Plant Pathology (2016) **65**, 196–204 Doi: 10.1111/ppa.12411

Genetic diversity and structure of *Hemileia vastatrix* populations on *Coffea* spp.

P. G. C. Cabral[a], E. Maciel-Zambolim[a], S. A. S. Oliveira[b], E. T. Caixeta[c] and L. Zambolim[a]*

[a]*Department of Phytopathology, Universidade Federal de Viçosa, Viçosa, Minas Gerais;* [b]*Embrapa Mandioca e Fruticultura, Cruz das Almas, Bahia; and* [c]*Embrapa Café, Brasília, Distrito Federal, Brazil*

Coffee leaf rust is the most limiting disease for coffee cultivation in Brazil. Despite its importance, relatively little is known about the genetic diversity of *Hemileia vastatrix*, the rust causal agent. In this work, the DNA from 112 monopustule isolates from different geographic locations and coffee genotypes were analysed by amplified fragment length polymorphisms (AFLP). The objectives were to assess the influence of the host and geographic origin on the diversity and population differentiation in *H. vastatrix*. The fungal population showed a low level of genotypic diversity. Gene diversity (h) was 0·027 and the hypothesis of random mating in the total population was rejected, but evidence for recombination was found for two subpopulations (São Paulo and Paraná). The analysis of molecular variance revealed that 90% of the genetic distribution of the pathogen occurs among isolates within the subpopulation (states or host of origin). There was no correlation between geographic and genetic distance ($r = -0{\cdot}024$, $P = 0{\cdot}74$), which together with the high number of migrants and the low degree of differentiation in populations of *H. vastatrix*, is consistent with the fact that the inoculum is probably easily dispersed by wind over long distances, allowing dispersal of the pathogen among coffee growing areas in Brazil. Therefore, it is difficult to predict the durability of resistant sources to coffee rust. The recommendation for the breeding programmes is thus to incorporate multigenic resistance as a control strategy.

Keywords: AFLP analysis, *Coffea arabica*, *Coffea canephora*, coffee leaf rust, durable resistance, population genetics

Introduction

Brazil is the largest coffee producer in the world, with a planted area of 2·33 million hectares (Beling & Neves da Silveira, 2013). Brazilian coffee crops are composed of 70% *Coffea arabica* (arabica coffee) and 30% *Coffea canephora* (conilon coffee). Arabica coffee is grown in regions above 800 m in the states of Minas Gerais, São Paulo, Paraná, Bahia, Rio de Janeiro and parts of Espírito Santo. Conilon predominates in regions of low altitude and higher temperatures in the states of Espírito Santo, Rondônia and parts of Bahia and Minas Gerais. Despite the importance of coffee production in Brazil and although resistant cultivars are available to farmers, 95% of planted varieties are susceptible to rust.

Rust, caused by the biotrophic fungus *Hemileia vastatrix*, is the main disease of coffee, and is widely distributed in producing regions. First reported in Brazil in 1970 in the state of Bahia, the disease spread rapidly throughout the country (Chaves *et al.*, 1970), leading to crop losses of 35–50% depending on weather conditions (Zambolim *et al.*, 1999).

*E-mail: zambolim@ufv.br

Published online 29 June 2015

Although the primary method of rust control is the application of protective or systemic fungicides, the most popular method for disease control is the use of resistant varieties, due to its efficiency, low cost and reduced impact on the environment (Zambolim *et al.*, 1999).

Híbrido de Timor, a natural hybrid between *C. arabica* and *C. canephora*, has been extensively used as the main source of resistance to rust. Resulting from this breeding strategy, different coffee cultivars resistant to rust have been released. However, over time the majority of these cultivars have become susceptible to disease due to the emergence of variants of the pathogen. The high genetic variability of *H. vastatrix* favours the emergence of new pathogen races capable of overcoming resistance in the field (Várzea & Marques, 2005). Cultivation over large areas also offers a favourable environment for rapid evolution of the pathogen, permitting the gradual movement of rust epidemics and the distribution of genotypes (Nunes *et al.*, 2009).

The mechanisms leading to the emergence of new races of *H. vastatrix* are still not clear. Because the sexual stage of the fungus has not been encountered (Gopalkrishnan, 1951; Várzea & Marques, 2005), mutation has been hypothesized as the principal mechanism for variability of the fungus (Várzea & Marques, 2005). Although morphological evidence has indicated the occurrence of karyogamy and meiosis in asexual spores

Plant Pathology

since 1967 (Rajendren, 1967), only in 2011 did an image cytometry study of DNA content reveal the presence of a novel type of sexual reproduction hidden within asexual spores of *H. vastatrix*, called cryptosexuality (Carvalho *et al.*, 2011). According to the authors, this type of reproduction could explain the frequent and rapid emergence of new physiological races of *H. vastatrix*.

It is therefore of great importance to investigate the patterns of genetic variability of *H. vastatrix*, allowing evaluation of the development potential of the predominant populations where breeding programmes are being developed, and in those where it is intended to introduce resistant material (Nunes *et al.*, 2009). Despite the importance of rust for coffee production in Brazil, little is known about the diversity of *H. vastatrix*. It is not known, for example, how the fungal population behaves in conilon and arabica coffee, if the fungal haplotypes are independent of the host species or geographic region, if there is dispersion of isolates among coffee species, or if mating occurs among the isolates. This information is important for development of rust control strategies, epidemiological studies of the disease and determination of evolutionary behaviour of the predominant fungal populations.

The objective of this study was therefore to assess the genetic diversity and population structure of *H. vastatrix* with respect to the host and geographical origin, using single-uredinial isolates from *C. arabica*, *C. canephora* and derivatives of Híbrido de Timor and Icatu (HDTI), from the five major producing states in Brazil.

