Central American and Caribbean population history of the Pseudocercospora fijiensis fungus responsible for the latest worldwide pandemics on banana

Jean Carlier, Stéphanie Robert, Véronique Roussel, Yolande Chilin-Charles, Nadia Lubin-Adjanoh, Aude Gilabert, Catherine Abadie

Central American and Caribbean population history of the *Pseudocercospora fijiensis* fungus responsible for the latest worldwide pandemics on banana

Authors:

Jean Carlier[1,2]*†, Stéphanie Robert[1,2]†, Véronique Roussel[1,2], Yolande Chilin-Charles[1,3], Nadia Lubin-Adjanoh [1,3], Aude Gilabert[1,2], Catherine Abadie[1,3]

Affiliations :

[1] BGPI, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France.

[2] CIRAD, UMR BGPI, F-34398 Montpellier, France.

[3] CIRAD, UMR BGPI, F-97130 Capesterre-Belle-Eau, Guadeloupe, France.

*Correspondence to: jean.carlier@cirad.fr

†These authors contributed equally to this work

Abstract Among the emerging fungal diseases threatening food security, the *Pseudocercospora fijiensis* fungus causing black leaf streak disease of banana is one of the most marked examples of a recent worldwide pandemic on a major crop. We assessed how this pathogen spread throughout the latest invaded region, i.e. Central America and the Caribbean. We retraced its population history combining detailed monitoring information on disease outbreaks and population genetic analyses based on large-scale sampling of *P. fijiensis* isolates from 121 locations throughout the region. The results first suggested that sexual reproduction was not lost during the *P. fijiensis* expansion, even in the insular Caribbean context, and a high level of genotypic diversity was maintained in all the populations studied. The population genetic structure of *P. fijiensis* and historical data showed that two disease waves swept northward and southward in all banana-producing countries in the study area from an initial entry point in Honduras, probably mainly through gradual stepwise spore dispersal. Serial founder events accompanying the northern and southern waves led to the establishment of two different genetic groups. A different population structure was detected on the latest invaded islands (Martinique, Dominica and Guadeloupe), revealing multiple introductions and admixture events that may have been partly due to human activities. The results of this study highlight the need to step up surveillance to limit the spread of other known emerging diseases of banana spread mainly by humans, but also to curb gene flow between established pathogen populations which could increase their evolutionary potential.

Keywords: Emerging disease, fungal plant pathogen, population history, dispersal, *Musa*, *Pseudocercospora fijiensis*.

1. Introduction

Of all plant pathogens, fungi are a highly serious threat to crops (Almeida et al., 2019). Global trade and large-scale standardized farming are conducive to fungal disease dispersal and establishment worldwide (McDonald and Stukenbrock, 2016). Thus many so-called 'emerging plant diseases' are caused by pathogens that have increased their geographical range and become invasive in new and wide areas (Anderson et al., 2004). Studying diseases that have recently emerged may help improve surveillance (Parnell et al., 2017) and they may also be useful models for understanding evolutionary and ecological processes involved in pathogen populations, such as dispersal, reproduction mode, population bottlenecks, admixture and changes in phenotypic traits (Stukenbrock and McDonald, 2008). It is first essential to reconstruct the history of pathogen populations during range expansions when documenting these processes in emerging diseases (Gladieux et al., 2015).

The detection dates of a given emerging disease across geographical areas could first be documented to retrace the history of pathogen populations. This spatiotemporal or historical data can be used to estimate the speed of past emergences, which is a key aspect for understanding and reacting to future invasive species (Evans, 2016). However, historical data are not informative enough to retrace the population history and identify the epidemiological and demographic processes that have accompanied the spread of emerging fungal diseases. Population genetics are often a valuable complementary approach for gaining clearer insight into the population biology of a fungal invasion (review in Gladieux et al., 2015; McDonald and Stukenbrock, 2016). Such studies have notably shown that, although fungi can be dispersed naturally through spore production, humans also play a role in the spread of diseases via the transport of agricultural products or infected plant materials. This role seems most evident between continents because it is harder for pathogens to travel several thousand kilometres by natural means alone. Yet it is not easy to disentangle human-mediated introductions from natural dispersal on smaller geographical scales. Several demographic or

66 evolutionary processes often accompany fungal invasions, which may promote adaptations
67 such as population bottlenecks, or admixture between populations following multiple
68 introductions. Lastly, changes in life history traits during fungal invasions have been
69 described and the loss of sexual reproduction is one of the most frequent.
70
71 Banana—a major high-value crop produced in tropical regions—is threatened by several
72 pathogens (Guzman et al., 2019), including the fungus *Pseudocercopora fijiensis* which
73 causes black leaf streak disease (BLSD). This emerging disease has spread from Asia
74 throughout the entire intertropical zone in the last 50 years and is considered to be one of the
75 most serious threats to food security (Pennisi, 2010). In some areas, over 50 fungicides are
76 applied annually to control the disease, which obviously has a major impact on human health
77 and the environment (de Lapeyre de Bellaire et al., 2010).
78 *P. fijiensis* is a heterothallic ascomycete fungus that produces two types of spores. Conidia
79 (from asexual reproduction) disperse over scales of only a few metres, whereas ascospores
80 (from sexual reproduction) are easily wind-dispersed over an average distance of 200 m, but
81 with a fat-tailed dispersal kernel that may extend several kilometres (Rieux et al., 2014).
82 Sexual and asexual reproduction concomitantly prevail in banana plantations (Guzman et al.,
83 2019). Human-mediated introductions may also accelerate regionwide disease spread through
84 the movement of infected leaf material (planting or packing materials, (Churchill, 2011). In a
85 continental context, sexual reproduction is not lost during invasion and—although genetic
86 discontinuity indicative of some stochastic long-distance dispersal events has been
87 reported—the genetic structure of *P. fijiensis* populations on a country scale suggest that the
88 disease is spread mainly by gradual stepwise spore dispersal (Halkett et al., 2010). Supporting
89 this observation, the rate of *P. fijiensis* spread estimated from the spatiotemporal data
90 published in (Halkett et al., 2010) was close to that predicted under certain conditions by a
91 modelling approach accounting for both asexual and sexual reproduction (Hamelin et al.,
92 2016).

In the Americas, BLSD was first detected in Honduras in 1972 and was most probably introduced by humans (Robert et al., 2012; Robert et al., 2015), spreading northwards and southwards and then reaching the Caribbean in 1990 (Guzman et al., 2019). Despite surveillance networks set up in the French West Indies (Teycheney et al., 2009), the last disease-free islands of Martinique, Dominica and Guadeloupe were reached and infested by BLSD from 2010 onwards (Ioos et al., 2011). This situation offered an opportunity to monitor a recent fungal invasion in an insular context, which could provide new information on the population biology of such organisms.

This study set out to retrace the population history of *P. fijiensis* throughout Central America and the Caribbean to gain insight into the epidemiological and demographic processes that have accompanied its spread in this latest invaded region. To that end, we gathered detailed monitoring information on disease outbreaks within the region. At the same time, we conducted a population genetic analysis using large-scale sampling throughout all banana-producing countries of mainland Central and South America and the Caribbean Basin. We used these data to address the following questions: What were the BLSD invasion routes throughout Central America and the Caribbean Islands? Was the invasion speed slower in the Caribbean insular context in comparison to mainland areas? Did successive bottlenecks accompany this invasion? Did the expansion have an impact on sexual reproduction when the fungus passed from island to island in the Carribean? Could this analysis generate insight into the potential invasion drivers, e.g. natural fungal dispersal by the wind, or human activities?

