



**HAL**  
open science

# The genetic structure of the plant pathogenic fungus *Melampsora larici-populina* on its wild host is extensively impacted by host domestication

Constance Xhaard, Bénédicte Fabre, Axelle Andrieux, P. Gladieux, Benoît Barres, Pascal Frey, Fabien Halkett

## ► To cite this version:

Constance Xhaard, Bénédicte Fabre, Axelle Andrieux, P. Gladieux, Benoît Barres, et al.. The genetic structure of the plant pathogenic fungus *Melampsora larici-populina* on its wild host is extensively impacted by host domestication. *Molecular Ecology*, 2011, 20 (13), pp.2739-2755. 10.1111/j.1365-294X.2011.05138.x . hal-01136973

**HAL Id: hal-01136973**

**<https://hal.science/hal-01136973>**

Submitted on 30 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The genetic structure of the plant pathogenic fungus *Melampsora larici-populina* on its wild host is extensively impacted by host domestication

C. XHAARD,\* B. FABRE,\* A. ANDRIEUX,\* P. GLADIEUX,† B. BARRÈS,¶ P. FREY\*<sup>1</sup> and F. HALKETT\*<sup>1</sup>

\*INRA, Nancy-Université, UMR 1136 Interactions Arbres-Microorganismes, F-54280 Champenoux, France, †Laboratoire Ecologie, Systématique et Evolution, Univ. Paris Sud, UMR8079, Orsay, France, CNRS, Orsay, France, AgroParisTech, Orsay, France, ¶INRA, UMR Biodiversité Gènes et Communautés, Domaine de Pierroton, F-33662 Cestas, France

## Abstract

Wild and cultivated plants represent very different habitats for pathogens, especially when cultivated plants bear qualitative resistance genes. Here, we investigated to what extent the population genetic structure of a plant pathogenic fungus collected on its wild host can be impacted by the deployment of resistant cultivars. We studied one of the main poplar diseases, poplar rust, caused by the fungus *Melampsora larici-populina*. A thousand and fifty individuals sampled from several locations in France were phenotyped for their virulence profile (ability to infect or not the most deployed resistant cultivar 'Beaupré'), and a subset of these was genotyped using 25 microsatellite markers. Bayesian assignment tests on genetic data clustered the 476 genotyped individuals into three genetic groups. Group 1 gathered most virulent individuals and displayed evidence for selection and drastic demographic changes resulting from breakdown of the poplar cultivar 'Beaupré'. Group 2 comprised individuals corresponding to ancestral populations of *M. larici-populina* naturally occurring in the native range. Group 3 displayed the hallmarks of strict asexual reproduction, which has never previously been demonstrated in this species. We discuss how poplar cultivation has influenced the spatial and genetic structure of this plant pathogenic fungus, and has led to the spread of virulence alleles (gene swamping) in *M. larici-populina* populations evolving on the wild host.

**Keywords:** clustering analysis, demographic history, disease spread, gene flow, rust fungus, selection

Received 24 August 2010; revision received 23 March 2011; accepted 31 March 2011

## Introduction

Human activities have become a major source of environmental and ecological changes. A particular example of this is the domestication of animals and plants, one of the most important technological innovations in human history (Purugganan & Fuller 2009). The domestication process led to marked shifts in the evolutionary dynamics of host–pathogen interactions in both animals (Daszak *et al.* 2000; Diamond 2002) and plants (Stukenbrock *et al.* 2007), and resulted in tremendous changes in pathogen popula-

tion structure (Lebarbenchon *et al.* 2008; Stukenbrock & McDonald 2008; Giraud *et al.* 2010).

