doi: 10.1111/j.1420-9101.2007.01462.x

Sympatric genetic differentiation of a generalist pathogenic fungus, *Botrytis cinerea*, on two different host plants, grapevine and bramble

E. FOURNIER* & T. GIRAUD†

*INRA Centre de Versailles, UMR 1290 Bioger-CPP, Versailles cedex, France
†Ecologie, Systématique et Evolution, UMR 8079 CNRS-UPS-ENGREF, Université Paris-Sud, Orsay, France

Keywords
ecological speciation;
fungi;
host races;
pathogen;
reduced viability of immigrants;
Rubus;
specialization;
sympatric speciation;
Vitis.

Abstract

Prime candidates for sympatric ecological divergence include parasites that differentiate via host shifts, because different host species exert strong disruptive selection and because both hosts and parasites are continually co-evolving. Sympatric divergence may be fostered even more strongly in phytopathogenic fungi, in particular those where sex must occur on the host, which allows adaptation alone to restrict gene flow between populations developing on different hosts. We sampled populations of *Botrytis cinerea*, a generalist ascomycete fungus, on sympatric grapes and brambles in six regions in France. Microsatellite data were analyzed using standard population genetics, a population graph analysis and a Bayesian approach. In addition to confirming that *B. cinerea* reproduces sexually, our results showed that the fungal populations on the two hosts were significantly differentiated, indicating restricted gene flow, even in sympatry. In contrast, only weak geographical differentiation could be detected. These results support the possibility of sympatric divergence associated with host use in generalist parasites.

Introduction

Speciation is a key problem in evolution. It has long been believed that species originate mostly through allopatric divergence (Mayr, 1963). The possibility of sympatric speciation in sexual populations had long been dismissed because recombination between different subsets of a population that have adapted to different resources or habitats was considered to counteract natural selection for locally adapted gene combinations (Felsenstein, 1981; Rice, 1984). Theoretical models have shown that the simplest way to eliminate the role of recombination in breaking down the effects of selection and thereby allow sympatric speciation, is to have the same gene(s) controlling pleiotropically both fitness and assortative mating, or both fitness and habitat preference if mating takes place within habitats (Rice, 1984). Another way to reduce recombination between two populations specialized on different niches is to build up an association between habitat-based fitness genes and either assortative mating genes and/or habitat preference genes if mating is restricted within habitats (e.g. Dickinson & Antonovics, 1973; Udovic, 1980; Rausher, 1984; Diehl & Bush, 1989; Johnson *et al.*, 1996). Theoretical models have shown that such sympatric speciation through ecological specialization is plausible but experimental examples are still dramatically scarce (Coyne & Orr, 1999).

Prime candidates for sympatric ecological speciation include parasites that speciate via host shifts, because different host species exert strong disruptive selection and because both hosts and parasites are continually evolving in response to each other (Kawecki, 1998). Among parasites, emphasis has been placed on phytophagous insects regarding the existence of host-associated genetic differentiation (Bush, 1994). However, phytopathogenic fungi also represent good models for studying incipient speciation by host shifts (Antonovics *et al.*, 2002; Lopez-Villavicencio *et al.*, 2005; Giraud, 2006). Indeed, a particularity of many of these organisms is that sex must occur on the host. A recent model has shown that, because of this key point in the life style, mutations for adaptation on a new host plant can affect

Correspondence: Elisabeth Fournier, *UMR 1290 Bioger-CPP*, Bâtiment 4, *INRA Centre de Versailles*, Route de Saint-Cyr, F-78026 *Versailles cedex*, France.
e-mail: Elisabeth.Fournier@supagro.inra.fr

both fitness on the host and the ability to mate on this host. Adaptation to a new host plant is thus sufficient to restrict gene flow in sympatry in some parasitic fungi, without requiring either assortative mating or host preference (Giraud *et al.*, 2006). Specialization would act in these fungi like a 'magic trait' (Gavrilets, 2004), allowing pleiotropically both adaptation to the host plant and reproductive isolation, thus facilitating sympatric speciation. In fact, many phytopathogenic fungi that had long been considered as broad generalists have recently been recognized as complexes of sibling species, specialized on different hosts (e.g. Schardl *et al.*, 1997; Le Cam *et al.*, 2002; Matsuda & Takamatsu, 2003; Le Gac *et al.*, 2007).

The filamentous fungus *Botrytis cinerea* is a phytopathogenic ascomycete responsible for grey mould on more than 200 host plants, causing severe damages on numerous crops (e.g. grapevine, kiwifruits, strawberries, tomatoes and lettuces) and weeds (e.g. brambles). It is also the agent of noble rot on grapevine, used for the elaboration of sweet wines. *Botrytis cinerea* has long been thought to be a single but morphologically variable and generalist species. Several recent studies have shown, however, that *B. cinerea* was likely to form a species complex, with restricted gene flow between different cryptic genetic groups (Giraud *et al.*, 1997; Albertini *et al.*, 2002; Fournier *et al.*, 2003, 2005). First, Giraud *et al.* (1997) identified two groups based on the presence or absence of two transposable elements (TEs) in the genome: *B. cinerea* var *vacuma* (isolates without either TE) and *B. cinerea* var *transposa* (isolates with both active TEs). The *vacuma* and *transposa* groups were frequently found in sympatry but seemed to exhibit differences in their host range (Giraud *et al.*, 1999; Muñoz *et al.*, 2002), their genetic diversity (Giraud *et al.*, 1997), their temporal succession (Giraud *et al.*, 1997; Martinez *et al.*, 2005) and their size of asexual spores (Giraud *et al.*, 1997).

More recently, studies of DNA polymorphism of different nuclear genes in *B. cinerea* populations showed that *B. cinerea* isolates consistently clustered in two different clades in the different gene phylogenies, Group I and Group II, which were therefore proposed to be phylogenetic species (Albertini *et al.*, 2002; Fournier *et al.*, 2003, 2005). Group I strains belonged exclusively to the *vacuma* TE type, whereas Group II strains included both *vacuma* and *transposa* TE types. The genetic diversity was lower within Group I, as revealed by DNA polymorphism and vegetative incompatibility tests. Groups I and II also exhibited differences in morphology (asexual spores being significantly smaller in Group II than in Group I), phenology (Group I isolates are mainly found in Spring on grapevine, whereas Group II isolates are equally present in Spring and Fall) and host range (narrower in Group I than in Group II). The differences previously reported between the *vacuma* and *transposa* strains may actually have been due to the differences between Group I and Group II and the different proportions of *vacuma* and *transposa* strains in the two cryptic species. The phylogenetic analyses could not detect any differentiation between *vacuma* and *transposa* strains within Group II, neither among strains from different hosts. This may have been due, however, to the low number of strains analyzed or to the lack of polymorphism in DNA sequences.