2. Materials and Methods

2.1. Historical data analysis

Historical *P. fijiensis* invasion data were collected in 12 Caribbean banana-producing countries. We noted the date and geographical coordinates (decimal units for latitude and longitude) for the locations corresponding to the first, middle and last detection of the disease

120 in each country (Supplementary Dataset 1). The historical dataset was divided into two
121 subsets corresponding to a northern front and a southern front. The northern front was
122 represented by five countries (Cuba, Jamaica, Haiti, Dominican Republican, Puerto Rico)
123 mainly distributed along a longitudinal east-west gradient and the southern front by seven
124 countries (Trinidad, Grenada, St Vincent, St Lucia, Martinique, Dominica, Guadeloupe)
125 mainly distributed along a latitudinal north-south gradient. The time of occurrence in each
126 location was calculated as the time between the date of detection in the location and the date
127 of the first appearance in the northern or southern Caribbean, i.e. the Esmeralda location in
128 Cuba (21.850°N, -78.100°W) and Coromandel location in Trinidad (10.117°N, -61.800°W),
129 respectively. These locations were considered as the origin of disease spread in the two
130 Caribbean areas. A linear model between the occurrence time and geographical distance from
131 the origin of each front was tested using the lm function of R 3.6.0 (http://www.R-
132 project.org). The Euclidian distance was chosen because the study locations for each of the
133 fronts were more or less distributed along the same line (from west to east for the northern
134 front and from south to north for the southern front). The speed of the invasion waves was
135 calculated as the inverse of the regression slope.
136
137 *2.2. Fungal material sampling and DNA extraction*
138 Leaf samples were collected from most banana-growing countries in mainland tropical
139 of Central, South America and the Caribbean Basin (Table1, Supplementary Dataset 1).
140 Overall, 2,188 fungal individuals from 121 different geographical locations and 19 countries
141 were analysed. The host plants were susceptible clones of Cavendish banana and plantain
142 cultivars. The samples from Central and South America (Mexico, Honduras, Costa Rica,
143 Panama, Columbia and Venezuela) had been previously collected between 1991 and 2009 for
144 the purposes of the study of Robert et al. (2012). All of the other Caribbean samples were
145 collected between 2009 and 2014 for this study. Overall, the samples first included 55
146 populations from 19 countries. The sampling effort was mainly focused on the Caribbean
147 islands (46 populations), including three populations collected from three different locations

per island, and even more were collected from two recently contaminated islands in the French West Indies (14 populations in Martinique and 6 in Guadeloupe). The populations consisted of around 18 to 60 fungal isolates collected simultaneously from different plants in the same banana cropfield. More intensive sampling was carried out in Martinique and Guadeloupe for the purpose of a landscape genetic analysis. About 397 isolates were collected from 27 locations in Guadeloupe and around 680 isolates were collected from 58 locations in Martinique (3 to 17 isolates per location) based on a two-dimensional sampling scheme designed to cover the two islands entirely (Supplementary Dataset 1).

Most of the populations were composed of fungal isolates grown in the laboratory. In this case, each individual corresponded to a strain derived from a single spore (sexual or asexual spores, depending on the case) isolated from an infected banana leaf. Growing mycelium resulting from spore germination was placed on solid medium (300 mL V8, 3 g $CaCO_3$, 20 g/L agar, pH 6) for 2 weeks at 25°C, before being dried for 2 days at 55°C and ground. DNA was extracted from growing mycelium as described in (Halkett et al., 2010). For some populations, young lesions ("leaf streak") were cut from banana leaves, dried for a day at 55°C and stored at –20°C. Each lesion was considered to have been caused by a single fungal strain. DNA was directly extracted from the sampled lesions using the protocol described in Robert et al. (2010).

2.3. Microsatellite marker genotyping

All individuals were genotyped using 16 microsatellite markers. These markers had already been used in other studies (Robert et al., 2010; Zapater et al., 2008). They were combined in three multiplex panels of four markers for the first panel, and six markers for the other two panels (Table S1, Supplementary file 1). PCR amplification and genotyping were carried out as described in (Robert et al., 2010). The genotype was obtained for at least 10 markers for all of the studied isolates, which was largely enough for genetic assignment (Arthofer et al., 2018).

176 *2.4. Linkage disequilibrium and genetic diversity analysis*

177 Multi-locus gametic linkage equilibrium was estimated using the standardized index of
178 association $\overline{r}_d$ (Agapow and Burt, 2001) and tested with the *poppr* R-package (Kamvar et al.,
179 2014). Gametic linkage disequilibrium between all pairs of loci was also tested in all
180 populations using Fisher's exact tests implemented in the *genepop* R-package with default
181 parameters (Rousset, 2008), and the false discovery rate (FDR) procedure implemented in the
182 *qvalue* R-package version 2.12.0 (http://github.com/jdstorey/qvalue).

183 Genetic variation within the 55 populations with more than 18 isolates was estimated
184 according to the number of polymorphic markers (out of 16), the number of multi-locus
185 genotypes (MLG), and by 3 descriptive indices averaged across markers: gene diversity index
186 (equivalent to the mean expected heterozygosity He (Nei, 1987), Simpson's index λ (Hill,
187 1973) computed using the *poppr* R-package (Kamvar et al., 2014) and the mean allelic
188 richness (Ar), calculated using the rarefaction method implemented in the *hierfstat* R-package
189 (Goudet, 2005). The cumulative geographic distance from the Honduras population HND3
190 (close to the site where the epidemic was assumed to have started in the Americas) was
191 calculated for each population following a point-by-point pathway northward and southward.
192 The correlation between that geographic distance and Ar or He was tested using the lm
193 function of R 3.6.0 (http://www.R-project.org).

194 The BOTTLENECK program (Piry et al., 1999) was used to detect a potential signal of
195 recent population bottlenecks based on the detection of excess heterozygosity (in the sense of
196 Nei's (1987) gene diversity) compared to heterozygosity expected under the mutation-drift
197 equilibrium assumption, from the observed number of alleles. The expected heterozygosity
198 distribution was estimated based on 1000 replications assuming a two-phase mutation model
199 (TPM: 70% stepwise mutation model (SMM), 30% IAM)), as recommended for
200 microsatellite markers (Dirienzo et al., 1994). We then used the Wilcoxon significance test to
201 determine whether loci displayed significant excess heterozygosity.

202

203 *2.5. Population structure analyses*

An admixture analysis was conducted by estimating the individual ancestral coefficients with the sNMF algorithm implemented in the LEA R-package (Frichot and François, 2015). This approach is similar to Bayesian clustering programs, such as STRUCTURE (Pritchard et al., 2000). The K value representing the number of ancestral populations ranged from 1 to 10, and 10 independent runs were performed for each K value. The number of ancestral populations that best explained the genotypic data was chosen based on the cross-entropy criterion. This criterion evaluates the quality of fit of the statistical model to the data using a cross-validation technique (Frichot and François, 2015). The maximum local contribution to ancestry was spatially interpolated at each geographical map point. Barplots were drawn up by taking the ancestry coefficient estimated for individuals or averaged for individuals belonging to the 55 populations presented in Table 2. Principal components analysis (PCA) of these populations was performed with the dudi.pca function using the R software ade4 package (Dray and Dufour, 2007).

To further evaluate the population structure and the degree of differentiation on different scales, we used an analysis of molecular variance (AMOVA) framework, as implemented in the HierFstat R-package (Goudet, 2005). The significance of the F-statistics estimated on different geographical levels was tested using 1,000 permutations between units. The 95% confidence intervals of these F-statistics were computed by boostrapping over loci 1,000 times. Pairwise *F*st (Weir and Cockerham, 1984) were also estimated and tested (Fisher exact test) using the *Genepop* R-package. Genetic distances between populations were estimated using the Cavalli-Sforza chord distance and represented using a neighbour-joining tree, as implemented in POPULATIONS1.2.31 software (www.bioinformatics.org/~tryphon/populations/).

2.6. Approximate Bayesian Computation (ABC) model choice analyses

A model choice analysis was carried out using the ABC random forest (ABC-RF) methodology developed by Pudlo et al. (2016), to test scenarios of *P. fijiensis* introduction

into Martinique. An ABC model choice analysis involves a comparison of observed data with data simulated under the evolutionary scenarios to be compared. Like Fraimout et al. (2017), we carried out two analyses with two different sets of populations as representatives of their genetic cluster (e.g. the Honduras ancestral population, the northern group, the southern group and the two populations from Martinique; Table S2, Supplementary File 1).