Some of the most devastating epidemics on agricultural crops were caused by pathogens that shifted from a (more or less related) wild host to a domesticated one (Anderson *et al.* 2004; Parker & Gilbert 2004; Desprez-Loustau *et al.* 2007). Examples of fungal (sensu lato, i.e. including oomycetes) plant diseases caused by host-shifts include rice blast (*Magnaporthe oryzae*; Couch *et al.* 2005), potato late blight (*Phytophthora infestans*; Gómez-Alpizar *et al.* 2007), wheat leaf blotch (*Mycosphaerella graminicola*; Stukenbrock *et al.* 2007) and leaf scald of barley (*Rhynchosporium commune*; Zaffarano *et al.* 2008). Alternatively, some pathogens appear to have followed their host along the domestication course through a

Correspondence: Fabien Halkett, Fax: +33 3 83 39 40 69; E-mail: halkett@nancy.inra.fr

<sup>1</sup>These authors co-directed this work.

process called host-tracking (e.g. *Ustilago maydis* on maize: Munkacsı *et al.* 2008; *Venturia inaequalis* on apple tree: Gladieux *et al.* 2008, 2010a). A drastic reduction in gene flow accompanying the host-shift and subsequent specialization on the domesticated plant can lead to the rapid formation of new pathogen species (Giraud *et al.* 2010). Even in the case of host-tracking, where wild and cultivated hosts are closely related, a strong genetic differentiation between populations from wild and cultivated hosts can be promoted by environmental differences existing between natural and domesticated habitats (Lebarbenchon *et al.* 2008; Stukenbrock & McDonald 2008).

In this context of pathogen evolution during domestication of cultivated plants, numerous studies have aimed to retrace the domestication routes of crop pathogens (Couch *et al.* 2005; Munkacsı *et al.* 2008; Zaffarano *et al.* 2008). However, only few studies interested in host-tracking pathogens have focused on gene flows occurring between pathogens from wild and cultivated plants (but see Kiros-Meles *et al.* 2011), certainly because of the difficulty to sample on wild host plants (Gladieux *et al.* 2008). Wild host plants can play the role of a pathogen reservoir (Wisler & Norris 2005; Lê Van *et al.* 2011), especially when both cultivated and wild hosts are sympatric. The wild hosts can be the source of new virulences that can cause drastic epidemics on cultivated plants (Couch *et al.* 2005; Gomez-Alpizar *et al.* 2007; Zaffarano *et al.* 2008) and also facilitate connectivity between crops. Conversely, pathogens from cultivated plants can spread onto wild plants and compete with pathogens from wild hosts, which can lead to extensive gene flow, including pathogenicity-related genes, and in turn an increase in the damages caused by pathogens on the wild plants (Rizzo *et al.* 2005).

Agricultural landscapes, especially since the development of intensive agricultural practices, offer new and more favourable environments for the development of plant pathogens (Yarwood 1970; Hansen 1987). Contrary to wild host populations, which are generally characterized by low plant density and high genetic diversity (Burdon 1993; Gilbert 2002; Stukenbrock *et al.* 2007; Burdon & Thrall 2008), human-managed agroecosystems are characterized by high plant density and a reduced genetic diversity, which enhances pathogen multiplication by favouring transmission between infected and uninfected plants (Burdon & Chilvers 1982; Stukenbrock & McDonald 2008). Therefore, plant-pathogen interactions in the wild have often been depicted as benign associations in which serious epidemics rarely occur, whereas differences in the physical properties and biology of human-managed environments can lead to devastating epidemics in agricultural

landscapes (Burdon 1993). Together with adaptation to host plants, differences in population dynamics between wild and domesticated ecosystems can not only impact pathogen population structure but also leave specific genetic signatures resulting from the demographic fluctuations, as was observed in meta-population systems with different host lifespans in the anther smut fungus *Microbotryum violaceum* (Bucheli *et al.* 2001).