The goal of the present study was to use highly polymorphic microsatellite markers (Fournier *et al.*, 2002) to investigate whether a genetic structure can be detected within *B. cinerea* Group II, for instance among strains from different host plants in sympatry and/or among the *transposa* and *vacuma* strains. Strains collected on two host plants, grapevine and bramble, in sympatry in each of six French regions were analyzed using eight microsatellite markers and typed for the presence of the TE *Flipper*. The main questions addressed were: (i) Is there a geographic differentiation among populations of *B. cinerea* Group II? (ii) Are the populations on the two hosts differentiated, i.e. are the populations present on the same host in different regions genetically more similar than populations on different hosts in sympatry? (iii) Is there a significant difference in the distribution of strains with and without TEs on the different hosts? (iv) Are the strains with and without TEs genetically differentiated? (v) Do the different populations undergo regular events of recombination despite the lack of sexual structure observation, as previously reported in the Champagne region (Giraud *et al.*, 1997)?

Materials and methods

Botrytis cinerea sampling

Isolates were sampled in September and October 2002 in six French regions: Alsace (near Riquewihr), Anjou (near Beaulieu-sur-Layon), Bourgogne (near Châlon-sur-Saône), Bordeaux (near Pessac), Champagne (near Epernay) and Flandres (near Cassel); see Fig. 1. In each area except Flandres, *B. cinerea* was sampled on a cultivated plant, grape (*Vitis vinifera*), and on a wild plant, bramble (*Rubus fructicosus*), directly surrounding the sampled vineyards, i.e. not farther than 200 m from the vineyard edges. Isolates from *V. vinifera* will be hereafter called 'V isolates'; isolates from bordering *Rubus* will be called 'bR isolates'. Additional isolates were sampled on *Rubus* spp. located far away from any vineyard, i.e. more than 50 km. Such isolates from insulated *Rubus* will be called hereafter 'iR isolates'. In Flandres, a region where no grape is cultivated, we collected a single iR population. All samples were taken from sporulating necroses on fruits using sterilized cotton buds. For each sample, spores were then cultured during 3 days at 21 °C on PDA medium (Petri dishes diameter 90 cm). Two uncontaminated zones of *B. cinerea* growing mycelium were then transferred on new PDA-Petri dishes (diameter 50 cm) and cultured again 3 days at

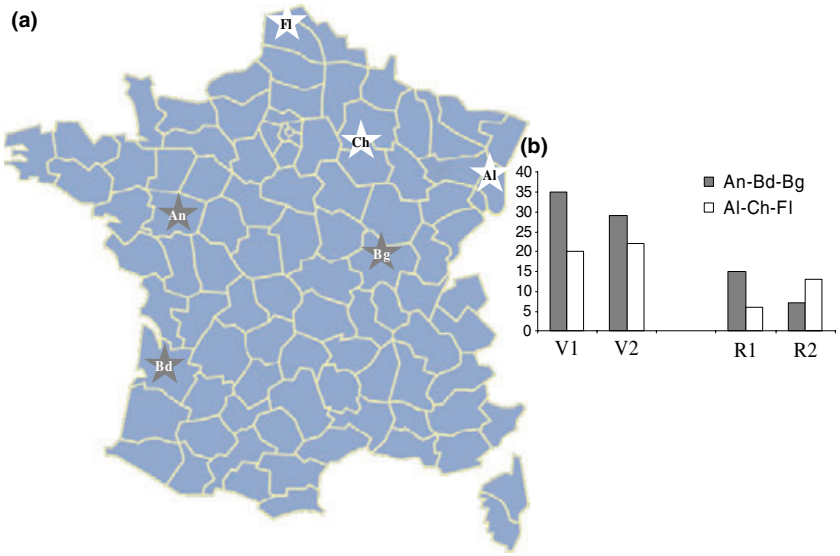

Fig. 1 (a) Geographic localization of sampling sites. An, Anjou; Al, Alsace; Bd, Bordelais; Bg, Bourgogne; Ch, Champagne; Fl, Flandres. (b) Geographic distribution of the isolates assigned to the subclusters V1 and V2 (R1 and R2) of the *Vitis* (*Rubus*, respectively) cluster. The *y*-axis represents the number of isolates from each subcluster in each region.

21 °C. This step was repeated twice in order to ensure genetic purity of the last culture. Isolates were then grown towards sporulation and spore suspensions were stored at -80 °C in 20% glycerol. A minimum of ten samples was taken per host plant and per site, but some strains were lost in the course of the purification process; the final number of isolates per host and per site varied from 4 to 35 (Table 1).

DNA extraction and amplification

For each of the 184 isolates, genomic DNA was extracted using the Chelex (Biorad) protocol described in Bucheli *et al.* (2001). Eight of the nine microsatellite markers developed by Fournier *et al.* (2002) were amplified separately for each isolate. Bc9 was excluded because preliminary experiments showed that this locus was difficult to amplify and was in linkage disequilibrium with the physically close Bc10 locus. Forward primers were labelled with a fluorescent probe. PCR amplifications were performed as in Fournier *et al.* (2002) and gel visualization as in Giraud (2004). On the basis of previous studies, *B. cinerea* Group II isolates were expected to represent around 95% of the isolates sampled in Fall. The few *B. cinerea* Group I isolates present in the sample were identified based on diagnostic alleles at the Bc6 microsatellite marker (see next section) and were excluded from further analysis.

Many repetitive elements are present within the *B. cinerea* genome, including the Fot-1 like TE *Flipper* (Lévis *et al.*, 1997) and the Gypsy-like LTR-retrotransposon *Boty* (Diolez *et al.*, 1995). To determine the TE status of each isolate, i.e. whether they were of *vacuma* or *transposa* type, we focused on the detection of *Flipper*, using the PCR approach described by Ma & Michailides (2005), with the two primers F-300 (5′-GCACAAAACCTACAGAAGA-3′) and F-1550 (5′-ATTCGTTTCTTGGACTGTA-3′). Isolates were noted *Flipper*+ when *Flipper* was successfully amplified, and *Flipper*- when no amplification was obtained.

Table 1 Number of *Botrytis cinerea* isolates genotyped for each French region and host-plant (total = 184 isolates).

	Host plant		
Region	*Vitis* (V isolates)	*Rubus* bordering vineyards areas (bR isolates)	insulated *Rubus* (iR isolates)
Alsace	18 [Al-V]	18 [Al-bR]	7 [Al-iR]
Anjou	35 [An-V]	4 [An-bR]	6 [An-iR]
Bordeaux	20 [Bd-V]	12 [Bd-bR]	
Bourgogne	13 [Bg-V]	15 [Bg-bR]	4 [Bg-iR]
Champagne	14 [Ch-V]	7 [Ch-bR]	6 [Ch-iR]
Flandres			5 [Fl-iR]

Codification of each population is indicated between brackets.