Materials and methods

Geographical and host origins

Coffee leaves infected by *H. vastatrix* were sampled in *C. arabica*, *C. canephora* and derivatives of Híbrido de Timor and Icatu in the main coffee producing regions of Minas Gerais (27 fields), Espírito Santo (31 fields), São Paulo (20 fields), Bahia (24 fields) and Paraná (10 fields). One hundred infected leaves were collected from 10 plants of the same coffee tree genotype within the same production field. Spores from the 100 leaves from each field were then bulked, and single uredinial isolates were produced by inoculating the bulked spores at a low density on a susceptible host. In total, 112 isolates of *H. vastatrix* were obtained using this strategy. A summary of the origins, hosts and georeferences of each *H. vastatrix* isolate is presented in Table S1.

Storage and inoculation

Isolates of *H. vastatrix* were stored according to the previously described methodology (Zambolim & Chaves, 1974). Inoculation of the plants was performed with the aid of a camel hair-brush, depositing urediniospores of *H. vastatrix* on the abaxial surface of young and completely expanded leaves of the cultivar Catuaí Vemelho IAC 44, a susceptible host. Next, distilled water was sprayed on the leaves with a spray bottle until slight wetness was achieved. After this procedure, the plants were covered with plastic bags and transferred to a dark cloud chamber at 22 ± 2°C for 48 h and relative humidity near 100%. After this period, the leaves were cleaned with cotton to remove urediniospores remaining on the abaxial side of the leaf, to prevent subsequent colonization by fungal hyperparasites such as *Lecanicillium lecanii*. The plants were maintained in growth chambers with individual sections, at 22 ± 2°C and a photoperiod of 12 h, until the onset of signs of the pathogen.

To obtain the single-uredinial isolates, the methodology used was the same as that described by Capucho *et al.* (2009) and similar to that described for other rust-causing pathogens such as *Puccinia coronata* (Leonard *et al.*, 2005) and *Puccinia triticina* (Elyasi-Gomari & Panteleev, 2006; Kolmer & Ordonez, 2007). The monopustule isolates were multiplied to obtain sufficient urediniospores for DNA extraction.

The fungal hyperparasite *L. lecanii* was isolated on potato dextrose agar (PDA) for extraction of its DNA.

DNA extraction

Genomic DNA from each of the monopustule isolates of *H. vastatrix* and *L. lecanii* was extracted from 30 and 50 mg of urediniospores and mycelium, respectively (Raeder & Broda, 1985). DNA quantification of each isolate was performed in a NanoDrop 2000 spectrophotometer and its quality confirmed by electrophoresis in 1% agarose gel. The DNAs were diluted to a final concentration of 100 ng μL^{-1} and stored at −20°C until use.

Genomic DNA of the host for pathogen multiplication, Catuaí Vemelho IAC 44, was extracted according to the protocol of Diniz *et al.* (2005).

AFLP markers

The DNA amplification reactions for each isolate of *H. vastatrix*, *L. lecanii* and the host for pathogen multiplication Catuaí Vemelho IAC 44 were performed according to the protocol described by Brito *et al.* (2010), with modifications. Samples of genomic DNA (800 ng) from the 112 monopustule isolates of *H. vastatrix* were cleaved with the restriction enzymes *Eco*RI and *Mse*I to obtain fragments that were bound to specific double-strand adapters by T4 DNA ligase. The samples were amplified by PCR using DNA bound to adapters. The ligation mixture was diluted to 1:10 in distilled water and stored at −20°C.

Pre-amplification was conducted using 2·5 μL of diluted template DNA, 2·5 μL of 10× PCR buffer (200 mM Tris-HCl (pH 8·4), 500 mM KCl), 0·75 μL $MgCl_2$ (50 mM), 0·5 μL primer *Eco*RI+A (75 ng μL^{-1}), 0·5 μL primer *Mse*I+C (75 ng μL^{-1}), 0·625 μL dNTPs (10 mM), 17·425 μL water and 5 U *Taq* DNA polymerase. The pre-selective amplification programme consisted of 23 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 60 s and extension at 72°C for 60 s. The mixture from pre-amplification was diluted to 1:40 in distilled water and stored at −20°C.

In selective amplification, seven primer combinations were used with three selective nucleotides (Table S2). The Eco (E) selective AFLP primers were marked with the fluorescent dyes FAM and NED at the 5′ end. Each amplification reaction contained 5 μL of the diluted pre-amplified DNA, 0·5 μL of the primer E-ANN (50 ng), 0·6 μL of the primer M-CNN (50 ng), 0·5 μL dNTPs (10 mM), 2 μL of 10× PCR buffer (200 mM Tris-HCl (pH 8·4), 500 mM KCl), 0·6 μL $MgCl_2$ (50 mM), 10·7 μL water and 5 U *Taq* DNA polymerase. The selective amplification programme included 13 cycles consisting of denaturation at 94°C for 30 s, annealing at 65°C for 30 s (with a decrease of 0·7°C every cycle) and extension at 72°C for 60 s. The amplification programme consisted of 32 denaturation cycles at 94°C

for 30 s, annealing at 56°C for 30 s and extension at 72°C for 60 s, and one final extension cycle at 72°C for 20 min.

Subsequently, the resulting amplification fragments were separated and detected by the capillary electrophoresis system with automated sequencer ABI 3130XL (Applied Biosystems, Inc.), using ROX500 as the internal standard.

Data analysis

Differences in electrophoretic patterns among the isolates were evaluated using the software GeneMapper v. 4.1 (Applied Biosystems). The AFLP fragments were coded as binary characters for the presence (1) and absence (0) of the fragment.

To increase the consistency of the data, only fluorescence emission peaks greater than 100 rfu within the range from 50 to 500 bp were included in the analysis. Reliability of the data was ensured by amplification of 10% of the isolates in duplicate for each primer combination, where the patterns that presented many polymorphic peaks were repeated. The peaks present in the *C. arabica* and *L. lecanii* controls were not considered for analysis.

Genotypic diversity was estimated by measures of richness, diversity and evenness. In order to reduce the error due to different sample sizes of the population under study, genotypic richness $E(g_n)$ was estimated based on rarefaction curves (Grünwald *et al.*, 2003), using the packages vegan and vegetarian implemented in the R software (R Development Core Team, 2014).