Poplar rust is a suitable model with which to study the interactions between pathogens evolving on wild and cultivated hosts, in particular because wild and cultivated poplars grow in sympatry in most parts of the Northern Hemisphere. The sympatry of wild and cultivated hosts enables both historical and contemporary gene flow to be assessed in the absence of confounding effect of large geographical distances on pathogen dispersal. Assuming recurrent gene flow exists between wild and cultivated hosts, it is thus possible to assess the consequences of the presence of cultivated hosts on the genetic structure of pathogen populations sampled on the wild host. Wild and cultivated poplars differ in their ecological characteristics, and thus in the nature of the niche they provide to the pathogen. On the one hand, the native Eurasian black poplar, *Populus nigra*, is a pioneer tree species of riparian forests that is distributed widely across Europe and Asia (Imbert & Lefevre 2003). Qualitative resistance to the Eurasian poplar rust fungus *Melampsora larici-populina* has not been detected so far in *P. nigra* (Pinon & Frey 2005), and rust incidence on wild poplar stands is generally low (Gérard *et al.* 2006). On the other hand, cultivated poplar stands are particularly exposed to the attacks of many pathogens because of their intensive monoclonal cultivation over several decades (Gérard *et al.* 2006). In Europe, rust caused by *M. larici-populina* is the main disease on cultivated poplars and causes severe damage and economic losses in poplar cultivation (Frey *et al.* 2005). Most of the rust-resistant poplar cultivars released so far by European poplar breeders have been shown to carry qualitative resistance genes, all of which have been overcome by *M. larici-populina* (Pinon & Frey 2005). The most damaging resistance breakdown occurred in the hybrid cultivar *Populus × interamericana* 'Beaupré', which carries the R7 resistance gene. This cultivar showed high yield performance and was grown successfully between 1980 and 1994. During this period, 'Beaupré' was the poplar cultivar planted most in northwestern Europe, accounting for 80% of new plantations in northern France (Pinon *et al.* 1998). In 1994, breakdown of the R7 resistance was detected in Belgium and northern France, and virulent (Vir7) *M. larici-populina* individuals spread all over Western Europe in <5 years, causing very destructive epidemics (Pinon *et al.* 1998; Pinon & Frey 2005). Contrary to a meta-

population system (or annual crop fields) that displays frequent extinction/recolonization by pathogens and rapid turn over of pathogen virulence (McDonald 2010), a perennial plant like a poplar tree exhibits a much higher longevity. Even if plantation of 'Beaupré' decreased after 1997 and totally ceased by 2005, this cultivar still predominates in poplar cultivated landscape in northern France, thus exerting a long lasting and constant selection pressure on *M. larici-populina* populations through the deployment of R7 resistance.

The rust fungus *M. larici-populina* is an obligate biotrophic pathogen belonging to the order Pucciniales (Basidiomycetes). It is a macrocyclic heteroecious fungus that produces five spore stages during its life cycle and which infects two host species. The sexual stage occurs on larch needles (*Larix* sp.) in spring—the impact on larch being insignificant. Then, *M. larici-populina* infects poplar leaves, on which many rounds of asexual multiplication occur between spring and autumn until leaf fall. The fungus then develops an overwintering form on senescent poplar leaves. In spring, the life cycle restarts with the infection of newly formed larch needles. The sexual stage on larch ensures a huge reshuffling of genetic material among individuals each year. Although native to the mountains of Central Europe (Alps, Carpathians, Sudetes, Tatras) and lowlands in northern Poland, the European larch (*Larix decidua*) has been widely planted in lowlands throughout Europe since the eighteenth century for timber and ornamental purposes. Given the overall distribution of larches, high levels of genotypic diversity are observed in all locations where *M. larici-populina* is sampled, even when individuals were collected on poplar at the end of the epidemic phase (Gérard *et al.* 2006; Pei *et al.* 2007; Barrès *et al.* 2008), in accordance with a predominantly sexual reproductive mode (Halkett *et al.* 2005a). Similar results were obtained with the North-American poplar rust fungus *Melampsora medusae* f. sp. *deltoidae*, when individuals were collected in poplar-larch sympatry area (Bourassa *et al.* 2007). In addition, host alternation and wind-dispersed spores foster gene flow across host plants (including between wild and cultivated poplars) and locations. Previous population genetic studies on *M. larici-populina* revealed a very low level of genetic differentiation between locations and demonstrated the lack of a relationship between geographical and genetic distances, both in France (Gérard *et al.* 2006; Barrès *et al.* 2008) and in the UK (Pei *et al.* 2007). Nonetheless, at a continental scale, isolation by distance was tentatively observed (Barrès *et al.* 2008), confirming a gradual dispersion of spores between locations. Even at this large spatial scale,  $F_{ST}$  values remained weak, indicating that even if dispersal of the pathogen is gradual, there is substantial gene mixing over the European continent.