For each of the three scenarios tested (Fig. S6, Supplementary File 1; see results for a description), we simulated 60,000 datasets (20,000 per scenario) using DIYABC v2.1.0 software (Cornuet et al., 2014). Parameters were drawn up from the prior distributions described in Table S3 (Supplementary File 1). Time was expressed as numbers of generations, assuming 10 successive generations per year (8). Historical data were used to define the bounds of the introduction events. A generalized stepwise mutation model as implemented in the software was used to simulate mutations at microsatellite markers. Observed and simulated datasets were summarized using all the one-sample and two-sample statistics for microsatellite markers, as well as the linear discriminant analysis (LDA) axes, calculated in DIYABC v2.1.0 (Cornuet et al., 2014). The quality of the simulated models (scenarios and priors) was checked using PCA on the space of the summary statistics, as implemented in DIYABC v2.1.0 (Cornuet et al., 2014) (Fig. S7, Supplementary File 1). After simulating the datasets, we carried out an ABC-RF model choice analysis using the abcrf R-package (Pudlo et al., 2016). An ABC-RF analysis provides a classification vote for competing scenarios where the number of times a scenario is selected represents the model best fitting the dataset from a forest of 900 trees, as well as the posterior probability of the best selected model. The consistency of the ABC-RF model choice analysis was ensured by replicating the simulations and ABC-RF analysis 10 times and calculating the prior error rate (Table S4, Supplementary File 1). The number of votes and the prior error rate for each scenario, as well as the posterior probability of the best selected model, were averaged over the 10 replicates. Preliminary analyses using different numbers of datasets per scenario were carried out to ensure that we used a sufficient number of datasets and of trees for the random forest analyses.

260

261 **3. Results**

262 Overall, 2,188 fungal individuals from 121 different geographical locations were
263 isolated (Table 1, Fig. 1A). The isolates were genotyped using 16 microsatellites. Historical
264 records were concurrently obtained from various locations (Table 1, Fig. 1A and
265 Supplementary Dataset 1). These data were used to conduct statistical and population genetic
266 analyses.

267

268 *3.1. Random mating reproduction mode in all invasive populations*

269 A within-population genetic analysis of 55 populations with more than 18 isolates
270 suggested the existence of sexual reproduction and random mating in all populations studied
271 (Table 2). Overall, 1,354 multi-locus genotypes (MLGs) were detected for 1,617 isolates.
272 Around 12.5% (170/1,354) of the total number of MLGs were repeated in the whole sample.
273 Of these repeated MLGs, 68.2% (116/170) were found only twice in the whole sample.
274 Around 37.6% (64/170) were repeated within and 62.4% (106/170) between populations.
275 Thus, most of the repeated MLGs were found only twice overall and were distributed in two
276 different populations. Around 38.2% (65/170) of the repeated MLGs included at least one
277 isolate which was cloned from ascospores derived from sexual reproduction. Taken together,
278 these observations did not enable us to conclude that isolates having the same MGLs were
279 clones. Due to the limited number of markers used, in most cases they may have
280 corresponded to MLGs that had been sampled twice by chance even in a sexual reproduction
281 setting. Regardless, the presence of repeated MLGs had little impact on the population
282 structure observed and the number of MLGs detected in each population was very close to the
283 total number of isolates sampled (Simpson's index > 0.93). Furthermore, in all the
284 populations, no pairs of loci were found to be in significant linkage disequilibrium using
285 Fisher's exact test. The multilocus association index r_d was significantly not null in 42%
286 (23/55) of the populations and 16% (9/55) of the populations, at 5% and 1% FDR,
287 respectively. However, r_d was significantly and negatively correlated with the gene (r=-0.46,

288 p=2.9 10^{-4}) and genotypic diversity index (r=-0.53, p=2.5 10^{-5}; Fig. S1, Supplementary File
289 1), suggesting that some linkage disequilibrium may have resulted from bottleneck effects, as
290 described in theoretical studies (Rogers, 2014).
291
292 *3.2. Two epidemic waves from Honduras with serial founder effects shaping the population*
293 *structure*
294 The chronology of the first disease occurrence dates in the different countries and
295 locations (Table 1, Fig. 1A) suggested that *P. fijiensis* spread via two separate epidemic
296 waves throughout the Central America-Caribbean Basin region. Starting from Honduras in
297 1972, a northern wave followed a route from Guatemala to Porto Rico (first detection in
298 2004), while a southern wave followed a route from Nicaragua to Guadeloupe (first detection
299 in 2012). In the Caribbean, significant linear regressions were obtained between the
300 occurrence times and the distances from the point considered as being the origin for each
301 wave (R^2=0.77 and 0.85 and p = 1.6e-05 and 1.9e-09, respectively, for the northern and
302 southern waves, Fig. 2). The speed of invasion was calculated as the inverse of the slope of
303 each linear regression and was very close for the two waves, with 6.8 km/month (81 km/year)
304 for the northern wave and 6.7 km/month (80 km/year) for the southern wave.
305
306 The population genetic structure of *P. fijiensis* in America was found to be in accordance with
307 historical data and showed that serial founder events accompanying the northern and southern
308 waves had led to the establishment of different genetic groups. The highest levels of allelic
309 richness (Ar) and Nei's gene diversity (He; (Nei, 1978) were detected in Honduras, where the
310 disease was first introduced in the area (Table 2). Significant signatures of bottlenecks on the
311 distribution of allele frequencies were observed in 34% (19/55) and 16% (9/55) of the
312 populations, at 5% and 1% FDR, respectively (Table 2). The number of microsatellite
313 markers and sample sizes used may have been too low to have sufficient statistical power for
314 tests on some samples (Peery et al., 2012). To further highlight the bottleneck effects, gene

315 diversity in each location, as estimated by the Ar or He parameter, was plotted against the
316 cumulative step-by-step geographical distance from Honduras following the migration waves
317 (Fig. 3). A significant negative correlation was detected when jointly considering the northern
318 and southern waves and genetic parameters. It was found that the population gene diversity
319 decreased with the distance from Honduras.
320 Individual ancestry coefficients were first estimated using the sNMF algorithm (Frichot and
321 François, 2015). The cross-entropy curve suggested a K value of 3 to 4, which could be
322 considered as representative of major ancestral populations (Fig. S2, Supplementary File 1).
323 The study populations were then considered to have derived from this ancestral gene pool and
324 the ancestry coefficients gave the proportions of individual genomes originating from the
325 different components of the pool. For a clear display of the results, barplots for the 55
326 populations with more than 18 isolates were drawn up from the mean values of the ancestry
327 coefficients with K ranging from 3 to 9 (Fig. S3, Supplementary File 1). The major population
328 structure in the study area was actually observed from a value of K=4 and a barplot of
329 individual ancestry coefficients for this value is also presented in Fig. S4 (Supplementary File
330 1). To highlight the population genetic structure, the maximum value of the ancestry
331 coefficients for each isolate was spatially interpolated (Fig. 1B), and a barplot of the mean
332 values of the ancestry coefficients for K=4 and for 55 populations with more than 18 isolates
333 is presented below (Fig. 1 C), as well as a neighbour-joining tree constructed from the genetic
334 distance (Fig. 1 D). These analyses highlighted the existence of three major genetic groups
335 corresponding to three geographical areas: a central group contained populations located in
336 Central America, Colombia and Jamaica; a second northern group contained populations
337 located in the northern Caribbean, from Cuba to Puerto Rico (there was some sub-structuring
338 in this group); and a third, fairly uniform, southern group contained populations from
339 Venezuela to the southern Caribbean islands to as far as Saint Lucia. PCA analysis revealed a
340 very close structure (Fig. S5, Supplementary File 1).
341 *F*-statistics were estimated at different hierarchical levels using AMOVA, while excluding
342 populations from Martinique, Dominica and Guadeloupe (Table 3). Three levels were

considered: populations, countries and groups corresponding to the one defined from the clustering analysis. Significant values were obtained at all the levels tested. However, a larger proportion of the genetic diversity was distributed between groups (F=0.23) rather than between countries within groups (F=0.05). Genetic differentiation between populations within groups and within countries was similar (F=0.16 and 0.12, respectively). These estimations seemed to indicate that the groups were genetically quite homogeneous and different from each other.