To measure the genotypic diversity, N_1 (Hill, 1973) and G indices (Stoddart & Taylor, 1988) were calculated. Confidence intervals for the values of N_1 and G were calculated from 1000 bootstrap, resampling the original frequency of the multilocus genotypes identified by the vegan and vegetarian packages of the R software (R Development Core Team, 2014).

Similarly, genotypic evenness was evaluated using the index $E_5 = [(G - 1)/(N_1 - 1)]$ (Grünwald *et al.*, 2003). After estimating the allelic frequency for each locus, analysis of the genetic diversity of Nei (1973) was also performed used the program popgene v. 1.3.2 (Yeh *et al.*, 1999).

To investigate the population structure of *H. vastatrix*, a *K*-means hierarchical clustering approach and DAPC were performed. To assess the genetic relationship among haplotypes (multilocus genotypes, MLGs) of *H. vastatrix* and the presence/absence of any structuration of the population, the Nei genetic distance (1973) was calculated and a minimum spanning network (MSN) was constructed using the packages poppr and adegenet (Kamvar *et al.*, 2014) in the R software.

Correlation between genetic similarity and geographic distance among isolates of *H. vastatrix* was analysed with the Mantel test (Mantel, 1967) in the R program (R Development Core Team, 2014). For this, a geographical distance matrix was generated by calculating the distance between isolates using the geographical coordinates, which was then compared with a matrix of genetic distances. Moreover, genetic differentiation (G_{ST}) and the number of migrants (N_m) were calculated pairwise between populations of *H. vastatrix* in the program popgene v. 1.3.2 (Yeh *et al.*, 1999).

The distribution analysis of genetic variability within and among populations was performed by analysis of molecular variance (amova), as suggested by Excoffier *et al.* (1992), using the program Arlequin v. 3.5 (Excoffier *et al.*, 2006). For this, populations of *H. vastatrix* classified by the state of origin were divided into subpopulations, which were defined by producing regions within each state. Those grouped by host were subdivided into subpopulations defined by the state of origin.

Evidence of sexual recombination in the total population of *H. vastatrix* ($n = 112$) was evaluated by the association index (I_A) and the $\bar{r}_d$ statistic (Maynard Smith *et al.*, 1993), by comparison among the observed values and expected values in 1000 replications by randomization of indices using the program multilocus v. 1.3 (Agapow & Burt, 2001), and the poppr package. Estimates of linkage disequilibrium (I_A and $\bar{r}_d$) for the total population, and for each of the hierarchical subsets (states or hosts), were compared with the expected distribution for each locus when in random association (null hypothesis). Deviation from the null hypothesis was evaluated with 1000 permutations of alleles among individuals of each locus. The expected values for 1000 permutations, for each subset, were used for constructing histograms for subsequent graphical analysis of the values of I_A and $\bar{r}_d$.

Results

AFLP markers

The seven primer pairs allowed the detection of 184 loci, of which 86 (46·7%) were polymorphic. The number of polymorphic fragments per primer combination ranged from 5 to 22, with sizes between 54 and 480 bp (Table S2). Among the 112 isolates analysed, 68 genotypes, i.e. AFLP patterns, were found, of which 59 (86·76%) were represented by a single isolate. Reliability of the data was confirmed by the same pattern presented for the amplified isolates in replicates and by the DNA controls (*L. lecanii* and Catuaí Vemelho IAC 44) that showed distinct AFLP patterns when compared to each other and to the *H. vastatrix* isolates, indicating no contamination.

Total population

Genetic diversity was assessed in the total population ($n = 112$) by the Hill's and Stoddart & Taylor's indices, which were, respectively, 40·394 and 18·447 (Table 1). Nei's gene diversity (h) for all loci was 0·027 (SD ± 0·064). A total of 68 haplotypes were detected in 112 isolates analysed, of which 59 were unique.

Subdivided populations

Populations from Espírito Santo and from *C. arabica* showed an elevated number of polymorphic loci, 34 and 39%, respectively. The lowest percentages of polymorphic loci were found in populations from Paraná and HDTI derivatives (3·8 and 6·5%, respectively) (Table S3). The total genetic diversity (H_T) in the pathogen population ranged from 0·0118 to 0·0352 (Table S3).

The number of genotypes present in each population was classified as the genotypic richness ('pobs'), and for the purpose of multiple comparison, the values were scaled by the rarefaction curve based on the lowest 'n' value of the populations ($E_g(n)$). In populations grouped according to hosts of origin, this value ranged from 5 in *C. canephora* to 6 in HDTI derivatives (Table 1). Considering the populations grouped by state, the $E_g(n)$ ran-

Table 1 Indices of richness, diversity and evenness for the populations of *Hemileia vastatrix* subdivided according to host species

	Population defined by host			
Statistic	*Coffea arabica*	*Coffea canephora*	Derivatives of Híbrido de Timor and Icatú	Overall
Sample size (n)[a]	68	38	6	112
pobs[b]	45	24	6	68
$E_g(n)$ for $n = 6$[c]	5·6	5	6	5·4
N_1[d]	33·986 (28·76–39·21)	15·816 (11·85–19·79)	6·000 (4·40–7·60)	40·394 (33·47–47·32)
G[e]	22·891 (16·84–28·95)	8·914 (4·8 –12·97)	6·000 (4·44–7·56)	18·447 (11·96–24·93)
E_5[f]	0·664	0·534	1·000	0·443

Numbers in parentheses indicate the confidence interval calculated from the bootstrapping approach with 1000 replicates.
[a]Number of individuals sampled.
[b]Observed number of genotypes.
[c]Expected number of genotypes based on the rarefaction curve for a sample of six individuals (smallest population).
[d]Hill's index of diversity.
[e]Stoddart & Taylor's index of diversity.
[f]E_5 evenness index calculated by $(G - 1)/(N_1 - 1)$.