Concerning phenotypic characteristics, no relationship between virulence profile and genotype, or between virulence profile and geographical location, has been evidenced (Gérard *et al.* 2006; Barrès *et al.* 2008). The lack of relationship between phenotype and genotype has been proposed to result from the extensive genetic reshuffling imposed by the sexual stage in larch.

Here we aimed to benefit from a very peculiar ecological situation (the large and long lasting selection pressure exerted by R7 cultivated poplars on the rust fungus *M. larici-populina*) to investigate the putative consequences of the impact of host domestication on population genetic structure of a plant pathogen on its wild host. We sampled more than twenty sites across France, where large differences in the acreage of R7 cultivated poplars exist between northern and southern locations. The individuals collected were genotyped using 25 polymorphic microsatellite markers and phenotyped for their Vir7 profile. The use of both putatively neutral and selected markers provided an opportunity to examine the relationship between population genetic structure and a key biological property (Barrett *et al.* 2008b), here the ability of *M. larici-populina* individuals to grow or not on R7 cultivated poplars. We first employed Bayesian clustering analyses to infer population structure in a set of *M. larici-populina* individuals collected on the wild host *P. nigra*. We then examined (i) whether the genetic differentiation matches the ability to grow on cultivated plants, (ii) whether identified genetic groups of the pathogen display a particular pattern of spatial distribution and how this relates to differences in cultivated host densities and finally (iii) whether genetic groups of the pathogen display differences in patterns of genetic variability consistent with a history of strong selection by host resistance and subsequent changes in demography.

## Materials and methods

### Sampling strategy

Poplar leaves infected with *M. larici-populina* were collected on *P. nigra* in 23 locations across France (and Aoste, Italy) between August and November 2009 (Table 1). Sampling was performed on wild riparian poplar stands, except for four locations (Amance, Charrey-sur-Saône, Guéméné-Penfao and Orléans) where wild *P. nigra* trees were grown in experimental poplar nurseries. These poplar nurseries also contained many different poplar cultivars harbouring eight distinct qualitative resistances. As no qualitative resistance to *M. larici-populina* has been reported so far in *P. nigra* (Pinon & Frey 2005), collecting on *P. nigra*—whatever the

**Table 1** Characteristics of sampling sites. The number in brackets after the number of genotyped individuals indicated the number of individuals with more than six loci that failed to amplify; G/N, Genotypic richness