3.3. Long-distance introduction and admixture events

We found a different population structure on islands consecutively located north of Saint Lucia (Martinique, Dominica and Guadeloupe), revealing multiple introductions and admixture events (Fig. 1). Based on some preliminary results, dense two-dimensional sampling was carried out in Martinique and Guadeloupe (Table 1, Fig. 1A), which enabled us to map the introduced genetic groups. Some individuals and populations from both ends of Martinique were related to the southern genetic group, while other samples located in the centre of the island were related to the northern genetic group, thus suggesting that there had been at least two independent introductions on that island (Fig. 1B and 1C). Three populations from Martinique (MTQ2, MTQ3, MTQ13) were grouped on the middle branches of the tree presented in Fig. 1 D. Geographically, these samples were located at junctions between differentiated populations belonging to the northern and southern genetic groups and we suspected some admixtures between them. *F*-statistics between populations in Martinique belonging to different genetic groups were sometimes very high, with a maximum at $F = 0.51$ (Table 3), and the highest values were obtained when comparing populations from the ends of Martinique (MTQ1 and MTQ14) with all the others (Supplementary Dataset S1). The PCA projection of the populations from Martinique, Dominica and Guadeloupe were either close to one of the three major groups (southern, northern and central) or between these groups, thus supporting the idea that there have been multiple introductions and admixture events on these islands (Fig. S5, Supplementary File 1). Based on this descriptive analysis, three different

scenarios for *P. fijiensis* introduction into Martinique were tested using an ABC procedure. The first scenario assumed two independent introductions into Martinique from two populations belonging to the northern and southern migration waves (Fig. S6, Supplementary File 1). Two other scenarios assumed a single introduction event from one population belonging to the northern or southern migration wave, with the second population from Martinique emerging from the first one after its establishment. In all the scenarios, Honduras was considered as the ancestral population and all the founded populations came through a bottleneck when introduced. In the first scenario, we considered that genetic differentiation between Martinican populations was the result of the combination of bottleneck effects and introductions from different origins, whereas in the other two scenarios genetic differentiation resulted only from bottleneck effects. The former scenario was found to be the most probable using two different sample sets (mean posterior probability of 0.95 ± 0.008; Table S4, Supplementary File 1).

Populations from Dominica and Guadeloupe showed intermediate patterns for ancestry coefficients (Fig. 1 C), and were all located in the middle of the tree in Fig. 1D, except population GLP5, which was more related to the southern genetic group. This distinct population had the lowest genetic diversity level in Guadeloupe (Table 2) and a very high level of genetic differentiation between all the other populations from Guadeloupe ($0.27 < F < 0.41$, $p<0.0001$; Supplementary Dataset 1). Thus, population GLP5 may have stemmed from a rare intense bottleneck and/or an independent introduction from the southern wave. It was not included in further analysis so as to be able to highlight the general trends. The range of pairwise *F*-statistics between populations from Guadeloupe and Dominica was lower or very close to that obtained between populations within countries in the central, northern or southern genetic groups ($F<0.14$, Table 3 and Supplementary Dataset 1). Taken together, all of the above observations suggested that the populations in Guadeloupe and Dominica were founded from an admixture between the northern and southern genetic groups.

The correlations between genetic diversity (estimated by Ar and He) and the cumulative geographical distance from Honduras for samples from Martinique, Dominica and Guadeloupe were significantly positive (Fig. 3), thus indicating that genetic diversity had increased on those islands, in contrast with samples from other countries. The estimated values of both parameters for the furthest samples from Honduras located in Guadeloupe were close to those obtained from Honduran samples. These results strongly indicated that there had been multiple introductions and/or admixtures on these islands.

4. Discussion

This study set out to gain insight into epidemiological and demographic processes that have accompanied the spread of the *P. fijiensis* fungus which causes black leaf streak disease of banana in the latest invaded region, i.e. Central America and the Caribbean. To that end, the population history of *P. fijiensis* in this region was retraced based on an analysis of successive recorded outbreaks and on the population genetic structure. Taken together, the findings of these analyses revealed that two different epidemic waves of *P. fijiensis* in the northern and southern Central America-Caribbean Basin region, starting from Honduras, shaped the *P. fijiensis* population structure through serial founder events. This led to the formation of relatively homogenous genetic groups in the northern and southern Caribbean region, with lower genetic diversity compared to Central America and with genetic differentiation between them. The situation was unique in Martinique, Dominica and Guadeloupe. The results suggested independent introductions of *P. fijiensis* in Martinique from both the northern and southern origins and an admixture between these introductions accompanying the consecutive spread of the disease in Dominica and Guadeloupe.

In contrast with what is frequently observed in fungi (Gladieux et al., 2015), the genotypic diversity and linkage disequilibirum analysis showed that sexual reproduction was not lost in *P. fijiensis* during the invasion in the study region, above all following the serial founder events that occurred in the Caribbean region. Thus, all the populations probably kept the

ability to produce wind-dispersed ascospores, which have been shown to disperse over long distances (Rieux et al., 2014). Ascospores might be essential for *P. fijiensis* to colonize new areas. Furthermore, the correlation between the occurrence time and the distance from a geographical point considered as the origin in each wave, and the close values obtained for the invasion speed between the two waves (about 80 km/year), suggested a similar and regular step-by-step migration trend. This speed was also close to that estimated in another study in a continental context (about 90 km/year) and using a modelling approach accounting for both sexual and asexual reproduction (about 100 km/year but with mate limitation; (Hamelin et al., 2016)). Lastly, the existence of homogeneous genetic groups within each wave was also in favour of regular step-by-step migration. Taken together, all the data obtained in this study suggested that *P. fijiensis* spread throughout the Caribbean mainly via windborne spore dispersal. Although insular, the Caribbean geography formed by around 30 islands distributed along arcs and separated by 50 to 200 km might have made this spreading pattern possible. However, we cannot rule out that there may have been some impacts due to human activities, especially with regard to island-to-island dissemination. Indeed, for human activities as well as windborne spore dispersal, the shorter the distance between islands, the more likely exchanges were and both means of dispersal might have had a similar impact on the pathogen population structure.

The latest islands to be invaded (Martinique, Dominica and Guadeloupe) appeared to have been as a special case and the existence of two differentiated genetic groups between the northern and southern Caribbean led us to detect multiple introductions. What historical scenario could explain the genetic pattern observed? Through surveillance networks (Teycheney et al., 2009), we found that Dominica and Guadeloupe—located north of Martinique—were disease-free when BLSD was first detected in Martinique in 2010 (Table 1). Evidence of the introduction of BLSD in Martinique was the detection of an admixture between the northern and southern genetic groups in Dominica and Guadeloupe, while these groups were not found to be admixed in different locations in Martinique. Populations

belonging to the southern genetic group might have been introduced into Martinique from Saint Lucia by windborne spore dispersal. However, an introduction into Martinique from the northern wave via wind would imply that the disease had first jumped to Guadeloupe and Dominica and then returned. This movement pattern would be contrary to the prevailing winds (east to west in the northern Caribbean) and hurricane trajectories (Andraca-Gómez et al., 2014; Varlas et al., 2018). Although the existence of rare and stochastic long-distance disease dispersal events cannot be ruled out, another likely introduction scenario for these islands that were invaded last might be: windborne introductions into Martinique from the south (via the south to north prevailing winds) or via human activities from the north. The disease could then have been introduced into Dominica and Guadeloupe from Martinique, likely by wind dispersal, accompanied by an admixture of the two introduced genetic groups.