Table 2 Indices of richness, diversity and evenness for the populations of *Hemileia vastatrix* subdivided according to the state of origin

	Population defined by state				
Statistic	Bahia	Espírito Santo	Minas Gerais	São Paulo	Paraná
Sample size (n)[a]	24	31	27	20	10
pobs[b]	12	24	22	17	8
$E_g(n)$ for $n = 10$[c]	5·8	8·7	8·9	9·3	8
N_1[d]	6·539 (4·03–9·05)	19·705 (15·92–23·49)	19·038 (15·34–22·74)	16·245 (13·52–18·97)	7·192 (5·24–9·15)
G[e]	3·646 (1·60–5·69)	14·343 (9·92–18·77)	14·878 (10·71–19·05)	15·385 (12·41–18·36)	6·250 (4·23–8·27)
E_5[f]	0·478	0·713	0·769	0·944	0·848

Numbers in parentheses indicate the confidence interval calculated from the bootstrapping approach with 1000 replicates.
[a]Number of individuals sampled.
[b]Observed number of genotypes.
[c]Expected number of genotypes based on the rarefaction curve for a sample of 10 individuals (smallest population).
[d]Hill's index of diversity.
[e]Stoddart & Taylor's index of diversity.
[f]E_5 evenness index calculated by $(G - 1)/(N_1 - 1)$.

ged from 5·8 for populations from Bahia to 9·3 from São Paulo (Table 2).

Genotypic diversity was estimated by the N_1 and G indices. When the genotypic diversity of isolates was analysed based on the host of origin, the population of HDTI derivatives showed higher genotypic diversity compared to populations of *C. arabica* and *C. canephora* (Table 1). The population from Bahia showed lower genotypic diversity compared to other populations when grouped by state of origin (Table 2).

The uniformity of distribution of genotypes within populations was assessed by the index of evenness (E_5), where the highest value was found in the population of HDTI when grouped by host (Table 1). In relation to the state of origin, the population of Bahia showed the lowest evenness compared to the other populations (Table 2).

The pairwise comparison of genetic differentiation and number of migrants among populations of *H. vastatrix* grouped both by host and by state of origin was reported to be low to moderate (Tables 3 & 4). When grouped by host species, the highest value of N_m was found between populations of *C. canephora* and *C. arabica* ($N_m = 19·61$; Table 3). Upon analysing grouping by state of origin, the populations of Bahia and Espírito Santo had the highest number of pairwise migrants among these populations ($N_m = 22·14$; Table 4).

From the analysis of molecular variance, the percentage of total variation was estimated among isolates grouped by host of origin, among isolates from different states and the same host, and among isolates belonging to the same state, with 90·39% of total variation due to variations within the state of origin (Table 5). Grouping of the data according to the state of origin of the isolates showed that 90·16% of the total variation was due to variation within producing regions of the same state (Table 6). The variations between populations of *H. vastatrix* belonging to the same host or the same state were low (0·94 and 0·31, respectively), indicating an absence of population structure in the set of isolates evaluated (Tables 5 & 6).

Table 3 Estimates of population differentiation (above the diagonal) and number of migrants (below the diagonal) between populations of *Hemileia vastatrix* defined by host

Population	*Coffea arabica*	*Coffea canephora*	Derivatives of Híbrido de Timor and Icatú
C. arabica	–	0·02	0·10
C. canephora	19·61	–	0·21
Derivatives of Híbrido de Timor and Icatú	4·67	1·92	–

G_{ST} total = 0·15.

Table 4 Estimates of population differentiation (above the diagonal) and number of migrants (below the diagonal) between populations of *Hemileia vastatrix* defined by state

Population	Bahia	Espírito Santo	Minas Gerais	São Paulo	Paraná
Bahia	–	0·02	0·11	0·06	0·17
Espírito Santo	22·14	–	0·06	0·03	0·08
Minas Gerais	3·94	8·04	–	0·05	0·04
São Paulo	8·30	16·24	10·42	–	0·07
Paraná	2·43	5·81	13·59	7·12	–

G_{ST} total = 0·12.

Table 5 Analysis of molecular variance (AMOVA) of isolates of *Hemileia vastatrix* grouped based on the host of origin

Factor of variation	df	SS[a]	CV[b]	% V[c]	FI[d]	*P*
Among hosts	2	11·27	0·024	0·94	0·009	0·285
Among states within the hosts	6	28·38	0·218	8·66	0·087	0·000
Within the state	102	231·74	2·272	90·39	0·096	0·000
Total	110	271·39	2·514	100·00	–	–

[a]Sum of squares.
[b]Coefficient of variance.
[c]Percentage variation.
[d]Fixation indices.

Table 6 Analysis of molecular variance (AMOVA) of isolates of *Hemileia vastatrix* grouped based on the state of origin

Factor of variation	df	SS[a]	CV[b]	% V[c]	FI[d]	*P*
Among states	4	22·46	0·008	0·31	0·003	0·283
Among regions within the states	6	23·98	0·240	9·53	0·096	0·000
Among the regions	100	226·73	2·267	90·16	0·098	0·000
Total	110	273·17	2·515	100·00	–	–

[a]Sum of squares.
[b]Coefficient of variance.
[c]Percentage variation.
[d]Fixation indices.

Cluster analysis

The cluster analysis performed by *K*-means and DAPC revealed the presence of three genetic groups, but they presented no relationship with the geographical origin and hosts, as can be observed in the minimum spanning network (MSN; Fig. 1a,b), based on the presence of the same haplotype in different regions and/or hosts. The Mantel test indicated that there was no correlation between genetic and geographic distance ($r = -0{\cdot}024$, $P = 0{\cdot}74$).