Site No	Location	Latitude	Longitude	No of individuals phenotyped	No of individuals genotyped	G/N
1	Amance*	48°45'N	6°20'E	80	15 (1)	1.00
2	Angers	47°23'N	0°39'W	26	16 (0)	0.75
3	Aoste (Italy)	45°46'N	7°18'E	16	14 (0)	0.93
4	Changy	46°09'N	3°53'E	46	14 (2)	0.93
5	Charrey-sur-Saône*	47°05'N	5°09'E	30	16 (0)	1.00
6	Croix	50°41'N	3°08'E	21	11 (5)	1.00
7	Guémené-Penfao*	47°38'N	1°54'W	61	15 (1)	0.93
8	Mirabeau	43°41'N	5°04'E	108	83 (13)	0.81
9	Montpellier	43°41'N	3°52'E	30	24 (2)	0.63
10	Nancy	48°43'N	6°10'E	48	16 (0)	1.00
11	Nievroz	45°49'N	5°04'E	88	16 (0)	1.00
12	Prelles	44°51'N	6°35'E	99	90 (6)	1.00
13	Ramières 1	44°45'N	4°53'E	39	16 (0)	0.31
14	Ramières 2	44°44'N	4°58'E	52	15 (1)	0.80
15	Rennes	48°06'N	1°48'W	34	16 (0)	1.00
16	Roubaix	50°41'N	3°10'E	12	12 (0)	0.92
17	Seilh	43°42'N	1°22'E	16	11 (5)	0.82
18	Saint-Ay	47°52'N	1°46'E	108	15 (1)	0.33
19	Saint-Laurent-du-Var	43°40'N	7°12'E	15	14 (1)	0.21
20	Ungersheim	47°51'N	7°17'E	26	16 (0)	0.88
21	Veurey-Voroise	45°16'N	5°37'E	80	14 (2)	1.00
22	Villenave d'Ornon	44°47'N	0°34'W	15	15 (0)	0.47
23	Orléans*	47°54'N	1°54'E	120	–	–

\*Experimental poplar nursery.

location type — avoided any specific selection pressure resulting from the qualitative resistance exerted by some cultivated poplars (Gérard *et al.* 2006). Hence, because of the extensive gene flow in this species, each collection site was assumed to reflect the local composition of *M. larici-populina* genotypes. At each location, 15–150 rust-infected leaves were collected. The highest sampling effort was conducted at Prelles (Table 1), located in the original distribution area of *M. larici-populina*, where larch and wild poplar are sympatric. In the following, samples from this location are referred to as the reference population.

#### Individual isolation

One single uredinium (sporulating lesion of *M. larici-populina* producing urediniospores) per leaf was selected randomly to avoid excessive clonal replicates, and grown on fresh leaf discs of *P. × euramericana* 'Robusta' as described by Barrès *et al.* (2008). Mono-uredinial isolation ensures that a single genotype per sampled leaf is isolated and multiplied, so that both virulence phenotype assessment and molecular analysis can be performed on the same individual.

#### Virulence phenotype assessment

All the *M. larici-populina* individuals (Table 1) were inoculated on two poplar cultivars, one susceptible (*P. × euramericana* 'Robusta') and the second carrying the R7 rust resistance gene (*P. × interamericana* 'Beaupré'), as described by Barrès *et al.* (2008). Phenotype of each individual was denoted Vir7 if both resistant and susceptible leaf discs were infected, and Avr7 if only the susceptible (control) leaf disc was infected. For each location, we calculated the proportion of Vir7 individuals. After virulence identification, leaf discs of 'Robusta' were frozen and stored at –20 °C until DNA extraction.

#### Microsatellite analysis

A subsample of 12–96 individuals per sampling site was defined randomly for microsatellite analysis (Table 1). Only 22 (of the 23) locations were considered for genetic analyses. DNA was extracted using the BioSprint 96 DNA plant kit used in combination with the BioSprint 96 automated workstation (Qiagen) following the BS96-DNA-plant protocol. Microsatellite analyses were performed at 25 polymorphic loci: MLP12 (Barrès