The *P. fijiensis* fungus may have been at least partly introduced into Martinique by humans despite the surveillance network operating in the Caribbean (Teycheney et al., 2009). Such events accompanying plant pathogen colonization on local or regional scales are difficult to detect and could be more frequent than imagined. Furthermore, the surveillance systems are not considered to be effective enough to avoid human-mediated introduction of crop diseases (Carvajal-Yepes et al., 2019). Only an estimated 2 to 6% of all cargo entering a country can be effectively controlled (Carvajal-Yepes et al., 2019) and there is a lot of uncontrolled boat traffic in the Caribbean, especially small private boats illegally transporting vegetables between neighbouring islands. Our results supported this view and highlighted the need to reinforce surveillance to cope with other emerging crop diseases that have yet to be spread worldwide. The surveillance system in Australia has been efficient enough to eradicate BLSD eight times with the application of strict and intensive measures (Guzman et al., 2019) requiring full collaboration of many actors (from growers to the Ministry of Agriculture). On a global scale, strong international collaboration between such actors is key for preventing and excluding emerging plant diseases. Furthermore, reinforcing surveillance could have more evident impacts on the invasion patterns of pathogens mainly spread through human

activities. In the case of banana, this could be crucial for limiting the spread of the very destructive soilborne *Fusarium* wilt tropical race 4 (*Fusarium oxysporum* fsp cubense TR4) fungus as a first outbreak in the Americas was suspected very recently in Colombia (Stokstad, 2019), and humans are considered to be the principal vector of dispersal (Dita et al., 2018; Qazi, 2015). The same reasoning could be applied to the banana bunchy top virus, which is not present in the Americas and is thought to be spread over long distances through infected plant material (Jacobsen et al., 2019).

Even after a serial founder effect during its expansion in Central and South America and the Caribbean, all of the *P. fijiensis* populations studied had maintained a relatively high level of genetic diversity and sexual reproduction. It has been suggested that *P. fijiensis* is among the plant pathogens with the highest evolutionary potential (McDonald and Linde, 2002) and this feature might remain true even in recently invaded regions. This assumption is supported by the erosion of quantitative resistance in newly disseminated hybrids on two Caribbean islands (Cuba and Dominican Republic), although local adaptation of *P. fijiensis* populations was demonstrated recently in cross-inoculation experiments (Dumartinet et al., 2019). Furthermore, admixtures between different genetic backgrounds, as observed in Dominica and Guadeloupe, increase the genetic diversity upon which selection can act (Lavergne and Molofsky, 2007) and produce novel genotypes that may facilitate adaptation to novel habitats (Rius and Darling, 2014), corresponding here to novel resistant varieties. This hypothesis could be tested by comparing the pathogenicity of non-admixed or admixed *P. fijiensis* genotypes on different hosts through cross-inoculation experiments. Furthermore, it can already be stressed that an effective surveillance system could be an integral part of sustainable strategies for plant resistance deployment, which can be seen as the most promising way of controlling emerging plant diseases (Mundt, 2014). Indeed, a second outcome of such a system might be to restrain gene flows between pathogen populations, which might affect their evolutionary potential.

5. Conclusions

The present study of the population history of the *P. fijiensis* fungus causing the latest major epidemic on bananas in the America-Caribbean area provided fresh insight into the epidemiological and demographic processes involved in this emerging plant disease. The results suggested that the fungus may have spread mainly through spores in this new area, which encompasses a large insular zone. However, probable human involvement in the spread of the disease was suspected on some islands. The evolutionary potential of new invasive *P. fijiensis* populations may remain high after serial founder events and could even increase following admixture, even on some islands. This study also highlighted that surveillance and quarantine measures need to be strengthened to cope with other emerging diseases of banana and potential gene flow between established pathogen populations. Possible pathogen adaptation following the release of new resistant hybrids should also be monitored in order to adjust resistance deployment strategies and hamper pathogen adaptation.

CRediT authorship contribution statement

Jean Carlier: Funding acquisition, Project administration, Conceptualization, Data curation, Formal analysis, Investigation, Writing - original draft. **Stéphanie Robert:** Conceptualization, Data curation, Formal analysis, Investigation, Writing - original draft. **Véronique Roussel** : Investigation, Data curation, Methodology, Resources. **Yolande Chilin-Charles[1,3], Nadia Lubin-Adjanoh** : Methodology, Resources. **Aude Gilabert :** Formal analysis, Investigation, Writing - review & editing. **Catherine Abadie:** Funding acquisition, Project administration, Formal analysis, Investigation, Writing - review & editing.

Declaration of competing interest

The authors declare that there is no conflict of interest.

Acknowledgments

We thank all of the regional partners who sent either leaf samples or historical data: L.Perez Vicente from INISAV in Cuba, R. Anselm from the Ministry of Agriculture in Dominica, E. Prophète from the Ministry of Agriculture in Haiti, Debbie Henry and Janet Conie from the

Ministry of Agriculture in Jamaica, K. Lombion, from the Ministry of Agriculture in Guadeloupe, T. Peter from the Ministry of Agriculture in Grenada, E. Husson from SICA-TG in Martinique, J. Iotti, from the Ministry of Agriculture in Martinique, M. Cortes from the Ministry of Agriculture in Porto Rico, L. Minier and J. Clase Salas, Secretary of Agriculture, Dominican Republic, L. Alexander from the Ministry of Agriculture in Saint Lucia, M. Richards from the Ministry of Agriculture in Saint Vincent, M. Fortune from the Ministry of Agriculture in Trinidad, M. Rivera from FHIA in Honduras, G. Rivas from CATIE in Costa Rica, and R. Ploetz from the University of Florida, USA. We would also like to thank Virginie Ravigné (CIRAD) for helpful discussions, Peter Biggins (CIRAD) and David Manley (freelance scientific translator) for their careful checking of the English.

Funding

S. R.'s PhD was supported by a doctoral fellowship from CIRAD and the French Occitanie region, and A. G. by a post-doctoral fellowship from *Agropolis Fondation* (grant number 1504-004-E-Space). Experiments and sampling were supported by the Emerfundis-ANR project (grant number 07-BDIV-003), the Interreg IV Caribbean Programme (grant number 31409-Cabaré), the *Prefecture de Guadeloupe* (grant number FCR 2010/38), European Regional Development Fund and the *Conseil Régional de Guadeloupe* through the MALIN project.

Data availability

All the data are available in the manuscript or in the supplementary data.

Appendix A. Supplementary materials

Supplementary Dataset 1

Supplementary File 1

References

Agapow, P. M., Burt, A., 2001. Indices of multilocus linkage disequilibrium. Molecular Ecology Notes. 1**,** 101-102.

Almeida, F., et al., 2019. The Still Underestimated Problem of Fungal Diseases Worldwide. Front Microbiol. 10**,** 214.

Anderson, P. K., et al., 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in Ecology & Evolution. 19**,** 535-544.

Andraca-Gómez, G., et al., 2014. A potential invasion route of Cactoblastis cactorum within the Caribbean region matches historical hurricane trajectories. Biological Invasions. 17**,** 1397-1406.

Arthofer, W., et al., 2018. Identifying the minimum number of microsatellite loci needed to assess population genetic structure: A case study in fly culturing. Fly (Austin). 12**,** 13-22.

Carvajal-Yepes, M., et al., 2019. A global surveillance system for crop diseases. Science. 364**,** 1237-1239.

Churchill, A. C. L., 2011. Mycosphaerella fijiensis, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. Molecular plant pathology. 12**,** 307-328.

Cornuet, J. M., et al., 2014. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 30**,** 1187-1189.

de Lapeyre de Bellaire, L., et al., 2010. Black Leaf Streak Disease is challenging the banana industry. Fruits. 65**,** 327-342.

Dirienzo, A., et al., 1994. MUTATIONAL PROCESSES OF SIMPLE-SEQUENCE REPEAT LOCI IN HUMAN-POPULATIONS. Proceedings of the National Academy of Sciences of the United States of America. 91**,** 3166-3170.

Dita, M., et al., 2018. Fusarium Wilt of Banana: Current Knowledge on Epidemiology and Research Needs Toward Sustainable Disease Management. Front Plant Sci. 9**,** 1468.

Dray, S., Dufour, A. B., 2007. The ade4 package: Implementing the duality diagram for ecologists. Journal of Statistical Software. 22**,** 1-20.

Dumartinet, T., et al., 2019. Pattern of local adaptation to quantitative host resistance in a major pathogen of a perennial crop. Evolutionary Applications.

Evans, A., 2016. The Speed of Invasion: Rates of Spread for Thirteen Exotic Forest Insects and Diseases. Forests. 7.

