Reticulation indicates likely occurrence of recombination.

unexpected for a recently introduced pathogen. However, multiple introductions of genetically diverse plant pathogens could increase the level of genetic variation of local populations (Ciampi-Guillardi et al. 2014; Leo et al. 2015). *C. truncatum* populations were also characterized by moderate to high haplotypic diversity and low to moderate clonal fractions, with some MLHs shared among areas separated by more than 800 km. The high genetic diversity observed in Brazilian populations of *C. truncatum* is in agreement with previous reports for populations of *C. truncatum* isolated from chili in China and India, using the same set of markers (Diao et al. 2015; Sharma et al. 2014). High levels of genetic variability in *C. truncatum* were also observed in previous studies using different molecular markers (Ford et al. 2004; Katoch et al. 2017; Rogério et al. 2016; Sant'anna et al. 2010; Sharma 2009; Vasconcelos et al. 1994).

Despite the fact that low population differentiation was detected between regions, clustering and molecular variance analyses indicated the existence of three strongly differentiated genetic clusters within each region. This finding is consistent with a possible history of introduction of *C. truncatum* from at least three distinct source populations, leading to three genetic clusters that spread throughout the country. The majority of soybean cultivars currently sown in Brazil are derived of cultivars introduced from the United States that were used in breeding programs (Wysmierski and Vello 2013). As *C. truncatum* is generally seedborne (Manandhar and Hartman 1999), it could be easily introduced in Brazil by infected seeds from the United States. Our previous work corroborates the hypothesis of three distinct source populations (Rogério et al. 2016). Sequencing data from a historical collection of *C. truncatum* isolates from different soybean-producing regions, revealed three genetic clades. A first clade encompassed *C. truncatum* strains isolated from soybean in the United States, which grouped together with the most widely distributed haplotype within that collection. Interestingly, this clade included strain CBS19532, which is apparently the same one regarded as the ex-holotype of *Glomerella glycines* (ATCC 1936) by Lehman and Wolf in 1926 (Damm et al. 2019). This strain also corresponds to drawings and description of the asexual morph of *G. glycines* (*C. glycines*), later classified as *C. truncatum* by Damm et al. (2009). It is possible that this strain has a connection with the *C. glycines* Hori previously identified at Korea by Hemmi in 1917, as described by Tiffany (1950) as the first causal agent of anthracnose disease in soybean. A second clade enclosed *C. truncatum* sequences from lima bean in Brazil and the United States that is also a host susceptible to the species, as some other Fabaceae (Carvalho et al. 2015; Damm et al. 2009). When soybeans were introduced in Brazil and became widely planted, other hosts as lima bean might have already established in the cropping system (Silva et al. 2017), and due to broad host range of the fungus, that lineage might be able to maintain itself on this new host. This hypothesis had already been considered for the widespread occurrence of *C. truncatum*, originally from lima beans, on

soybeans in the United States (Tiffany 1950). The third clade included strains from hosts as weeds and other leguminous plants. The clade that tends to be exclusively associated with soybean tends to be the less variable, while the cluster associated with soybean and weeds was the most variable. These observations suggest that future studies on the origin of invasive *C. truncatum* populations will need to cover multiple geographic regions and multiple hosts.

Although the three clusters detected in Brazilian *C. truncatum* were found to coexist in the majority of populations, admixture between clusters was not detected, indicating the existence of barriers to gene flow between them. In contrast with our initial hypothesis of geographical populations of *C. truncatum* being genetically homogeneous due to intense movement of infected material among fields, our results do not indicate a strong impact of intraregional gene flow on patterns of genetic structure. Our findings do suggest, however, that haplotype flows have played a greater role in structuring populations mainly in the maintained the genetic distribution of the three clusters within fields. Differentiation between regions (Mato Grosso and Goiás) was low to moderate, a substantial proportion of MLHs were shared among populations across regions and consecutive years, and levels of variability in the two regions were similar, suggesting extensive transportation of the pathogen across regions, possibly via contaminated seed (Ciampi et al. 2008). The fact that populations from Mato Grosso showed higher differentiation among themselves than observed in Goiás suggests a greater impact of founder events in the former region. This observation is not consistent with strict host-tracking of the pathogen during the spread of soybean cultivation in Brazil, which began in the south and extended toward the Cerrado, reaching Mato Grosso first (in the late 1980s) and only later in Goiás (Fearnside 2001).