*et al.* 2006), MLP49, MLP50, MLP54, MLP55, MLP56, MLP57, MLP58, MLP59, MLP66, MLP68, MLP71, MLP73, MLP77, MLP82, MLP83, MLP87, MLP91, MLP92, MLP93, MLP94, MLP95, MLP96, MLP97 (Xhaard *et al.* 2009) and MLP100 (this study). Microsatellite loci were amplified in a GeneAmp PCR system 9700 thermocycler (Applied Biosystems) under the following conditions: an initial denaturing step of 5 min at 95 °C; 32 cycles including 30 s of denaturation at 95 °C, 90 s annealing at 60 °C and 60 s extension at 72 °C; and a final 30 min extension step at 60 °C. PCR was carried out in a 10- $\mu$ l final reaction volume containing 5  $\mu$ l of Qiagen® Multiplex PCR kit, 2  $\mu$ l of DNA, 0.2  $\mu$ M of each reverse and forward primers. Forward primers were labelled with a fluorescent tag (FAM, NED, VIC or PET). Three multiplex PCR were run, comprising 9, 7 and 9 loci, respectively (Table 2). PCR products from the three multiplex reactions (3  $\mu$ l PCR1, 4  $\mu$ l PCR2 and 5  $\mu$ l PCR3) were pooled and loaded on an ABI 3730 Genetic Analyzer (Applied Biosystems). Fragments

were sized with a LIZ-1200 size standard, and alleles were scored using GENEMAPPER 4.0 (Applied Biosystems). Individuals for which more than six loci failed to amplify were omitted from further analyses.

#### Data analysis

*Identification of repeated multilocus genotypes.* When pooling all multilocus genotypes from all locations, we identified individuals carrying repeated multilocus genotypes. Given the life cycle of *M. larici-populina*, such repeated multilocus genotypes can result from clonal multiplication or can simply occur by chance given the allelic frequency in the population (Halkett *et al.* 2005a). To distinguish true clonal copies from individuals carrying the same multilocus genotype by chance, we used MLGSIM software (Stenberg *et al.* 2003) that computes the probability of observing a multilocus genotype  $n$  times in a population and performed Monte Carlo simulations (10 000 iterations) to determine the

**Table 2** Characteristics of the microsatellite loci used

Locus	Repeat unit	Allele range	Multiplex PCR No.	All individuals ( $n = 476$ )			Prelles ( $n = 90$ )			Reference
				A	$H_E$	$H_O$	A	$H_E$	$H_O$	
MLP12	Complex (AAG) <sub>n</sub>	217–265	2	9	0.688	0.542***	7	0.735	0.634 <sup>ns</sup>	1
MLP49	(GAT)18	344–407	3	19	0.732	0.731***	13	0.766	0.762 <sup>ns</sup>	2
MLP50	(TCT)16	262–297	2	9	0.614	0.633***	7	0.415	0.471*	2
MLP54	(ATG)14	126–141	1	6	0.551	0.416***	5	0.559	0.477 <sup>ns</sup>	2
MLP55	(ATC)15	133–167	1	11	0.521	0.517***	6	0.614	0.682***	2
MLP56	(AAC)7	269–306	2	11	0.625	0.594***	9	0.661	0.640 <sup>ns</sup>	2
MLP57	(CAA)13	154–178	1	6	0.302	0.304 <sup>ns</sup>	6	0.402	0.438 <sup>ns</sup>	2
MLP58	(AAG)12	235–259	2	13	0.692	0.553***	9	0.727	0.771 <sup>ns</sup>	2
MLP59	(ATC)13	305–335	2	12	0.760	0.625***	11	0.808	0.440 <sup>ns</sup>	2
MLP66	(GAT)12	189–213	1	7	0.688	0.678***	4	0.675	0.616 <sup>ns</sup>	2
MLP68	(TTGA)12	412–472	3	9	0.455	0.450***	7	0.332	0.310 <sup>ns</sup>	2
MLP71	(AG)14	358–376	2	8	0.429	0.443***	5	0.317	0.310***	2
MLP73	(AC)14	419–449	3	17	0.842	0.829***	11	0.822	0.738 <sup>ns</sup>	2
MLP77	(AT)10	431–437	3	6	0.374	0.336***	4	0.506	0.451 <sup>ns</sup>	2
MLP82	(TAC)10	172–220	1	16	0.797	0.621***	12	0.740	0.581 <sup>ns</sup>	2
MLP83	(ATG)10	145–163	1	7	0.476	0.453 <sup>ns</sup>	4	0.505	0.591 <sup>ns</sup>	2
MLP87	(TGT)8	338–353	3	6	0.465	0.370***	3	0.464	0.427 <sup>ns</sup>	2
MLP91	(GTT)10	125–134	1	5	0.184	0.202 <sup>ns</sup>	3	0.055	0.056 <sup>ns</sup>	2
MLP92	(TTG)11	324–336	2	5	0.267	0.127***	5	0.265	0.035 <sup>ns</sup>	2
MLP93	(GTT)11	125–149	1	7	0.572	0.629***	6	0.601	0.636***	2
MLP94	(TTC)10	452–477	3	10	0.666	0.592***	6	0.686	0.605 <sup>ns</sup>	2
MLP95	(ACA)9	441–469	3	14	0.743	0.633***	11	0.652	0.583 <sup>ns</sup>	2
MLP96	(TGG)9	403–413	3	5	0.558	0.478***	3	0.500	0.434 <sup>ns</sup>	2
MLP97	(GATT)9	380–416	3	8	0.627	0.680***	6	0.586	0.528 <sup>ns</sup>	2
MLP100	(ACA)18	178–231	1	14	0.706	0.509***	12	0.597	0.455 <sup>ns</sup>	3