Standard population genetics analyses

The software FSTAT version 2.9.3 (Goudet, 1995; Available from http://www.unil.ch/izea/softwares/fstat.html) was used to compute observed (H_{exp}) and unbiased (H_{nb}) genetic diversity (Nei, 1987), and allelic richness over all loci in the total sample and in each of the 15 populations, i.e. for each region and host plant combination (Table 1). FSTAT was also used to calculate pairwise Weir and Cockerham's (Weir & Cockerham, 1984) F_{ST} values between all pairs of populations and to evaluate their significance by random permutations and progressive Bonferroni correction as described by Rice (1989). Mean values of F_{ST}, H_{nb} and allelic richness were compared between groups of populations using one-sided tests based on 500 permutations.

The relative contributions of the factors plant species and sampling location to the genetic variance were

estimated with hierarchical analyses of molecular variance (AMOVA) using ARLEQUIN Version 2.000 (Schneider *et al.*, 2000). Negative variance components can sometimes occur and they indicate a lack of genetic structure, the parameter's value being then 0.

Genic differentiation between each pair of populations was assessed using an unbiased estimate of the *P*-value of the probability test (or Fisher's exact test), implemented in the online version of the GENEPOP software, available at http://genepop.curtin.edu.au/. The null hypothesis for this test was 'the allelic distribution does not differ across populations'.

The number of different multilocus genotypes (GML) and the index of association (I_A) were calculated using Multilocus 1.3b (Agapow & Burt, 2001) available freely at http://www.agapow.net/software/multilocus/. The I_A is a measure of the multilocus linkage disequilibrium (Brown *et al.*, 1980; Maynard-Smith *et al.*, 1993; Haubold *et al.*, 1998), which gives information on whether two different individuals sharing the same allele at one locus are more likely to share the same allele at another locus. For any pair of individuals, the number of loci with respect to which they differ is calculated and the variance of this number is compared with that expected if there is no linkage disequilibrium. I_A is equal to zero if there is no linkage disequilibrium and increases as linkage disequilibrium increases. The null hypothesis of complete panmixia ($I_A = 0$) was tested with the procedure implemented in the software, by comparing the observed data set to 100 randomized data sets in which infinite recombination has been imposed by randomly shuffling the alleles among individuals, independently for each locus.

Graph-theoretic approach

We performed a population graph analysis (Dyer & Nason, 2004) using the PopGraph software available at http://dyerlab.bio.vcu.edu/wiki/index.php/Software#Population_Graphs. This method examines the distribution of intraspecific genetic structure without assuming any *a priori* model of population arrangement. In the resulting graphical topology, populations are represented by nodes connected by the minimal set of edges that sufficiently describes the total among-population covariance structure. In the graph theoretical framework, when a particular barrier to gene flow between two groups of population is examined, the null hypothesis states that edges connecting nodes across this hypothesized barrier should be as numerous as the number of edges in other portions of the graph.

Bayesian inference

The existence of a population structure in the total sample was investigated using the Bayesian approach implemented in Structure Version 2.1 (Pritchard *et al.*, 2000; Falush *et al.*, 2003). This clustering algorithm assumes a model in which there are K populations (where K may be unknown), each of which is characterized by a set of allele frequencies at each locus. Individuals in the sample are assigned probabilistically to these K populations, or jointly to two or more populations if their genotypes indicate that they are admixed, without consideration of their region or host of origin. Loci are assumed at Hardy–Weinberg equilibrium and linkage equilibrium within the K populations, but this approach proved to be robust to deviations from these assumptions (Falush *et al.*, 2003). We used the model with admixture, recommended when little is known about the existence of admixture (Falush *et al.*, 2003), and assumed uniform priors for the vector of proportion q_i of the individual i's genome in each cluster. The scores of individuals in the clusters (i.e. the posterior estimates of the q_i) correspond to the probability of ancestry in any one of them. We varied K from 1 to 20, each with three independent simulations to check the consistency of the results. Each simulation consisted in 100 000 Monte-Carlo Markov Chain (MCMC) iterations preceded by a burn-in period of 50 000 iterations (the burn-in period is the first set of iterations of the MCMC that is dependent on the configuration at the start – the iterations are not incorporated in the final calculation of the posterior probability). Using the distribution of maximum likelihoods, we calculated the posterior probability for each K to determine the most probable structure. Following Frantz *et al.* (2006), we also analyzed the distribution of the highest value of the ancestry coefficient for each K. We considered individuals to be assigned to a single cluster when the proportion of ancestry in this cluster was greater than 80%. On the basis of this threshold, we computed the assignment rate for each K as the proportion of individuals assigned to a single cluster (i.e. with a proportion of ancestry over the 80% threshold). The most probable structure was then given by the value of K with the highest assignment rate.

Results

Genetic diversity

Among the 184 genotyped isolates, we discriminated 180 different multilocus genotypes (GML) and the two multicopy genotypes were each represented only twice, indicating a low incidence of clonality. The overall genetic diversities calculated for the total sample and for the 15 populations were high, with nonbiased genetic diversity ranging from 0.52 to 0.79 (0.75 in the total sample) and mean allelic richness ranging from 3.12 to 6.4 (13.0 in the total sample) (Table 2). The index of association was not significantly different from zero for 12 of 15 populations, suggesting panmixia. The three remaining populations (An-V, Bd-bR and Bd-V) exhibited significant but low values of I_A (0.21, 1.419 and

Table 2 Genetic diversity indexes within the total sample and each population.

Population	*n*	H_{exp}	H_{nb}	Mean nb of alleles per locus	GML	I_A
Total	184	0.7482 (0.1888)	0.7503 (0.1893)	13	180	0.255 (<0.01)*
An-bR	4	0.6823 (0.0587)	0.7893 (0.0665)	3.3750	4	-0.72 (1)
An-iR	6	0.6597 (0.1555)	0.7197 (0.1697)	4.1250	6	0.57 (0.1)
An-V	35	0.5838 (0.2742)	0.5926 (0.2782)	4.8750	34	0.21 (0.01)*
Bd-bR	12	0.7338 (0.1093)	0.7665 (0.1143)	5.7500	11	1.419 (<0.01)*
Bd-V	20	0.6769 (0.1605)	0.6942 (0.1646)	5.7500	20	0.206 (0.03)*
Bg-bR	15	0.7031 (0.1313)	0.7280 (0.1359)	5.8750	15	0.065 (0.35)
Bg-iR	4	0.6302 (0.2592)	0.7250 (0.2973)	3.3750	4	-0.428 (1)
Bg-V	13	0.4986 (0.2775)	0.5188 (0.2889)	3.6250	13	0.11 (0.18)
Ch-bR	7	0.6227 (0.2023)	0.6720 (0.2191)	4.0000	7	0.607 (0.06)
Ch-iR	6	0.6409 (0.1671)	0.7065 (0.1822)	3.6250	6	0.162 (0.38)
Ch-V	14	0.6206 (0.1454)	0.6438 (0.1506)	4.7500	14	0.278 (0.11)
Al-bR	18	0.7210 (0.1468)	0.7421 (0.1511)	6.3750	18	0.205 (0.08)
Al-iR	7	0.6224 (0.2476)	0.6703 (0.2666)	4.2500	7	0.004 (0.53)
Al-V	18	0.6247 (0.2221)	0.6440 (0.2293)	5.0000	18	-0.069 (0.67)
Fl-iR	5	0.5956 (0.1335)	0.6667 (0.1425)	3.1250	5	0.645 (0.08)