Colletotrichum species have been predominantly observed in the asexual or vegetative state. Only a small number of species has been associated with the teleomorph, *Glomerella*, and in general, these forms are rarely observed in nature, with a few notable exceptions (Barcelos et al. 2014; Damm, et al. 2012). The genus *Colletotrichum* has an unusual mating system and to date all the strains sampled possess only one idiomorph (Barcelos et al. 2014; Chen et al. 2002; Crouch et al. 2008). For *C. truncatum*, the identity of the sexual stage is still unclear, since it has not been described either in laboratory conditions or in nature (Damm et al. 2009; Hyde et al. 2009; Katoch et al. 2017). A recent study found population genetic evidence for sexual recombination in *C. truncatum* populations infecting chili from China (Diao et al. 2015), which supported our initial hypothesis that *C. truncatum* infecting soybean in Brazil may be recombinant in structure. However, this hypothesis was not fully corroborated by our population genetic analyses, as estimates of multilocus linkage disequilibrium were low in all three clusters, but the hypothesis of random association of alleles was still rejected. The limited evidence for admixture between clusters is also consistent with lack of recombination resulted from sexual reproduction. The moderate to high levels of haplotypic diversity

TABLE 3. Estimates of genetic and genotypic diversity of the genetic clusters of *Colletotrichum truncatum* associated to soybean in Brazil

Clusters	Code	N^a	MLHs[b]	Genotypic diversity[c]	Clonal fraction[d]	H^e	$N_a{}^f$	$A_R{}^g$	$r_D{}^h$ (P)[i]
Cluster 1	C1	81	51	0.63	0.37	0.24	3.85 (0.66)	2.51	0.03 (0.003)
Cluster 2	C2	26	21	0.81	0.19	0.34	2.54 (0.31)	2.54	0.08 (0.001)
Cluster 3	C3	59	59	1.00	0.00	0.61	5.46 (0.72)	4.65	0.07 (0.001)
Total	–	166	131	0.79	0.21	0.70	3.95 (0.39)	8.31	0.39 (0.001)

[a] Sample size.
[b] Number of distinct multilocus haplotypes (MLHs).
[c] Number of distinct genotypes/sample size or fraction of nonrepeated genotypes in the sample.
[d] 1 – haplotypic diversity.
[e] Nei's unbiased gene diversity (Nei 1978).
[f] Number of distinct alleles averaged across loci and standard deviation in parenthesis.
[g] Allelic richness estimated by rarefaction to $N = 21$ (Szpiech et al. 2008).
[h] Multilocus linkage disequilibrium (Agapow and Burt 2001).
[i] P = associated probability estimated after 1,000 randomizations.

detected in clonal pathogens could be generated by mechanisms such as extensive chromosome rearrangements mediated by transposable elements (Stukenbrock 2016). A population genomic approach could make it possible to resolve the reproductive mode of the three Brazilian clusters of *C. truncatum*, by giving access to the mating types, their frequencies, and genomic rates of recombination.

This study is the first to use SSR markers and analyze the population structure of *C. truncatum* isolates infecting soybean in Brazil. The results obtained here provide novel insights about the pathogen introduction in Brazil and their implication in distribution of genetic variation across populations from two important areas of soybean production. The current genetic structure of the pathogen reflects possible multiple recent introduction events, represented by three genetic groups widely distributed across soybean fields, raising questions with respect to the factors allowing their maintenance in syntopy without evidence of admixture between them. These observations can have a direct impact on disease management strategies. Although there is currently no breeding program aiming anthracnose resistance for soybean, future breeding efforts should take into account the current population structure and the genetic diversity levels of the pathogen, for instance by using isolates that are representative of the genetic variability of the species in Brazil for screening resistant varieties. The possible lack of sexual reproduction is also important from the plant pathology point of view, as it may contribute to prevent admixture between clusters and the emergence of recombinant haplotypes with increased virulence. Population genomics studies using higher resolution molecular markers such as single nucleotide polymorphisms may improve the discrimination of clonal lineages, the detection of recombination events in *C. truncatum* populations in Brazil and can provide further insight into the history and population biology of this important plant pathogen.

ACKNOWLEDGMENTS

LITERATURE CITED

Agapow, P. M., and Burt, A. 2001. Indices of multilocus linkage disequilibrium. Mol. Ecol. Notes 1:101-102.