A, total number of alleles;  $H_E$ , expected heterozygosity calculated according to Nei (1978);  $H_O$ , observed heterozygosity. Probability test for Hardy–Weinberg (HW) equilibrium: significant levels of  $P$  value were corrected for multiple comparisons using false discovery rate control (\* $Q < 0.05$ ; \*\*  $Q < 0.01$ ; \*\*\* $Q < 0.001$ ; ns, nonsignificant). References: (1) Barrès *et al.* 2006; (2) Xhaard *et al.* 2009; (3) this study.

probability threshold below which each repeated multilocus genotype can be significantly considered as a true clone (significance level was set to 0.05).

To minimize departure from Hardy–Weinberg (HW) expectations and linkage disequilibrium (LD) on which several analyses are based, we considered only one copy of each multilocus genotype (found to be significantly over-represented) for all subsequent genetic analyses. An exception is the computation of the genetic characteristics of genetic groups, for which summary statistics were calculated both with and without clonal copies.

*Clustering analysis.* We investigated population subdivision using the Bayesian clustering method implemented in GENELAND (Guillot *et al.* 2005b; Guillot 2008). GENELAND offers two types of analyses, taking –or not– information on the geographical origin of individuals into account to assign individuals to genetic groups. Assignment tests were conducted using both methods, which can be more or less sensitive to the within-site genetic admixture. Another strength of GENELAND is that it infers the best number of genetic groups,  $K$ , during a single run using the specific Reverse Jump MCMC (Markov chain Monte Carlo) algorithm, which enables the Markov chain to explore the whole range of putative cluster numbers at each step of the simulation run. The number of genetic groups that best explain the data (i.e. how confidently individuals are sorted) is given by the value of  $K$  in which the Markov chain spent most iterations (Guillot *et al.* 2005a; b).  $K$  was allowed to vary from 1 to 10 and initialized at eight. Three independent runs were performed to check for convergence of Markov chains. This first batch of runs with varying  $K$  enabled the best putative number of genetic groups to be assessed. Once the most likely  $K$  value was determined, three additional runs were performed, with the corresponding value of  $K$  fixed, to better infer the assignment level of each individual to the  $K$  genetic groups (Guillot *et al.* 2005b). For the three runs with varying  $K$ , we used the uncorrelated allele frequency model, and for those with fixed  $K$ , we used the correlated allele frequency model according to Guillot (2008) with default settings, except for the number of nuclei, which was increased to 500 (in the nonspatial mode), given the number of individuals. Run length was set to 550 000 iterations with a burning-in period of 50 000 iterations. The threshold of assignment coefficient ( $q$ ) for individual membership to a genetic group was set to 0.8.