n, sample size; H_{exp}, observed genetic diversity; H_{nb}, nonbiased genetic diversity (standard deviations between brackets); GML, number of distinct multilocus genotypes; I_A, index of associaton (associate probabilities between brackets).

0.206 respectively). The index of association computed on the total sample was significantly different from zero ($I_A = 0.255$, $P < 0.01$, Table 2), which was likely because of a Walhund effect, thus indicating genetic differentiation between populations.

Distinction between *vacuma* and *transposa* isolates

Because of the limitation in the amount of available DNA for some isolates, the presence or absence of *Flipper* could not be determined in 13 of the 184 isolates. The *Flipper* element was successfully amplified in 143 of the 171 remaining isolates (83.6%). The frequency of *Flipper*+ isolates was not significantly different among the V, bR and iR populations ($\chi^2_2 = 4.8$, $P = 0.09$, Fig. 2a). In contrast, the prevalence of *Flipper* significantly differed according to geographic origin ($\chi^2_5 = 21.3$, $P = 0.0007$, Fig. 2b). The χ^2 test remained significant when the smallest population, Flandres, was removed ($\chi^2_4 = 12.3$, $P = 0.015$).

Standard population genetics approaches to assess the respective effects of geography and host plant on population structure

The hierarchical analysis of molecular variance was performed first using plant species as a grouping factor (Table 3). Although the main contribution to the genetic variance was due to the variation within sampling sites, the effect of plant species was significant ($P < 0.0001$) as well as the effect of sampling location within each plant species ($P < 0.0001$). When sampling location was used

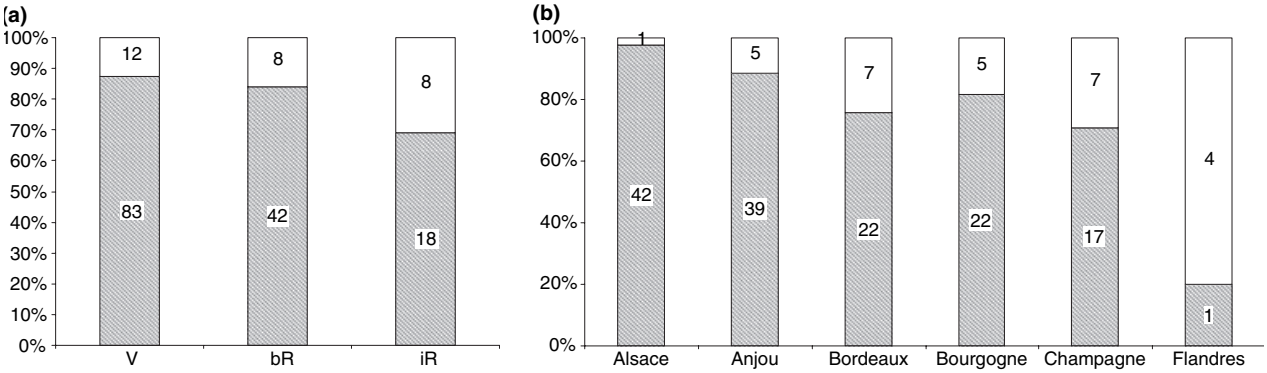

Fig. 2 Distribution of *Flipper*+ isolates (in grey) according to (a) host plant species and (b) geographic origin. V, isolates collected from *Vitis*; bR, isolates collected from *Rubus* bordering *Vitis* plantations; iR, isolates collected from insulated *Rubus*.

Table 3 Hierarchical analysis of molecular variance (AMOVA) with plant species and geographic location as a grouping factor.

Source of variation	d.f.	Sum of squares	Variance component	Variation (%)	*P*
Plant species	1	10.953	0.08150	2.87	<0.0001
Region within plant species	8	38.389	0.14689	5.18	<0.0001
Within region	146	380.721	2.60768	91.95	0.0059
Total	155	430.064	2.83607	100	
Region	4	20.049	-0.07328	-2.63	0.8533
Plant species within region	6	29.294	0.25623	9.18	<0.0001
Within plant species	146	380.721	2.60768	93.44	<0.0001
Total	155	430.064	2.79063	100	

as a grouping factor (Table 3), only the effect of plant species remained significant ($P < 0.0001$). Altogether, these results mean that the effect of plant species on the partitioning of total genetic variance is greater than the effect of geography and that the effect of plant species is significant in sympatry in each region.

F_{ST} values were computed for each pair of populations (Fig. 3). The highest values of F_{ST} were consistently found between populations from different hosts, either in sympatry or not. The means ± SD of F_{ST} were 0.101 ± 0.057 between V and bR populations (15 significant F_{ST} values of 25) and 0.136 ± 0.06 between V and iR populations (12 significant F_{ST} values of 25), whereas it was only 0.03 ± 0.048 between bR and iR populations (3 significant F_{ST} values of 25). Pairs of populations from different hosts in sympatry, i.e. V and bR isolates from the same localities, were always significantly differentiated, except in the Bourgogne region. Furthermore, in the four regions where V, bR and iR populations were all available, F_{ST} values between V and bR were always higher than those between bR and iR. Altogether, these results further indicate that the plant effect is stronger than the geographic effect on the population structure of *B. cinerea* Group II: populations collected from different but sympatric hosts were genetically more differentiated than populations coming from geographically distant but identical host species.

Between geographically distant populations from the same host, the means ± SD of F_{ST} were 0.058 ± 0.037, 0.045 ± 0.030 and 0.033 ± 0.050 for V, bR and iR (with 4, 3 and 0 significant F_{ST} values of 10 respectively). The geographic differentiation appeared greater among *Botrytis* populations coming from *Vitis* than among those from bordering or isolated *Rubus* but the F_{ST} values were not significantly different one from each other. One-sided tests performed to compare these statistics in a pairwise manner were indeed not significant (null hypothesis H_0: 'F_{ST} (bR) > F_{ST}(iR)': $P = 0.41$; H_0: 'F_{ST} (bR) > F_{ST}(V): $P = 0.57$; H_0: 'F_{ST}(iR) > F_{ST}(V): $P = 0.66$).