Sexual recombination

The hypothesis of sexual recombination in the total population ($n = 112$) of *H. vastatrix* was rejected, as this is not expected for a population in random mating ($I_A = 2{\cdot}2670$, $P < 0{\cdot}001$; $\bar{r}_d = 0{\cdot}0322$, $P < 0{\cdot}001$). However, when I_A and $\bar{r}_d$ where calculated for isolates within each state, there were indications of recombination in the subpopulations of São Paulo and Paraná ($I_A = 0{\cdot}0014$, $P = 0{\cdot}4505$; $\bar{r}_d = 0{\cdot}0001$, $P = 0{\cdot}4505$ and $I_A = 0{\cdot}479$, $P = 0{\cdot}1476$; $\bar{r}_d = 0{\cdot}0401$, $P = 0{\cdot}1476$, respectively), but not for the other three populations (Minas Gerais, Espírito Santo and Bahia; Table 7). The same hypothesis was tested with clone-corrected data and similar results were found (data not show).

Discussion

The results presented here showed low levels of population differentiation and no apparent geographic pattern in genetic diversity. The assessment of seven primer pairs allowed the detection of 184 loci, of which 46·7% were polymorphic. A total of 68 AFLP patterns were detected in 112 isolates analysed, and of these 86·7% were represented by a single isolate. In previous studies, six AFLP primer combinations were evaluated for *H. vastatrix*, detecting 349 fragments, of which only 13·2% were polymorphic (Rozo *et al.*, 2012). On the other hand, 77% of unique genotypes in 120 isolates of *H. vastatrix* were found with RAPD markers (Nunes *et al.*, 2009). Therefore, in comparison with the studies above, the AFLP-based genetic markers used in the current study were sufficiently polymorphic for discriminating all isolates of *H. vastatrix*. The results suggest that AFLP markers can be used to evaluate genetic diversity and population structure of *H. vastatrix* through estimates of variation at multiple loci across the genome.

The analysis of genotypic diversity was complemented by the analysis of gene diversity considering the total population studied. The results showed low values of Nei's gene diversity ($h = 0{\cdot}027 \pm 0{\cdot}064$), in contrast with those obtained by Maia *et al.* (2013) and Nunes *et al.* (2009), who found values seven ($h = 0{\cdot}194 \pm 0{\cdot}130$) and 10 times higher ($h = 0{\cdot}280 \pm 0{\cdot}150$), respectively. However, medium genotypic diversity was found in the total population studied ($G = 18{\cdot}447$ and $N_1 = 40{\cdot}394$), differing from those reported by the

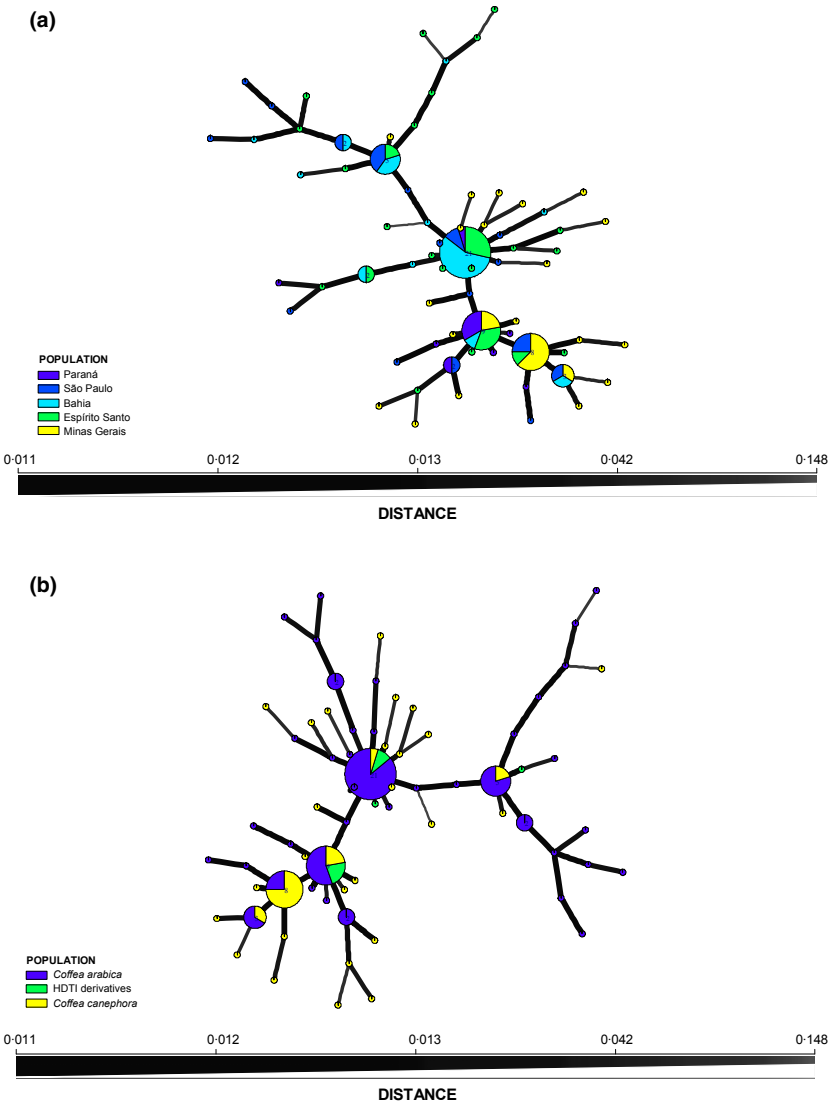

Figure 1 Minimum spanning network of 68 haplotypes detected in the studied population (112 isolates) of *Hemileia vastatrix*. Each circle represents a unique haplotype and the colours represent the sampling sites. The circle size represents the haplotype frequency. Line widths and shading represent relatedness of the haplotypes based on Nei's genetic distance. Line length is arbitrary. Grouped isolates of *Hemileia vastatrix* based on the state of origin (a) and based on the host species (b).