Alfenas, A. C., and Mafia, R. G. 2007. Armazenamento de microorganismos em cultura com Ênfase em fungos fitopatogênicos. Pages 91-102 in: Métodos em Fitopatologia. R. C. Gonçalves, A. C. Alfenas, and R. G. Mafia, eds. UFV, Viçosa.

Ali, S., Gladieux, P., Leconte, M., Gautier, A., Justesen, A. F., Hovmoller, M. S., Enjalbert, J., and De Vallavieille-Pope, C. 2014. Origin, migration routes and worldwide population genetic structure of the wheat yellow rust pathogen *Puccinia striiformis* f. sp. *tritici*. PLoS Pathog. 10:e1003903.

Andrus, C. F., and Moore, W. D. 1935. *Colletotrichum truncatum* (Schw.), n. comb., on garden and lima beans. Phytopathology 25:121-125.

Arantes, N. E., and Miranda, M. A. C. 1993. Melhoramento genético e cultivares de soja para o cerrado da região sudeste do Brazil. Pages 225-266 in: Cultura da Soja nos Cerrados. N. E. Arantes and P. I. M.Souzaed, eds. Potafos, Piracicaba.

Araújo, A. G., Café-Filho, A. C., and Cupertino, F. P. 1988. Antracnose da soja na região geoeconômica do Distrito Federal. Fitopatol. Bras. 13:130.

Barcelos, Q. L., Pinto, J. M. A., Vaillancourt, L. J., and Souza, E. A. 2014. Characterization of *Glomerella* strains recovered from anthracnose lesions on common bean plants in Brazil. PLoS One 9:e90910.

Baroncelli, R., Zapparata, A., Sarrocco, S., Sukno, S. A., Lane, C. R., Thon, M. R., Vannacci, G., Holub, E., and Screenivasaprasad, S. 2015. Molecular diversity of anthracnose pathogen populations associated with UK strawberry production suggests multiple introductions of three different *Colletotrichum* species. PLoS One 10:e0129140.

Cannon, P. F., Damm, U., Johnston, P. R., and Weir, B. S. 2012. *Colletotrichum*–Current status and future directions. Stud. Mycol. 73:181-213.

Carvalho, E. M. S., Beserra, J. E. A., and Barguil, B. M. 2015. Lima bean diseases. Pages 13-113 in: *Phaseolus lunatus*: Diversity, Growth and Production. A. C. A. Lopes, A. S. A. Ferreira, and R. L. F. Gomes, eds. Nova Science Publishers, New York.

Castellani, A. 1939. Viability of some pathogenic fungi in distilled water. Am. J. Trop. Med. Hyg. 24:270-276.

Chen, F., Goodwin, P. H., Khan, A., and Hsiang, T. 2002. Population structure and mating-type genes of *Colletotrichum graminicola* from *Agrostis palustris*. Can. J. Microbiol. 48:427-436.

Ciampi, M. B., Meyer, M. C., Costa, M. J. N., Zala, M., McDonald, B. A., and Ceresini, P. C. 2008. Genetic structure of populations of *Rhizoctonia solani* anastomosis group-1 IA from soybean in Brazil. Phytopathology 98:932-941.

Ciampi-Guillardi, M., Baldauf, C., Souza, A. P., Silva-Junior, G. J., and Amorim, L. 2014. Recent introduction and recombination in *Colletotrichum acutatum* populations associated with citrus postbloom fruit drop epidemics in São Paulo, Brazil. Phytopathology 104:769-778.

Crouch, J. A., Glasheen, B. M., Giunta, M. A., Clarke, B. B., and Hillman, B. I. 2008. The evolution of transposon repeat-induced point mutation in the genome of *Colletotrichum cereale*: Reconciling sex, recombination and homoplasy in an "asexual" pathogen. Fungal Gen. Biol. 45:190-206.

Damm, U., Cannon, P. F., Woudenberg, J. H. C., and Crous, P. W. 2012. The *Colletotrichum acutatum* species complex. Stud. Mycol. 73:37-113.

Damm, U., Sato, T., Alizadeh, A., Groenewald, J. Z., and Crous, P. W. 2019. The *Colletotrichum dracaenophilum*, *C. magnum* and *C. orchidearum* species complexes. Stud. Mycol. 92:1-46.