We also tested the existence of genetic differentiation between *Flipper*+ and *Flipper*- isolates by calculating pairwise F_{ST} in the populations on each host plants or geographic origin. When *Flipper*+ and *Flipper*- populations were separated according to the host of origin, only 5 of 15 F_{ST} values were significant, after progressive Bonferroni correction. When *Flipper*+ and *Flipper*- isolates were separated according to geography, 3 of 66 comparisons were significant. Altogether, these results indicate that there is no particular genetic differentiation between *Flipper*+ and *Flipper*- isolates, beyond the effect

	Vitis			Bordering *Rubus*					Isolated *Rubus*					
Bd-V	Bg-V	Ch-V	Al-V	An-bR	Bd-bR	Bg-bR	Ch-bR	Al-bR	An-iR	Bg-iR	Ch-iR	Al-iR	Fl-iR	
0.0360	0.0942*	0.0644*	0.0455	0.1735*	0.1307*	0.0790*	0.1482*	0.0607*	0.1905*	0.1068*	0.1228*	0.1985*	0.1917*	An-V
	0.0619	0.0300	0.0191	0.1008	0.0734*	0.0131	0.1027*	0.0300*	0.1090	0.0487	0.0316	0.1321	0.1254	Bd-V
		0.1556*	0.0595	0.2337	0.1276	0.0859	0.2376*	0.1365*	0.2017	0.1353	0.1210*	0.2418*	0.2733*	Bg-V
			0.0844*	0.1011	0.1136*	0.0622	0.1144*	0.0089	0.1363	0.0838	0.0951	0.1465*	0.1943*	Ch-V
				0.1262*	0.0907*	0.0271	0.1063	0.0583*	0.1355*	0.0397	0.0500	0.1270	0.1651*	Al-V
					0.0164	0.0363	–0.0072	0.0090	–0.0109	–0.0772	–0.0224	0.0087	0.0081	An-bR
						0.0631	0.0519	0.0542*	–0.0093	–0.0103	–0.0046	0.0842	0.0684	Bd-bR
							0.1049*	0.0340*	0.0932*	0.0297	0.0265	0.1043	0.1179*	Bg-bR
								0.0265	0.0468	–0.0245	0.0002	0.0631	0.0716	Ch-bR
									0.0508*	–0.0269	0.0276	0.0471	0.0976	Al-bR
										–0.0257	–0.0121	0.0740	0.0767	An-iR
											–0.0440	–0.0055	0.0268	Bg-iR
												0.0240	0.0447	Ch-iR
													0.1210	Al-iR

Fig. 3 Weir and Cockerham pairwise F_{ST} calculated with FSTAT. Stars indicate significant values (assessed after 2100 permutations and progressive Bonferroni correction). White panels: F_{ST} values between 0 and 0.1; light grey: F_{ST} values between 0.1 and 0.2; dark grey: F_{ST} values > 0.2.

of their different frequencies in the various geographic locations.

Genic differentiation was therefore tested across all population pairs, without considering the TE status of isolates. All the 105 pairwise comparisons were significant, the χ^2 statistics being computable for 31 and reaching infinity for the others. Interestingly, only 2 of the 31 computable comparisons (6.4%) involved V against bR or iR populations, whereas 48 of the 74 highly significant comparisons (64.9%) implied V against bR or iR populations. This further supported the idea that genetic differentiation was the highest between populations coming from different hosts.

Population graph analysis

The graph theoretical framework allowed assessing the genetic relationships between populations without any *a priori* hypothesis on population arrangements. In the resulting graph, the 15 nodes, i.e. populations, were connected by 19 edges. The iR and V populations appeared differentiated and the bR populations seemed to form an intermediate group between the iR and V populations (Fig. 4). We followed the method advocated by Dyer & Nason (2004) to test the existence of a restriction in gene flow between V and bR + iR populations on one hand and V + bR and iR populations on the other hand. In the graph, V populations (five nodes, four edges) were separated from bR + iR populations (ten nodes, eleven edges) by four intermediate edges. Under the null hypothesis of no differentiation, the probability of obtaining an edge connecting V and bR + iR populations equals 0.476 and the probability of observing four edges or fewer connecting V and iR + bR populations equals 0.0003. Similarly, V + bR populations (ten nodes, ten edges) are separated from iR populations (five nodes, five edges) by four intermediate edges; the probability of observing four edges or fewer connecting V + iR and bR populations was again equal to 0.0003. These significant results regarding the differentiation between V and iR + bR populations on the one hand, and V + bR and iR populations on the other hand, indicate again that the host plant acts as a significant source of restriction to gene flow. The graph analysis further suggests that bR populations probably result from admixture between V and iR populations.

Bayesian inference of population structure

We performed Bayesian clustering using the 184 isolates. The highest posterior probability given by the maximum likelihood distribution was obtained for $K = 10$, but the assignment rate for this value of K was only 57.37%. The best assignment rate (92.4%) was obtained for $K = 2$ (Fig. 5). The smallest cluster (43 individuals with an ancestry coefficient over 80% in this cluster) was composed of 88.4% iR + bR isolates and was therefore called the '*Rubus* cluster'. The other cluster contained 74% of V isolates and was therefore called the '*Vitis* cluster' (127 individuals with an ancestry coefficient over 80% in this cluster). The remaining 15 individuals (14 bR + iR and 1 V) could be considered as admixed because they could not be assigned to either cluster. This analysis further supported the existence of two genetically distinct clusters of individuals, separated mainly by their host of origin. To confirm the genetic differentiation of the populations on the two host plants, we performed Bayesian assignment analyses separately in the three regions where V and bR isolates were equally represented, i.e. Alsace, Bordeaux and Bourgogne (Table 4), without considering iR isolates. In each region, the best assignment rate was always obtained for $K = 2$, resulting in a *Rubus* cluster (where the proportion of bR isolates was 87.5%, 80% and 87.5% for Alsace, Bordeaux and

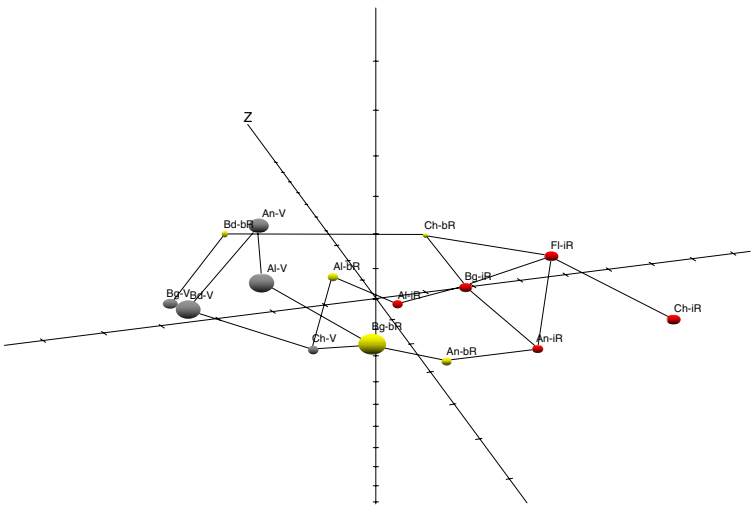

Fig. 4 Analysis of the inter-population genetic structure using the graph theoretical framework implemented in the PopGraph programme. Circles represent the different populations and have sizes proportional to effectives. (Online edition of article: grey=V populations, yellow=bR populations, red=iR populations).