Table 7 The observed values of association index (I_A) and analysis of $\bar{r}_d$ and the *P* value for each estimate for the total population ($n = 112$) of *Hemileia vastatrix* and the subpopulations subdivided according the state of origin

Population	I_A	$P - I_A$	$\bar{r}_d$	$P - \bar{r}_d$
Total	2·267	0·0009	0·0322	0·0009
Paraná	0·479	0·1478	0·0401	0·1478
São Paulo	0·001	0·4505	0·0001	0·4505
Bahia	1·057	0·0029	0·0732	0·0029
Espírito Santo	3·024	0·0009	0·0514	0·0009
Minas Gerais	1·898	0·0009	0·0468	0·0009

former authors. In other rusts, such as *Phragmidium violaceum*, Nei's gene diversity was also low when using AFLP markers and single lesion isolates (Gomez *et al.*, 2006). The differences in the results may be explained by the method of collecting the isolates. In the present study and those of Gomez *et al.* (2006), all isolates evaluated were from single lesions (genetic uniform material), whereas the methods of Maia *et al.* (2013) and Nunes *et al.* (2009) consisted of bulk isolates. This fact may mask the true genetic variability due to the presence of different genotypes of the pathogen in a single sample; consequently, the band patterns analysed were of DNA fragments from several individuals.

The MSN analysis reinforces the low gene and medium genotypic diversity found, consistent with a single introduction event, and evolution of the Brazilian population with a short genetic background, characteristic of a founder effect. This fact is verified by the high degree of similarity among the isolates (>90%), due to the introduction of isolates among the states as a result of the exchange of germplasm of the hosts or pathogen dispersal over long distances. The high number of unique haplotypes may be due to the effects of mutation events.

The lack of formation of a spatial group, indicated by no correlation between genetic and geographic distances

with the Mantel test ($r = -0{\cdot}024$, $P = 0{\cdot}74$), is indicative of the dispersal of urediniospores over long distances. The importance of wind in the dispersal of *H. vastatrix* was proven by the presence of urediniospores at an elevation of 1000 m with good viability and germination (Martinez *et al.*, 1975). Wind dispersal has been suggested as the cause of introduction and rapid spread of coffee rust in South America (Bowden *et al.*, 1971), especially in Brazil (Chaves *et al.*, 1970). The absence of a correlation between geographic distance and genetic distance was also reported for other rust fungi, including *Cronartium ribicola*, *Melampsora epitea*, *P. triticina* and *Melampsora larici-populina*. These fungi also have an efficient mechanism for dissemination over long distances (Hurtado & Ramstedt, 2002; Hamelin *et al.*, 2005; Mebrate *et al.*, 2006; Pei *et al.*, 2007).

When the total population ($n = 112$) of *H. vastatrix* was studied for the reproduction regime, the hypothesis of sexual recombination was rejected. Maia *et al.* (2013) found a similar result, indicating the absence of sexual recombination in the population. In another study, Nunes *et al.* (2009) reported linkage disequilibrium in only 22% of pairs for the tests performed. In the other 78% of tests, the null hypothesis of random association of alleles was accepted. However, when a set of isolates collected in São Paulo and Paraná were analysed in the present study, the presence of a recombination signal was observed. If sexual recombination does exist, it may be explained by the phenomenon of cryptosexuality, which consists of the occurrence of a recombination within an asexual structure, the urediniospore (Carvalho *et al.*, 2011).

The effect of the host on genetic diversity and population structure of *H. vastatrix* was verified in the subdivided population of 68 isolates collected from *C. arabica*, 38 from *C. canephora* and six from HDTI. The statistical analysis of these isolates showed that the highest value for the variables richness and evenness were found in the population derived from HDTI. This fact suggests the presence of a larger number of unique genotypes in these populations, and that these genotypes are more evenly distributed in relation to other populations (Grünwald *et al.*, 2003).

When a pairwise comparison was conducted between these isolate groups, low values of genetic differentiation (G_{ST}) were found, indicating that there is a low variation among isolates of *H. vastatrix* in the host species. Maia *et al.* (2013) also found a low genetic differentiation among isolates grouped according to the host species ($G_{ST} = 0{\cdot}026$). In contrast, Gouveia *et al.* (2005) reported a high genetic differentiation among populations from *C. arabica* ($G_{ST} = 0{\cdot}788$) and HDTI ($G_{ST} = 0{\cdot}783$). The high value of G_{ST} reported in this study may be explained by the origins of the isolates of *H. vastatrix*, which were from three continents: Africa, Asia and South America. Analysing only isolates from Minas Gerais, Nunes *et al.* (2009) found a high average genetic differentiation ($G_{ST} = 0{\cdot}220$), similar to that found by Gouveia *et al.* (2005) ($G_{ST} = 0{\cdot}260$), but this value was low compared to isolates from Africa ($G_{ST} = 0{\cdot}865$) and Asia ($G_{ST} = 0{\cdot}768$).

The AMOVA and cluster analyses also suggest that the host does not affect the genetic structure of the *H. vastatrix* population. This is consistent with coffee genetic breeding and the co-evolution of plant pathogens. In Brazil, breeders use natural and artificial hybrids of *C. canephora* and *C. arabica* (Híbrido de Timor or Icatu) to transfer the rust resistance genes to arabica cultivars. In this case some *C. arabica* cultivars have genes of both species. Therefore, selection pressure exerted on the pathogen population with high genotypic richness can promote the selection of new and complex races of the pathogen that are capable of infecting both species.

The 112 *H. vastatrix* isolates were also subdivided into the state in which they were collected, so as to evaluate the geographical structure of diversity. Using this strategy, 34 isolates were grouped in the state of Bahia, 31 in Espírito Santo, 27 in Minas Gerais, 20 in São Paulo and 10 in Paraná. The lowest values for genetic richness, diversity and evenness were found for isolates from Bahia. The first report of *H. vastatrix* in Brazil occurred in this state, and it is consistent with what is expected for the founder population. During dispersion and colonization of the new habitat, geographic isolation will generally reduce the genetic diversity of the population by means of genetic drift (Nei *et al.*, 1975). Thus, few haplotypes are present in Bahia state, but in high frequencies. However, these haplotypes are also common to other coffee producing regions.