Fraimout, A., et al., 2017. Deciphering the Routes of invasion of Drosophila suzukii by Means of ABC Random Forest. Mol Biol Evol. 34**,** 980-996.

Frichot, E., François, O., 2015. LEA: AnRpackage for landscape and ecological association studies. Methods in Ecology and Evolution. 6**,** 925-929.

Gladieux, P., et al., 2015. The population biology of fungal invasions. Mol Ecol. 24**,** 1969-86.

Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes. 5**,** 184-186.

Guzman, M., et al., Black leaf streak. In: D. R. Jones, (Ed.), Handbook of diseases of banana, abaca and enset. CABI, Wallingford, 2019, pp. 41-115.

Halkett, F., et al., 2010. Genetic discontinuities and disequilibria in recently established populations of the plant pathogenic fungus Mycosphaerella fijiensis. Mol Ecol. 19**,** 3909-23.

Hamelin, F. M., et al., 2016. Mate Finding, Sexual Spore Production, and the Spread of Fungal Plant Parasites. Bull Math Biol. 78**,** 695-712.

Hill, M. O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology. 54**,** 427-432.

Ioos, R., et al., 2011. First Report of Black Sigatoka Disease in Banana Caused by Mycosphaerella fijiensis on Martinique Island. Plant Disease. 95**,** 359-359.

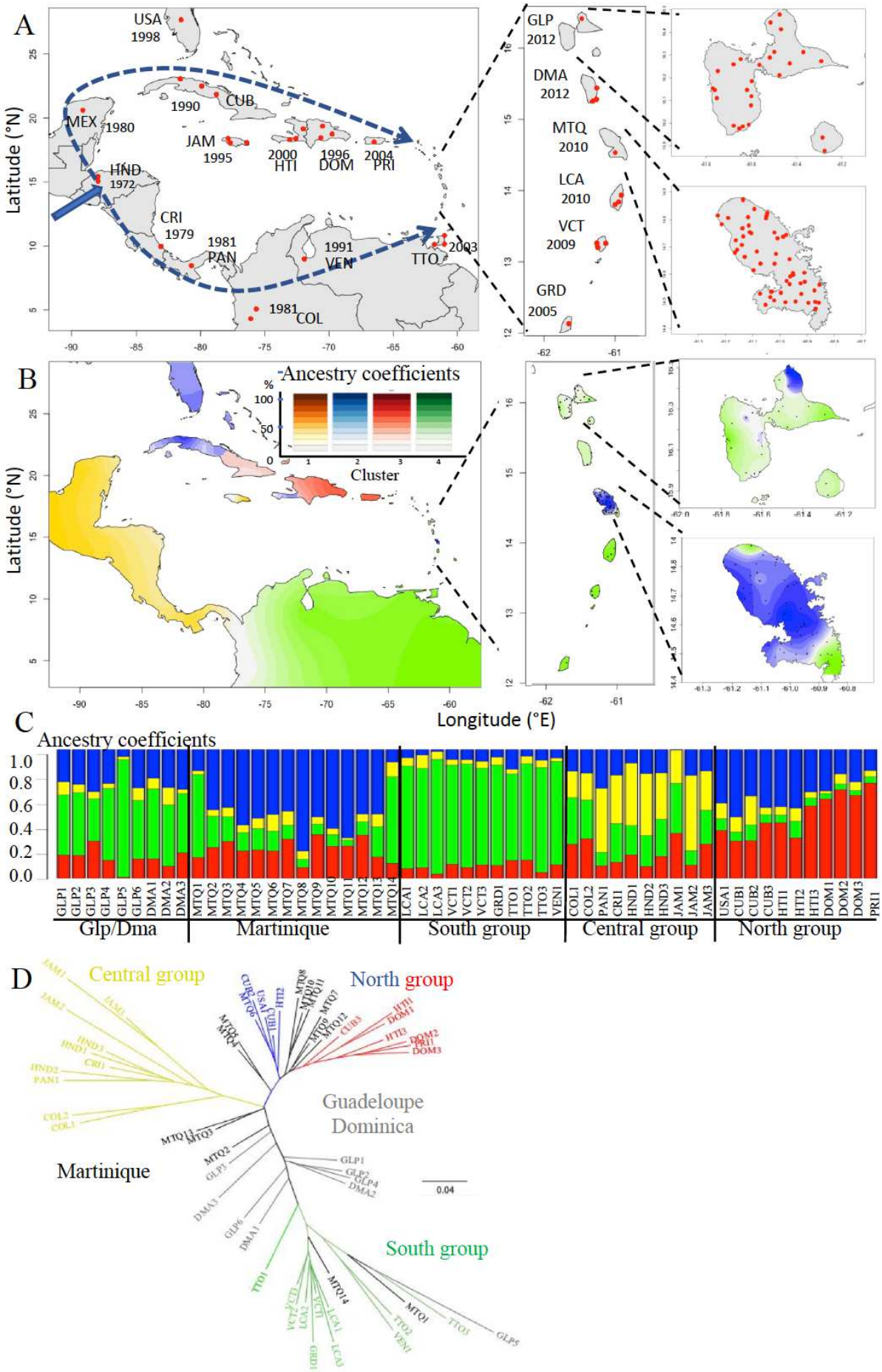
Jacobsen, K., et al., 2019. Seed degeneration of banana planting materials: strategies for improved farmer access to healthy seed. Plant Pathology. 68**,** 207-228.

Jones, D. R., 2019. Handbook of diseases of banana, abaca and enset. CABI, Wallingford.

Kamvar, Z. N., et al., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2**,** e281.