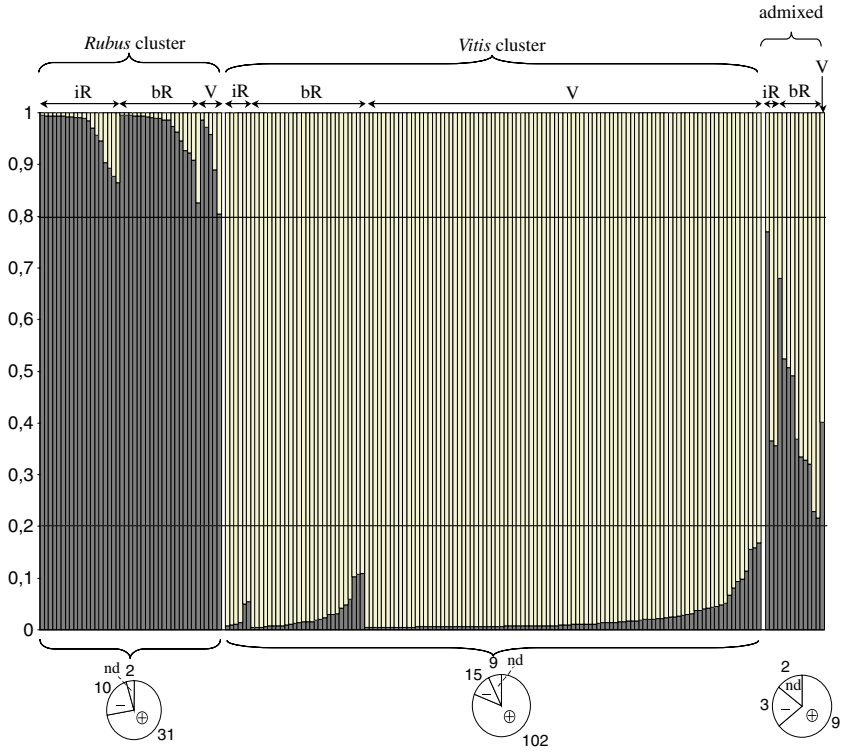

Fig. 5 Bayesian assignment of individuals into two clusters. Black (light grey) areas of vertical columns indicate the probability of ancestry in cluster 1 (cluster 2, respectively). The black lines represent the 20% and 80% thresholds of probability of ancestry to be assigned to one cluster. Below is indicated the TE status of isolates composing the two clusters (⊕ : *Flipper*+, –: *Flipper*-, nd: undetermined).

Bourgogne respectively) and a *Vitis* cluster (where the proportion of V isolates was 61.5%, 80.9% and 75% for Alsace, Bordeaux and Bourgogne respectively). To further investigate the structure of pathogen populations within each host, we analyzed separately the *Vitis* and *Rubus* clusters obtained from the first Bayesian analysis.

Table 4 Bayesian analysis using the Structure programme, performed separately in the regions of Alsace, Bordeaux and Bourgogne.

	Alsace	Bordeaux	Bourgogne
Nb of V isolates/nb of bR isolates	18/18	20/12	13/15
K with best posterior probability	5	5	5
K with best assignment rate	2	2	2
Composition of subclusters			
'Vitis' subcluster:			
Total number of isolates assigned to this cluster	26	21	16
V isolates	16	17	12
R isolates	10	4	4
R isolates assigned to the '*Vitis*' cluster	10	4	4
'Rubus' subcluster:			
Total number of isolates assigned to this cluster	8	10	8
V isolates	1	2	1
R isolates	7	8	7
V isolates assigned to the '*Rubus*' cluster	1	2	1

K, number of clusters.

For the *Rubus* cluster, the posterior probability peaked at $K = 6$, but only 67.4% individuals were correctly assigned. Again, the best assignment rate (95.2%) was obtained for $K = 2$, resulting in two subclusters, R1 and R2. These subclusters did not differ significantly in their composition regarding the TE status ($\chi^2_2 = 0.01$, $P = 0.99$). For the *Vitis* cluster, the most probable K was 7, but assignment rate was only 48%. The best assignment rate (83.5%) was for $K = 2$, resulting in two subclusters V1 and V2. To assess whether the subclusters corresponded to geographic differentiation, we tested different geographic subdivisions (West/East, North/-South, North-West/South-East, North-East/South-West). A marginally significant west-east geographic gradient was observed within the '*Rubus*' cluster but not within the '*Vitis*' cluster (Fig. 1): the subcluster R1 (R2) tended to include significantly more (less, respectively) individuals coming from Anjou, Bordeaux and Bourgogne than from Flandres, Champagne and Alsace ($\chi^2_1 = 5.47$, $P = 0.019$), with a threshold probability of 0.0125 for significance, whereas within the '*Vitis*' cluster the numbers of individuals on the two sides of the separation line were not significantly different in subclusters V1 and V2 ($\chi^2_1 = 0.51$, $P = 0.476$).

Discussion

Before the use of neutral genetic markers, the reproductive mode of the fungus *B. cinerea* was thought to be purely asexual, because of the huge number of conidias

produced by mature vegetative mycelium in vineyards. Studies of the genetic structure of French populations on grapes based on PCR-RFLP (Giraud *et al.*, 1997), however, revealed that the level of linkage disequilibrium in these populations was consistent with regular recombination. The present work, using more powerful genetic markers, confirms the regular occurrence of sexual reproduction within *B. cinerea* Group II populations in France, evidenced by the lack of genetic disequilibrium among loci and by the absence of over-represented multilocus genotypes. In a recent paper, Ma & Michailides (2005) studied the genetic structure of *B. cinerea* Californian populations on different hosts using microsatellite primed (MP)-PCR and concluded in contrast that their populations may have been largely clonal. However, the MP-PCR technique uses primers designed in the core repeat of microsatellite loci [e.g. $(AAC)_8$, $(AAG)_8$…], thereby amplifying the interval between two inversely orientated microsatellite loci. Therefore, MP-PCR markers reveal mostly the presence/absence of microsatellite loci between different strains, which is much less variable than the length of the microsatellite loci that we used here. Moreover, amplicons generated with MP-PCR are generally separated with low-resolution techniques (1.5% agarose gels in the case of Ma & Michailides, 2005), which is likely to yield homoplasy. We therefore believe that there is strong evidence for the predominance of recombination in *B. cinerea* Group II populations.