The low value of genetic differentiation (G_{ST}) was also observed in isolates grouped by geographic region, as in those grouped by host. In clonal populations, migration contributes to prevent genetic differentiation among populations, because there can occur a 'mixture' of different pathogen haplotypes. The fact that the inoculum is easily dispersed by wind over long distances allows dispersion of the pathogen and spread of the disease (Bowden *et al.*, 1971). Thus, the facilitated movement of propagules may have contributed to dispersion of *H. vastatrix* in Brazil.

Another factor that facilitates dispersion of the urediniospores of *H. vastatrix* is the presence of large continuous areas of coffee plantations in Brazil, which is proven by the higher number of migrants ($N_m = 22{\cdot}14$) found in the states of Bahia and Espírito Santo. For other rusts two situations were observed, high and low migration (measured as gene flow), depending on the presence or absence of large continuous areas of the hosts (Hamelin *et al.*, 2000; Keiper *et al.*, 2006). For example, there was a low migration rate ($N_m < 1$) among isolates of *C. ribicola* (white pine rust) where the absence of hosts between the east and west regions of the US prevented migration of this fungus (Hamelin *et al.*, 2000). In this study, it was also possible to observe an association between genetic differentiation and the geographic origin of the pathogen populations. The presence of multiple clonal lineages and subpopulations of the pathogen, as well as wide dispersion of races and identical genotypes were also found for *Puccinia graminis* f. sp. *avenae* in all

grain-producing regions of Australia, confirming the high rate of migration for this pathogen (Keiper *et al.*, 2006).

Thus, areas of continuous coffee plantations allow a gradual movement of coffee leaf rust epidemics and distribution of genotypes. If any genetic differentiation is detected in populations of *H. vastatrix* it is probably caused by the cultural practices used (Nunes *et al.*, 2009).

By means of the AMOVA it was shown that there was no significant difference between the pathogen populations defined by the state of origin of the isolates. The highest percentage of genetic variance (90%) occurred among isolates within the subpopulation. Thus, mechanisms of genetic variability, together with dispersion of inoculum, promote high genetic variability within the subpopulations. Variability is conserved among populations due to clonal admixture that maintained the same allele frequencies in populations of *H. vastatrix*. The *K*-means and DAPC approach confirmed that there is no structuring of the populations in relation to the state of origin of the isolates.

Genetic drift and migration are two evolutionary mechanisms that may be associated with these results. Coffee rust was first reported in Brazil in 1970, about 100 years after its report in Sri Lanka. Therefore, recently established populations still reflect the founder effect, with less genetic diversity among populations (Nunes *et al.*, 2009). On the other hand, it was found that there is migration between populations, which contributed to a lower differentiation between them.

The results obtained with the complete and subdivided populations are in agreement with those observed by Gouveia *et al.* (2005) and Maia *et al.* (2013), confirming that populations of *H. vastatrix* are dysfunctional in relation to the type of host and geographical origin. Three hypotheses are considered to explain these finds. First, the introgression of rust resistance genes of *C. canephora* into genotypes of *C. arabica* contributes to the lack of population structure of *H. vastatrix* considering the host. Secondly, AFLP markers used in this study were not associated with regions of the genome under selection imposed by host resistance genes; and thirdly, of high migration rates are occurring in the studied populations.

A low degree of differentiation in populations of *H. vastatrix* shown in this study, together with facilitated dispersal of the inoculum over long distances, allows for pathogen dispersal among coffee-growing areas. Furthermore, the large variation within subpopulations of the fungus suggests a high evolutionary rate and may explain the breakdown of resistance in cultivars derived from HDTI to the rust. Due to the occurrence of migration between populations and the large variability within subpopulations of the pathogen, it is recommended that breeding programmes adopt the incorporation of horizontal resistance as a control strategy.

Acknowledgements

This work was supported by grants from the National Council of Scientific and Technological Development (CNPq), Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café (CBP&D/Café), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) and Financiadora de Estudos e Projetos (FINEP).

References

Agapow P, Burt A, 2001. Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**, 101–2.

Beling RR, Neves da Silveira D, 2013. *Anuário Brasileiro do Café 2013*. Santa Cruz do Sul, Brasil: Editora Gazeta Santa Cruz Ltda.

Bowden J, Gregory PH, Johnson CG, 1971. Possible wind transport of coffee leaf rust across the Atlantic Ocean. *Nature* **229**, 500–1.

Brito GG, Caixeta ET, Gallina AP *et al.*, 2010. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. *Euphytica* **173**, 255–64.

Capucho AS, Caixeta ET, Zambolim EM, Zambolim L, 2009. Herança da resistência do 'Híbrido de Timor' UFV443-03 à ferrugem-do-cafeeiro. *Pesquisa Agropecuária Brasileira* **44**, 276–82.

Carvalho CR, Fernandes RC, Carvalho GMA, Barreto RW, Evans HC, 2011. Cryptosexuality and the genetic diversity paradox in coffee rust, *Hemileia vastatrix*. *PLoS ONE* **6**, e26387.

Chaves GM, Cruz Filho J, Carvalho MG, Matsuoka K, Coelho DJ, Shimoya CA, 1970. Ferrugem do cafeeiro (*Hemileia vastatrix*): revisão de literatura com observações e comentários sobre a enfermidade no Brasil. *Seiva* **30**, 1–75.

Diniz LEC, Sakiyama NS, Lashermes P *et al.*, 2005. Analysis of AFLP markers associated to the *Mex-1* resistance locus in Icatú progenies. *Crop Breeding and Applied Biotechnology* **5**, 387–93.

Elyasi-Gomari S, Panteleev VK, 2006. Virulence polymorphism of *Puccinia recondita* f. sp. *tritici* and effectiveness of *Lr* genes for leaf rust resistance of wheat in Ukraine. *Plant Disease* **90**, 853–7.

Excoffier L, Smouse PE, Quattro JM, 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–91.

Excoffier L, Laval G, Schneider S, 2006. ARLEQUIN *ver 3.01. An Integrated Software Package for Population Genetics Data Analysis*. Bern, Switzerland: Computational and Molecular Population Genetics Lab.