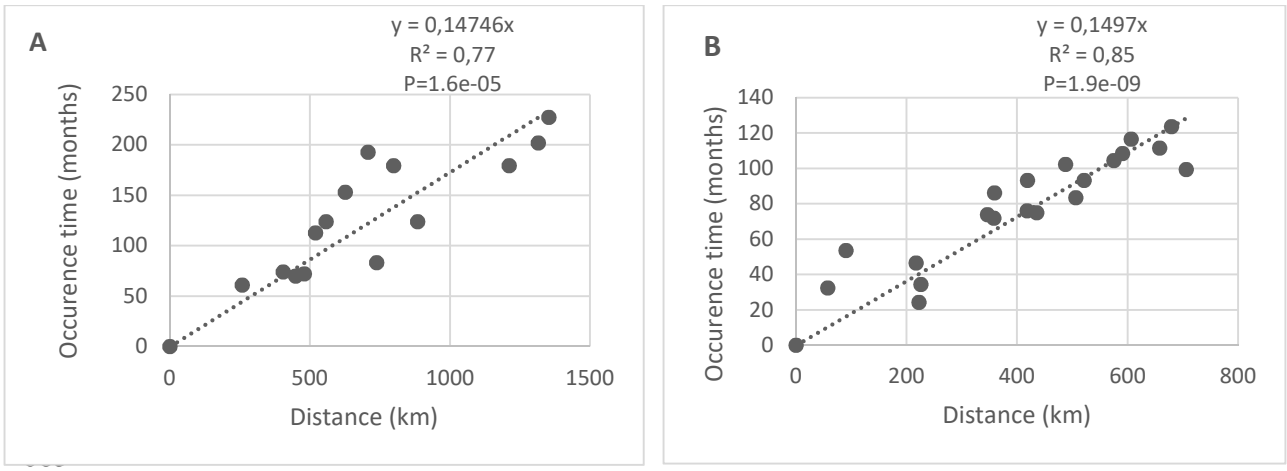
Lavergne, S., Molofsky, J., 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc Natl Acad Sci U S A. 104, 3883-8.
McDonald, B. A., Linde, C., 2002. Pathogen population genetics, evolution potential, and durable resistance. Ann. Rev. Phytopathol. 40, 349-379.
McDonald, B. A., Stukenbrock, E. H., 2016. Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. Philos Trans R Soc Lond B Biol Sci. 371.
Mundt, C. C., 2014. Durable resistance: a key to sustainable management of pathogens and pests. Infect Genet Evol. 27, 446-55.
Nei, M., 1978. Estimation of average heterozygosity and genetic distances from a small number of individuals. Genetics. 89, 583-590.
Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, NY.
Parnell, S., et al., 2017. Surveillance to Inform Control of Emerging Plant Diseases: An Epidemiological Perspective. Annual Review of Phytopathology, Vol 55. 55, 591-610.
Peery, M. Z., et al., 2012. Reliability of genetic bottleneck tests for detecting recent population declines. Molecular Ecology. 21, 3403-3418.
Pennisi, E., 2010. Armed and Dangerous. Science. 327, 804-805.
Piry, S., et al., 1999. Bottleneck: A computer program for detecting recent reductions in the effective population size using allele frequency data. The Journal of Heredity. 90, 502-503.
Pritchard, J. K., et al., 2000. Inference of population structure using multilocus genotype data. Genetics. 155, 945-959.
Pudlo, P., et al., 2016. Reliable ABC model choice via random forests. Bioinformatics. 32, 859-66.
Qazi, J., 2015. Banana bunchy top virus and the bunchy top disease. Journal of General Plant Pathology. 82, 2-11.
Rieux, A., et al., 2014. Long-distance wind-dispersal of spores in a fungal plant pathogen: estimation of anisotropic dispersal kernels from an extensive field experiment. PLoS One. 9, e103225.
Rius, M., Darling, J. A., 2014. How important is intraspecific genetic admixture to the success of colonising populations? Trends Ecol Evol. 29, 233-42.
Robert, S., et al., 2012. Contrasting introduction scenarios among continents in the worldwide invasion of the banana fungal pathogen Mycosphaerella fijiensis. Mol Ecol. 21, 1098-114.
Robert, S., et al., 2010. Optimized genotyping with microsatellite markers in the fungal banana pathogen Mycosphaerella fijiensis (Mycosphaerellaceae). Am J Bot. 97, e130-2.
Robert, S., et al., 2015. MULTIPLE INTRODUCTIONS AND ADMIXTURE AT THE ORIGIN OF THE CONTINENTAL SPREAD OF THE FUNGAL BANANA PATHOGEN MYCOSPHAERELLA FIJIENSIS IN CENTRAL AMERICA: A STATISTICAL TEST USING APPROXIMATE BAYESIAN COMPUTATION. Revue D Ecologie-La Terre Et La Vie. 70, 127-138.
Rogers, A. R., 2014. How population growth affects linkage disequilibrium. Genetics. 197, 1329-41.
Rousset, F., 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour. 8, 103-6.
Stokstad, E., 2019. Banana fungus puts Latin America on alert. Science. 365, 207-208.
Stukenbrock, E. H., McDonald, B. A., 2008. The origins of plant pathogens in agro-ecosystems. Annual Review of Phytopathology. 46, 75-100.
Teycheney, P.-Y., et al., Development of plant epidemiological surveillance networks, data exchanges and joint response strategies in the Caribbean: the french experience. In: G. B. R. Cfcs, (Ed.), Annual Meeting of the Caribbean Food Crops Society. s.n., Basseterre, Saint-Kitts-et-Nevis, 2009, pp. 8 p.
Varlas, G., et al., 2018. An analysis of the synoptic and dynamical characteristics of hurricane Sandy (2012). Meteorology and Atmospheric Physics. 131, 443-453.
Weir, B. S., Cockerham, C. C., 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38, 1358-1370.
Zapater, M. F., et al., 2008. Microsatellite markers for the fungal banana pathogens Mycosphaerella fijiensis, Mycosphaerella musicola and Mycosphaerella eumusae. Molecular Ecology Resources. 8, 1121-1125.

671 Tables and Figures:



672

673 **Fig. 1.** Invasion and population genetic structure of *Pseudocercospora fijiensis* in the
674 Americas. A, geographical locations of the samples analysed and year of first official
675 detection of BLSD in the locations; B, interpolation of sNMF results with K=4 of the most
676 represented cluster in all of the 2,188 isolates analysed; C, barplots with K=4 of the mean
677 ancestry coefficients for the 55 samples with more than 18 isolates; D, unrooted neighbour-
678 joining tree based on Cavalli-Sforza distances between the same 55 samples. Abbreviations:
679 GLP = Guadeloupe, DMA = Dominica, MTQ = Martinique, LCA = Saint Lucia, VCT = Saint
680 Vincent, GRD = Grenada, TTO = Trinidad, VEN = Venezuela, COL = Colombia, PAN =
681 Panama, CRI= Costa Rica, HND= Honduras, MEX= Mexico, USA= United States of
682 America, CUB=Cuba, JAM=Jamaica, HTI=Haiti, DOM= Dominican Republic, PRI= Puerto
683 Rico.
684



686 **Fig. 2.** Linear regression between occurrence times in Caribbean locations and Euclidian
687 distances between these locations and a geographical point considered as the origin of the
688 northern (A) or southern (B) migration front of *P. fijiensis* in the Caribbean.
689

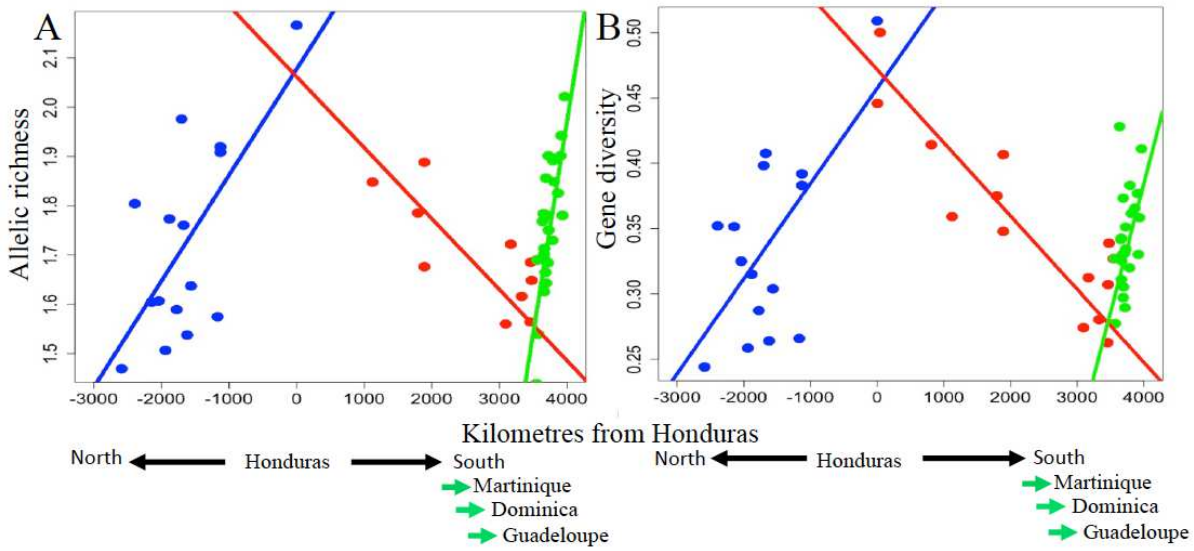


Fig. 3. Linear regression between cumulative step-by-step distance from Honduras with allelic richness (A) or (B) Nei's gene diversity calculated for the 55 *P. fijiensis* samples with more than 18 isolates. Correlation tests were: (A), northern direction (r=-0.67, p=6 10^{-3}), southern direction (r=-0.86, p=2.9 10^{-6}), Martinique to Guadeloupe (r=-0.65, p=4.6 10^{-4}); (B), northern direction (r=-0.63, p=0.011), southern direction (r=-0.89, p=6.4 10^{-7}), Martinique to Guadeloupe (r=-0.44, p=0.02).

Table 1 Summary of the historical data and *Pseudocercospora fijiensis* samples collected.