To determine the presence or absence of TEs in the sampled isolates, we focused on the detection of the *Flipper* element following the PCR-based method proposed by Ma & Michailides (2005). *Flipper* was successfully amplified in 83.6% of the 171 tested isolates, confirming the prevalence of *transposa* isolates at the end of the season already observed by several authors (Giraud *et al.*, 1997; Ma & Michailides, 2005; Martinez *et al.*, 2005). We did not detect significant differences in the distribution of *Flipper*+ isolates on the different host plants, whereas their distribution differed significantly according to the geographic location. Beyond these differences of geographic distribution, we did not detect genetic differentiation between *Flipper*+ and *Flipper*− isolates, suggesting unrestricted gene flow.

The results of our study show that geography does not greatly differentiate the *B. cinerea* Group II populations, at least at the studied scale, indicating great spore dispersal ability. In contrast, host plant significantly restricts gene flow within *B. cinerea* Group II. Although the barrier to gene flow between populations from different hosts was not complete [F_{ST} of 0.10 (0.13) between V and bR (V and iR, respectively) populations], it was detectable even in sympatry. The PopGraph analysis further showed that bR populations were intermediate between V and iR populations. Bayesian analysis without any *a priori* partition lead to consistent results: the assignment rate of isolates was optimized with two clusters, one cluster being composed of a majority of isolates coming from *Vitis* and the other of a majority of *Rubus* isolates. A small part of the sample could not be assigned to any cluster, suggesting that some individuals may result from admixture (recombination) between the two populations. Altogether, these results strongly indicate that *B. cinerea* Group II populations are structured according to the host plants, which means that isolates from the same host exchange genes more frequently with each other than with isolates from another host. This is consistent with previous results using PCR-RFLP markers on different hosts (Giraud *et al.*, 1999; Muñoz *et al.*, 2002).

The Bayesian analysis further suggested the existence of an asymmetry in the direction of gene flow between the populations from the two hosts. Indeed, of the 99 assigned V isolates, five (5%) were assigned to the cluster, showing that their host of origin was likely to be *Rubus* despite the fact that there were sampled on *Vitis* (Table 4). These 'misassigned' isolates may be considered as migrants from the *Rubus* population towards *Vitis* hosts. Similarly, of the 61 iR + bR assigned isolates, 33 (54%) were assigned to the '*Vitis*' cluster but were in fact sampled on *Rubus*, thus likely representing migrants from the *Vitis* pool towards *Rubus* (Table 4). Similar results were obtained for the three regions analyzed separately (Alsace, Bordeaux and Bourgogne). Thus, dispersal between pathogen populations developing on different hosts may be asymetric, i.e. larger from *Vitis* hosts towards *Rubus* hosts than the reverse, even in sympatry. This makes sense in the view of the large inoculum pressure produced in vineyard compared with that produced on wild hosts. On the contrary, the 15 isolates (14 iR or bR, and 1 V) that could not be assigned to any of the two clusters may represent admixed individuals. However, these results may also be due to the fact that the eight microsatellites used in this study are not powerful enough to correctly assign every individual, and thus have to be interpreted carefully.

Our results raise the question of what mechanisms can explain the observed restriction in gene flow between the pathogen populations from different hosts in sympatry. As both *Vitis* and *Rubus* are permanent plants, this is likely not due to a difference in the phenology of the hosts. Because the populations are not clonal, only a mechanism of preferential crosses between isolates from the same host can explain the observed pattern. In *B. cinerea*, ascospore dispersal is strictly passive, occurring via rain, wind or animals. Ascospores then germ and develop on the substrate where they land, if they are able to. The resulting mycelium then produces asexual spores. Therefore, there cannot be active choice of the habitat by the fungus, but there is a limitation of dispersal between development on a particular host and the mating process. Merely because of this biology departing from free-living organisms, *B. cinerea* individuals are forced to mate on the substrate where they initially developed, which is also true for many other filamentous fungi. The model

developed by Giraud *et al.* (2006) showed that under these conditions, if there is a sufficient level of selection by the host, a significant reduction of gene exchange between pathogens from the different hosts can rapidly be reached, even without active assortative mating. Alternatively, some extant of assortative mating could have been selected for between isolates adapted to different hosts if a strong disruptive selection exists (Kawecki & Ebert, 2004). The exact mechanism allowing gene flow restriction between isolates from *Vitis* and *Rubus* remains to be explored using pathogenicity tests and mating experiments.

Acknowledgments

We thank A. Gautier, O. Jonot, A-S. Walker, R. Meziani, E. Balsemin and Y. Brygoo for technical assistance and useful discussions. We thank G. Aguileta for comments on an earlier draft of the manuscript. This work was supported by the Bureau des Ressources Génétiques, as well as an 'ACI Jeunes chercheurs' Grant.

References

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

Albertini, C., Thébaud, G., Fournier, E. & Leroux, P. 2002. Eburicol 14a-demethylase gene (cyp51) polymorphism and speciation in *Botrytis cinerea*. *Mycol. Res.* **106**: 1171–1178.

Antonovics, J., Hood, M.E. & Partain, J. 2002. The ecology and genetics of a host-shift: *Microbotryum* as a model system. *Am. Nat.* **160**: S40–S53.

Brown, A.H.D., Feldman, M.W. & Nevo, E. 1980. Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics* **96**: 523–536.

Bucheli, E., Gautschi, B. & Shykoff, J.A. 2001. Differences in population structure of the anther smut fungus *Microbotryum violaceum* on two closely related host species, *Silene latifolia* and *S. dioica*. *Mol. Ecol.* **10**: 285–294.

Bush, G.L. 1994. Sympatric speciation in animals: new wine in old bottles. *Trends Ecol. Evol.* **9**: 285–288.

Coyne, J.A. & Orr, H.A. 1999. The evolutionary genetics of speciation. In: *Evolution of Biological Diversity* (A. E. Magguran & R. M. Ray, eds), pp. 1–36. Oxford University Press, Oxford, UK.

Dickinson, H. & Antonovics, J. 1973. Theoretical considerations of sympatric divergence. *Am. Nat.* **107**: 256–274.

Diehl, S.R. & Bush, G.L. 1989. The role of habitat preference in adaptation and speciation. In: *Speciation and its Consequences* (D. Otte & J. A. Endler, eds), Sinauer, Sunderland, MA, pp. 345–365.