Gomez DR, Evans KJ, Harvey PR *et al.*, 2006. Genetic diversity in the blackberry rust pathogen, *Phragmidium violaceum*, in Europe and Australasia as revealed by analysis of SAMPL. *Mycological Research* **110**, 423–30.

Gopalkrishnan KS, 1951. Notes on the morphology of the genus *Hemileia*. *Mycologia* **43**, 271–83.

Gouveia MMC, Ribeiro A, Várzea VMP, Rodrigues CJ Jr, 2005. Genetic diversity in *Hemileia vastatrix* based on RAPD markers. *Mycologia* **97**, 396–404.

Grünwald NJ, Goodwin SB, Milgroom MG, Fry WE, 2003. Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology* **93**, 738–46.

Hamelin RC, Hunt RS, Geils BW, Jensen GD, Jacobi V, Lecours N, 2000. Barrier to gene flow between Eastern and Western populations of *Cronartium ribicola* in North America. *Phytopathology* **90**, 1073–8.

Hamelin RC, Allaire M, Bergeron MJ, Nicole MC, Lecours N, 2005. Molecular epidemiology of white pine blister rust: recombination and spatial distribution. *Phytopathology* **95**, 793–9.

Hill MO, 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427–32.

Hurtado S, Ramstedt M, 2002. AFLP comparison of distant *Melampsora epitea* (willow rust) population. *Mycological Research* **106**, 1400–7.

Kamvar ZN, Tabima JF, Grunwald NJ, 2014. POPPR: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**, e281.

Keiper FJ, Haque MS, Hayden MJ, Park RF, 2006. Genetic diversity in Australian populations of *Puccinia graminis* f. sp. *avenae*. *Phytopathology* **96**, 96–104.

Kolmer JA, Ordonez ME, 2007. Genetic differentiation of *Puccinia triticina* populations in Central Asia and the Caucasus. *Phytopathology* **97**, 1141–9.

Leonard KJ, Anikster Y, Manisterski J, 2005. Virulence associations in oat crown rust. *Phytopathology* **95**, 53–61.

Maia TA, Zambolim EM, Caixeta ET, Mizubuti ESG, Zambolim L, 2013. The population structure of *Hemileia vastatrix* in Brazil inferred from AFLP. *Australasian Plant Pathology* **42**, 533–42.

Mantel N, 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209–20.

Martinez JA, Palazzo DA, Karazawa M, Monteiro MVM, Reu NRN, 1975. Presença de esporos de *Hemileia vastatrix* Berk & Br. agente causal da ferrugem do cafeeiro, em diferentes altitudes nas principais áreas cafeeiras dos Estados de São Paulo e Paraná (Brasil). *O Biológico* **41**, 77–88.

Maynard Smith J, Smith NH, O'Rourke M, Spratt BG, 1993. How clonal are bacteria? *Proceedings of the National Academy of Sciences, USA* **90**, 4384–8.

Mebrate SA, Dehne HW, Pillen K, Oerke EC, 2006. Molecular diversity in *Puccinia triticina* isolates from Ethiopia and Germany. *Journal of Phytopathology* **154**, 701–10.

Nei M, 1973. Analysis of gene diversity in subdivision populations. *Proceedings of the National Academy of Sciences, USA* **70**, 3321–3.

Nei M, Maruyama T, Chakraborty R, 1975. The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10.

Nunes CC, Maffia LA, Mizubuti ESG, Brommonschenkel SH, Silva JC, 2009. Genetic diversity of populations of *Hemileia vastatrix* from organic and conventional coffee plantations in Brazil. *Australasian Plant Pathology* **38**, 445–52.

Pei MH, Bayon C, Ruiz C, Tubby I, 2007. Population structure of poplar rust *Melampsora larici-populina* in the UK inferred from AFLP. *Plant Pathology* **56**, 472–9.

R Development Core Team, 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raeder U, Broda P, 1985. Rapid preparation of DNA from filamentous fungi. *Letters in Applied Microbiology* **1**, 17–20.

Rajendren RB, 1967. New type of nuclear life cycle in *Hemileia vastatrix*. *Nature* **213**, 105–6.

Rozo Y, Escobar C, Gaitán A, Cristancho M, 2012. Aggressiveness and genetic diversity of *Hemileia vastatrix* during an epidemic in Colombia. *Journal of Phytopathology* **160**, 732–40.

Stoddart JA, Taylor JF, 1988. Genotypic diversity: estimation and prediction in samples. *Genetics* **118**, 705–11.

Várzea VMP, Marques DV, 2005. Population variability of *Hemileia vastatrix* vs. coffee durable resistance. In: Zambolim L, Maciel-Zambolim EM, Várzea VMP, eds. *Durable Resistance to Coffee Leaf Rust*. Viçosa, Brazil: Universidade Federal de Viçosa, 53–74.

Yeh FC, Yang R, Boyle T, 1999. *POPGENE: Microsoft Window-Based Freeware for Population Genetic Analysis*. Alberta, Canada: Molecular Biology and Biotechnology Centre, University of Alberta.

Zambolim L, Chaves GM, 1974. Efeito de baixas temperaturas e do binômio temperatura-umidade relativa sobre a viabilidade dos uredosporos de *Hemileia vastatrix* Berk. et Br. e *Uromyces phaseolityca* Arth. *Experientiae* **17**, 151–84.

Zambolim L, Vale FXR, Pereira AA, Chaves GM, 1999. Manejo integrado das doenças do cafeeiro. In: Zambolim L, ed. *Produção de Café com Qualidade*. Viçosa, Brazil: Universidade Federal de Viçosa, 134–215.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Table S1. Isolates of *Hemileia vastatrix* used in the AFLP study, their host coffee genotypes and geographic location

Table S2. Description of fragments obtained by seven combinations of AFLP primers in 112 *Hemileia vastatrix* isolates

Table S3. Estimates of the percentage of polymorphic loci and total genetic diversity within populations of *Hemileia vastatrix* based on 184 AFLP markers