Country/Place	First official detection (month/year)	Sampling (Number)			
		Locations	Isolates	Collected years	Reference
France/Guadeloupe	01/2012	27	397		This study
Dominica	06/2012	3	91		This study
France/Martinique	09/2010	58	680		This study
Saint Lucia	01/2010	3	78		This study
Saint Vincent	10/2009	3	94		This study
Grenada	11/2005	1	25		This study
Trinidad	11/2003	3	68		This study
Venezuela	1991*	1	26		Robert et al., 2012
Colombia	1981*	2	46		Robert et al., 2012
Panama	1981*	1	31		Robert et al., 2012
Costa Rica	1979*	1	30		Robert et al., 2012
Honduras	1972*	3	80		Robert et al., 2012
Mexico	1980*	1	11		Robert et al., 2012
USA/Florida	1998*	1	60		This study
Cuba	11/1989	3	138		This study
Jamaica	08/1995	3	95		This study
Haiti	01/2000	3	85		This study
Dominican Rep.	09/1996	3	127		This study
USA/Puerto Rico	08/2004	1	26		This study
Total		121	2188		

*According to Jones (Jones, 2019)

Table 2 Genetic statistics calculated in *Pseudocercospora fijiensis* populations.

Country	Code	Number of isolates	Number of MLG[1]	P[2]	Ar[3]	λ[4]	He[5]	Test of He excess[6]	$\overline{r}_d$[7]
France/ Guadeloupe	GLP1	32	31	93.8	2.02	0.97	0.41	**	0.00
	GLP2	20	17	87.5	1.94	0.94	0.33	NS[8]	0.02
	GLP3	25	24	87.5	1.90	0.96	0.38	**	0.02

	GLP4	38	37	87.5	1.78	0.97	0.36	NS	0.01
	GLP5	18	16	31.3	1.28	0.93	0.21	NS	0.08**
	GLP6	32	29	81.3	1.83	0.96	0.37	**	0.00
Dominica	DMA1	37	37	81.3	1.85	0.97	0.36	**	0.01
	DMA2	26	25	87.5	1.89	0.96	0.38	*	0.01
	DMA3	28	26	81.3	1.73	0.96	0.32	NS	0.03*
France/ Martinique	MTQ1	23	17	43.8	1.36	0.93	0.20	NS	0.10**
	MTQ2	21	20	81.3	1.75	0.95	0.33	NS	0.05**
	MTQ3	29	28	93.8	1.90	0.96	0.35	NS	0.00
	MTQ4	36	29	75.0	1.68	0.96	0.29	NS	0.04**
	MTQ5	27	26	75.0	1.78	0.96	0.33	*	0.02
	MTQ6	18	18	87.5	1.86	0.94	0.37	NS	-0.01
	MTQ7	19	19	68.8	1.69	0.95	0.31	*	-0.02
	MTQ8	32	30	68.8	1.64	0.96	0.30	*	0.01
	MTQ9	18	18	68.8	1.66	0.94	0.33	NS	0.00
	MTQ10	47	41	87.5	1.71	0.97	0.33	NS	0.01
	MTQ11	20	19	81.3	1.70	0.95	0.31	NS	0.04*
	MTQ12	25	23	68.8	1.63	0.95	0.34	NS	0.05**
	MTQ13	19	17	81.3	1.78	0.93	0.34	NS	0.04*
	MTQ14	18	17	81.3	1.77	0.94	0.43	NS	0.06**
Saint Lucia	LCA1	23	22	62.5	1.54	0.95	0.28	NS	0.04*
	LCA2	31	31	68.8	1.69	0.97	0.33	*	0.02
	LCA3	24	20	56.3	1.44	0.94	0.22	NS	0.07**
Saint Vincent	VCT1	46	44	68.8	1.65	0.98	0.34	NS	0.02*
	VCT2	24	24	68.8	1.56	0.96	0.26	NS	0.01
	VCT3	24	24	68.8	1.68	0.96	0.31	**	-0.01
Grenada	GRD1	25	25	62.5	1.62	0.96	0.28	NS	0.01
Trinidad	TTO1	20	20	75.0	1.72	0.95	0.31	NS	0.00
	TTO2	26	23	56.3	1.56	0.95	0.27	NS	0.01
	TTO3	22	17	37.5	1.32	0.92	0.23	NS	0.12**

Venezuela	VEN1	26	26	62.5	1.68	0.96	0.35	NS	0.05**
Colombia	COL1	23	23	75.0	1.79	0.96	0.38	NS	0.01
	COL2	23	23	81.3	1.89	0.96	0.41	NS	0.00
Panama	PAN1	31	31	93.8	1.85	0.97	0.36	NS	0.00
Costa Rica	CRI1	30	30	100.0	1.91	0.97	0.41	NS	0.03*
Honduras	HND1	30	30	100.0	2.11	0.97	0.50	**	0.01
	HND2	23	22	87.5	2.01	0.95	0.45	**	0.01
	HND3	27	27	100.0	2.17	0.96	0.51	**	0.01
USA	USA1	60	60	87.5	1.80	0.98	0.35	**	0.00
Cuba	CUB1	46	43	81.3	1.77	0.98	0.32	*	0.00
	CUB2	40	40	87.5	1.98	0.98	0.40	**	0.01
	CUB3	52	52	62.5	1.64	0.98	0.30	*	0.01*
Jamaica	JAM1	32	27	87.5	1.57	0.96	0.27	NS	0.06**
	JAM2	30	30	93.8	1.91	0.97	0.38	NS	0.03*
	JAM3	33	33	87.5	1.92	0.97	0.39	NS	0.03**
Haiti	HTI1	25	25	75.0	1.54	0.96	0.26	NS	-0.01
	HTI2	28	28	75.0	1.76	0.96	0.41	*	0.01
	HTI3	32	27	68.8	1.59	0.96	0.29	*	0.01
Dominican	DOM1	45	41	68.8	1.51	0.97	0.26	NS	0.04**
Republic	DOM2	46	46	68.8	1.61	0.98	0.33	NS	0.04**
	DOM3	36	34	68.8	1.60	0.97	0.35	NS	0.02*
USA/Puerto Rico	PRI1	26	22	50.0	1.47	0.95	0.24	NS	0.07**

[1] Number of multi-locus genotypes (MLG), [2] Percentage of polymorphic loci, [3] Allelic richness, [4] Simpson's index (Hill, 1973), [5] Nei's gene diversity (Nei, 1978), [6] Wilcoxon test implemented in Bottleneck software (Piry et al., 1999), [7] Standardized index of association (Agapow and Burt, 2001),[8] Not significant, * Significant at a false discovery rate (FDR) of 5%, ** Significant at a false discovery rate (FDR) of 1%.

Table 3 Genetic differentiation between *Pseudocercospora fijiensis* populations.

Comparisons	*F*-statistic[1]	CI-95%[2]	Test
Global analysis: AMOVA[3]			
Between populations	0.36	0.30/0.43	Perm[4]***

Between populations within groups[5]	0.16	0.14/0.19	Perm***
Between populations within countries	0.12	0.06/0.15	Perm***
Between countries	0.27	0.22/0.33	Perm***
Between countries within groups	0.05	0.01/0.10	Perm***
Between groups	0.23	0.18/0.32	Perm***
Guadeloupe/Dominica: pairwise estimate[6]			
Min	0.00	-0.025/0.01	Fet[7] NS[8]
Max	0.14	0.06/0.21	Fet***
Martinique: pairwise estimate			
Min	0.00	-0.03/0.00	Fet NS
Max	0.51	0.31/0.68	Fet***
MTQ1 vs others	0.31-0.51		Fet***
MTQ14 vs others	0.21-0.39		Fet***
MTQ1 vs MTQ14	0.21		Fet***
Combinations excluding MTQ1 and MTQ14	0.00-0.11		Fet NS to Fet***

[1] Hierarchical *F*-statistics from the AMOVA or pairwise estimate of Weir and Cockerham (1984), [2] 95% confidence interval, [3] AMOVA carried out without samples from Guadeloupe, Martinique, Dominica, [4] Test using permutations, [5] Groups corresponding to those defined in figure 2C: central group (Colombia, Panama, Costa Rica, Honduras, Jamaica), northern group (Florida, Cuba, Haiti, Dominican Republic, Porto Rico), southern group (Saint Lucia, Saint Vincent, Grenada, Trinidad, Venezuela), [6] range of *F*-statistics estimated between population pairs. Estimates for all population pairs are given in Data S1. [7] Fisher exact test between population pairs, [8] Not significant, *** $p < 0.0001$.