Diolez, A., Marches, F., Fortini, D. & Brygoo, Y. 1995. Boty, a long-terminal-repeat retroelement in the phytopathogenic fungus *Botrytis cinerea*. *Appl. Environ. Microbiol.* **61**: 103–108.

Dyer, R.J. & Nason, J.D. 2004. Population graphs: the graph theoretic shape of genetic structure. *Mol. Ecol.* **13**: 1713–1727.

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

Felsenstein, J. 1981. Scepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**: 124–138.

Fournier, E., Giraud, T., Loiseau, A., Vautrin, D., Estoup, A., Solignac, M., Cornuet, J.M. & Brygoo, Y. 2002. Characterization of nine polymorphic microsatellite loci in the phytopathogenic fungus *Botrytis cinerea* (Ascomycota). *Mol. Ecol. Notes* **2**: 253–255.

Fournier, E., Lévis, C., Fortini, D., Giraud, T., Leroux, P. & Brygoo, Y. 2003. Characterization of Bc-*hch*, the *Botrytis cinerea* homolog of the *Neurospora crassa het-c* vegetative incompatibility locus, and its use as a population marker. *Mycologia* **95**: 951–961.

Fournier, E., Giraud, T. & Brygoo, Y. 2005. Partition of the *Botrytis cinerea* complex in France using multiple gene genealogies. *Mycologia* **97**: 1251–1267.

Frantz, A., Plantegenest, M. & Simon, J.-C. 2006. Ecological specialization correlates with genotypic differentiation in sympatric host-populations of the pea aphid. *J. Evol. Biol.* **19**: 392–401.

Gavrilets, S. 2004. *Fitness Landscapes and the Origin of Species*. Monographs in population biology. Princeton University Press, Princeton, NJ.

Giraud, T. 2004. Patterns of within population dispersion and mating of the fungus *Microbotryum violaceum* parasitizing the plant *Silene latifolia*. *Heredity* **93**: 559–565.

Giraud, T. 2006. Speciation: selection against migrant pathogens: the immigrant inviability barrier in pathogens. *Heredity* **97**: 316–318.

Giraud, T., Fortini, D., Lévis, C., Leroux, P. & Brygoo, Y. 1997. RFLP markers show genetic recombination in *Botryotinita fuckeliana* (*Botrytis cinerea*) and transposable elements reveal two sympatric species. *Mol. Biol. Evol.* **14**: 1177–1185.

Giraud, T., Fortini, D., Lévis, C., Lamarque, C., Leroux, P., LoBuglio, K. & Brygoo, Y. 1999. Two sibling species of the *Botrytis cinerea* complex, *transposa* and *vacuma*, are found in sympatry on numerous host plants. *Phytopathology* **89**: 967–973.

Giraud, T., Villaréal, L.M.M.A., Austerlitz, F., Le Gac, M. & Lavigne, C. 2006. Importance of the life cycle in sympatric host race formation and speciation of pathogens. *Phytopathology* **96**: 280–287.

Goudet, J. 1995. FSTAT (vers. 1.2): a computer program to calculate *F*-statistics. *J. Heredity* **86**: 485–486.

Haubold, B., Travisiano, P.B. & Hudson, R.R. 1998. Detecting linkage disequilibrium in bacterial populations. *Genetics* **150**: 1341–1348.

Johnson, P.A., Hoppensteadt, F.C., Smith, J.J. & Bush, G.L. 1996. Conditions for sympatric speciation: a diploid model incorporating habitat fidelity and non-habitat assortative mating. *Evol. Ecol.* **10**: 187–205.

Kawecki, T.J. 1998. Red Queen meets Santa Rosalia: arms races and the evolution of host specialization in organisms with parasitic lifestyles. *Am. Nat.* **152**: 635–651.

Kawecki, T. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecol. Lett.* **7**: 1225–1241.

Le Cam, B., Parisi, L. & Arene, L. 2002. Evidence of two formae speciales in *Venturia inaequalis*, responsible for apple and pyracantha scab. *Phytopathology* **92**: 314–320.

Le Gac, M., Hood, M.E., Fournier, E. & Giraud, T. 2007. Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. *Evolution* **16**: 15–26.

Lévis, C., Fortini, D. & Brygoo, Y. 1997. Flipper, a mobile Fot1-like transposable element in *Botrytis cinerea*. *Mol. Gen. Genet.* **254**(6): 674–680.

Lopez-Villavicencio, M., Enjalbert, J., Hood, M.E., Shykoff, J.A., Raquin, C. & Giraud, T. 2005. The anther smut disease on *Gypsophila repens*: a case of parasite sub-optimal performance following a recent host shift? *J. Evol. Biol.* **18**: 1293–1303.

Ma, Z. & Michailides, T.J. 2005. Genetic structure of *Botrytis cinerea* populations from different host plants in California. *Plant Dis.* **89**: 1083–1089.

Martinez, F., Dubos, B. & Fermaud, M. 2005. The role of saprotrophy and virulence in the population dynamics of *Botrytis cinerea* in vineyards. *Phytopathology* **95**: 692–700.

Matsuda, S. & Takamatsu, S. 2003. Evolution of host-parasite relationships of Golovinomyces (Ascomycete: Erysiphaceae) inferred from nuclear rDNA sequences. *Mol. Phylogenet. Evol.* **27**: 314–327.

Maynard-Smith, J., Smith, N.H., O'Rourke, M. & Spratt, B.G. 1993. How clonal are bacteria? *Proc. Natl. Acad. Sci. USA* **90**: 4384–4388.

Mayr, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge.

Muñoz, G., Hinrichsen, P., Brygoo, Y. & Giraud, T. 2002. Genetic characterisation of *Botrytis cinerea* populations in Chile. *Mycol. Res.* **106**: 594–601.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Pritchard, J.K., Stephens, P. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Rausher, M.D. 1984. The evolution of habitat preference in subdivided populations. *Evolution* **38**: 596–608.

Rice, W.R. 1984. Disruptive selection on habitat preference and the evolution of reproductive isolation: a simulation study. *Evolution* **38**: 1251–1260.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223–225.

Schardl, C.L., Leuchtmann, A., Chung, K.-R., Penny, D. & Siegel, M.R. 1997. Coevolution by common descent of fungal symbiont (*Epichloë* spp.) and grass hosts. *Mol. Biol. Evol.* **14**: 133–143.

Schneider, S., Roessli, D. & Excoffier, L. 2000. *Arlequin ver 2.000: A Software for Population Genetics Analysis*. Genetic and Biometry Laboratory, University of Geneva, Switzerland.

Udovic, D. 1980. Frequency-dependent selection, disruptive selection and the evolution of reproductive isolation. *Am. Nat.* **116**: 621–641.

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

Received 12 April 2007; revised 20 September 2007; accepted 4 October 2007