Damm, U., Woudenberg, J. H. C., Cannon, P. F., and Crous, P. W. 2009. *Colletotrichum* species with curved conidia from herbaceous hosts. Fungal Divers. 39:45-87.

Desprez-Loustau, M. L., Robin, C., Buee, M., Courtecuisse, R., Garbaye, J., Suffert, F., Sache, I., and Rizzo, D. M. 2007. The fungal dimension of biological invasions. Trends Ecol. Evol. 22:472-480.

Diao, Y., Zhang, C., Xu, J., Lin, D., Liu, L., Mtung'e, O. G., and Liu, X. 2015. Genetic differentiation and recombination among geographic populations of the fungal pathogen *Colletotrichum truncatum* from chili peppers in China. Evol. Appl. 8:108-118.

Dias, M. D., Pinheiro, V. F., and Café-Filho, A. C. 2016. Impact of anthracnose on the yield of soybean subjected to chemical control in the north region of Brazil. Summa Phytopathol. 42:18-23.

El Mousadik, A., and Petit, R. J. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. Theor. Appl. Genet. 92:832-839.

EMBRAPA. 1996. Cultura da soja. edited by Ministério da Agricultura Pecuária e Abastecimento. Empresa Brasileira de Pesquisa Agropecuária-Embrapa Soja, Embrapa Cerrados, Embrapa Agropecuária Oeste, Brazil.

Excoffier, L., and Lischer, H. E. L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564-567.

Excoffier, L., Smouse, P., and Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Falush, D., Stephens, M., and Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.

Fearnside, P. M. 2001. Soybean cultivation as a threat to the environment in Brazil. Environ. Conserv. 28:23-38.

Ford, R., Banninza, C., Photita, W., and Taylor, P. W. J. 2004. Morphological and molecular discrimination of *Colletotrichum truncatum* causing anthracnose on lentil in Canada. Australas. Plant Pathol. 33:559-569.

Gladieux, P., Feurtey, A., Hood, M. E., Snirc, A., Clavel, J., Dutech, C., Roy, M., and Giraud, T. 2015. The population biology of fungal invasions. Mol. Ecol. 24:1969-1986.

Gladieux, P., Zhang, X.-G., Afoufa-Bastien, D., Sanhueza, R. M. V., Sbaghi, M., and Cam, B. L. 2008. On the origin and spread of the scab disease of apple: Out of central Asia. PLoS One 3:e1455.

Grünwald, N. J., Garbelotto, M., Goss, E. M., Heungens, K., and Prospero, S. 2012. Emergence of the sudden oak death pathogen *Phytophthora ramorum*. Trends Microbiol. 20:131-138.

Hallatschek, O., and Nelson, D. R. 2008. Gene surfing in expanding populations. Theor. Popul. Biol. 73:158-170.

Hirimoto, D. M., and Vello, N. A. 1986. The genetic base of Brazilian soybean cultivars. Braz. J. Genet. 9:295-306.

Hyde, K. D., Cai, L., McKenzie, E. H. C., Yang, Y. L., Zhang, J. Z., and Prihastuti, H. 2009. *Colletotrichum*, a catalogue of confusion. Fungal Divers. 29:1-17.

Hymowitz, T. 1970. On the domestication of the soybean. Econ. Bot. 24: 408-421.

Hymowitz, T. 1990. Soybeans: The success story. Pages 159-163 in: Advances in new crops. J. Janick and J. E. Simon, eds. Timber Press, Portland, OR.

Jombart, T., and Ahmed, I. 2011. Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071.

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

Kamvar, Z. N., Tabima, J. F., and Grunwald, N. J. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281.

Katoch, A., Sharma, P., Padder, B. A., and Sharma, P. N. 2017. Population structure of *Colletotrichum truncatum* in Himachal Pradesh and identification of broad-spectrum resistant sources in *Capsicum*. Agric. Res. 6: 296-303.

Lehman, S. G., and Wolf, F. A. 1926. Soybean anthracnose. J. Environ. Agric. Res. 33:381-390.

Leo, A. E., Ford, R., and Linde, C. C. 2015. Genetic homogeneity of a recently introduced pathogen of chickpea, *Ascochyta rabiei*, to Australia. Biol. Invasions 17:609-623.

Mahmodi, F., Kadir, J. B., Puteh, A., Pourdad, S. S., Nasehi, A., and Soleimani, N. 2014. Genetic diversity and differentiation of *Colletotrichum* spp. isolates associated with Leguminosae using multigene loci, RAPD and ISSR. Plant Pathol. J. 30:10-24.

Manandhar, J. B., and Hartman, G. L. 1999. Anthracnose. Pages 13-14 in: Compendium of Soybean Diseases. G. L. Hartman, J. B. Sinclair, and J. C. Rupe, eds. American Phytopathological Society, St. Paul, MN.

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

Milgroom, M. G. 1996. Recombination and the multilocus structure of fungal populations. Annu. Rev. Phytopathol. 34:457-477.

Milgroom, M. G., and Fry, W. E. 1997. Contributions of population genetics to plant disease epidemiology and management. Adv. Bot. Res. 24:1-30.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a number of individuals. Genetics 89:538-590.

Peakall, R., and Smouse, P. E. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6: 288-295.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Ranathunge, N. P., Ford, R., and Taylor, P. W. J. 2009. Development and optimization of sequence-tagged microsatellite site markers to detect genetic diversity within *Colletotrichum capsici*, a causal agent of chilli pepper anthracnose disease. Mol. Ecol. Resour. 9:1175-1179.

Rogério, F., Ciampi–Guillardi, M., Barbieri, M. C., Bragança, C. A. D., Seixas, C. D. S., Almeida, A. M. R., and Massola, N. S., Jr. 2016.

Phylogeny and variability of *Colletotrichum truncatum* associated to soybean anthracnose in Brazil. J. Appl. Microbiol. 122:402-415.

Sant'anna, J. R., Miyamoto, C. T., Rosada, L. J., Franco, C. C. S., Kaneshima, E. N., and Castro-Prado, M. A. A. 2010. Genetic relatedness of Brazilian *Colletotrichum truncatum* isolates assessed by vegetative compatibility groups and RAPD analysis. Biol. Res. 43:51-62.

Sharma, G., Pinnaka, A. K., and Shenoy, B. D. 2014. Infra-specific diversity of *Colletotrichum truncatum* associated with chilli anthracnose in India based on microsatellite marker analysis. Arch. Phytopathol. Plant Prot. 47: 2509-2523.

Sharma, R. 2009. Genetic differentiation of host limited forms of *Colletotrichum truncatum* from northwestern Himalayas. Arch Phytopathol. Plant Prot. 42:960-966.

Silva, R. N. O., Burle, M. L., Pádua, J. G., Lopes, A. C. A., Gomes, R. L. F., and Castillo, J. M. 2017. Phenotypic diversity in lima bean landraces cultivated in Brazil, using the Ward-MLM strategy. Chil. J. Agric. Res. 77: 35-40.

Stukenbrock, E. H. 2016. The role of hybridization in the evolution and emergence of new fungal plant pathogens. Phytopathology 106:104-112.

Stukenbrock, E. H., and McDonald, B. A. 2008. The origins of plant pathogens in agro-ecosystems. Annu. Rev. Phytopathol. 46:75-100.

Szpiech, Z. A., Jakobsson, M., and Rosenberg, N. A. 2008. ADZE: A rarefaction approach for counting alleles private to combinations of populations. Bioinformatics 24:2498-2504.

Tiffany, L. H. 1950. The anthracnose complex on soybeans. Retrospective Theses and Dissertations. Iowa State University. UMI Number DP12218.

Vasconcelos, M. J. V., Machado, M. A., Almeida, A. M. R., Henning, A. A., Barros, E. G., and Moreira, M. A. 1994. Differentiation of *Colletotrichum truncatum* isolates by random amplified polymorphic DNA. Fitopatol. Bras. 19:520-523.

Weidemann, G. J., Tebeest, D. O., and Cartwright, R. D. 1988. Host specificity of *Colletotrichum gloeosporioides* f. sp. *aeschynoemene* and *C. truncatum* in the Leguminosae. Phytopathology 78:986-990.

Weir, B. S., and Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.

Wysmierski, P. T., and Vello, N. A. 2013. The genetic base of Brazilian soybean cultivars: Evolution over time and breeding implications. Genet. Mol. Biol. 36:547-555.

Yarwood, C. E. 1970. Man-made plant diseases. Science 168:218-220.