*Population genetics analyses.* We computed summary statistics of genetic variability within (and its appar-

tionment among) groups identified using Bayesian clustering methods. Expected and observed heterozygosities were calculated with GENETIX (4.05) (Belkhir *et al.* 1996–2004). We calculated allelic richness ( $A_r$ ) using FSTAT version 2.9.3 (Goudet 1995) that implements a rarefaction procedure to account for differences in sample size. Differentiation coefficient between populations ( $F_{ST}$ ) and deviation from random mating within populations ( $F_{IS}$ ) was estimated according to Weir & Cockerham (1984) using GENETIX. Linkage disequilibrium between loci and deviations from HW equilibrium expectation (heterozygote excess and deficit) were calculated using GENEPOP on the Web (<http://genepop.curtin.edu.au>) (Rousset 2008). To adjust the  $P$  value for multiple tests for linkage disequilibrium, we used the false discovery rate (FDR) procedure (Benjamini & Yekutieli 2001), which controls the proportion of significant results ('false discovery rate') instead of controlling the chance of making even a single error. The resulting adjusted  $P$  values are called  $Q$  values. This procedure is implemented in the R package QVALUE (Storey & Tibshirani 2003).

*Analysis of isolation by distance.* Isolation by distance was assessed within the most widespread group (group 1) identified using Bayesian clustering methods (after removing locations with less than five assigned individuals) using the continuous method (Rousset 2000) of the 'isolde' procedure implemented in GENEPOP (Rousset 2008; option 'Analysis of isolation by distance between groups'). The significance of Mantel tests between pairwise genetic distances — analogue of  $F_{ST}/(1-F_{ST})$  ratios — and geographical distances were performed using the ISOLDE program (10 000 permutations). Geographical distances were log-transformed because *M. larici-populina* individuals evolve in a two-dimensional space (Rousset 1997). For geographical distances, we considered Euclidian distances between pairs of sampling sites.

*Test for loci under selection.* We tested for signs of positive selection on each locus, using the  $F_{ST}$ -outlier approach (Lewontin & Krakauer 1973), as implemented in FDIST2 (Beaumont & Nichols 1996). This method uses the expected distribution of  $F_{ST}$  vs.  $H_E$  to identify outlier loci that have excessively high (positive selection) or low (balancing selection)  $F_{ST}$  compared with neutral expectations. We ran 20 000 simulations to compute the 95% confidence intervals of the neutral value distribution of  $F_{ST}$  values under both the stepwise mutation model (SMM) and the infinite allele model (IAM).

*Tests for demographic equilibrium.* The demographic history of the groups of *M. larici-populina* identified using Bayesian clustering methods was assessed using two different methods designed to detect historical changes in population size, based on deviation from mutation–drift equilibrium. The first method relies on the fact that populations that have recently experienced a reduction in effective population size should exhibit larger values of gene diversity (i.e. heterozygosity level) than expected from the number of alleles at mutation–drift equilibrium (Cornuet & Luikart 1996). Conversely, expanding populations should exhibit deficits in heterozygosity levels. We used BOTTLENECK 1.2.02 (Piry *et al.* 1999) to assess the strength of the heterozygosity excess/deficit, i.e. expected heterozygosity ( $H_E$ ) in the sense of Nei (1978), relative to the expected equilibrium heterozygosity ( $H_{eq}$ ). The distribution of  $H_{eq}$  was obtained through simulation of the coalescent process given the observed allelic richness. The tests were performed under the SMM, as well as under a two-phase model (TPM), allowing for 30% of multistep changes (IAM). We performed all analyses with 2000 iterations of the coalescent process. Significance of the deviation from mutation–drift equilibrium was assessed using Wilcoxon signed rank tests, which are most powerful and robust when used with few polymorphic loci (Cornuet & Luikart 1996).

The second method is dedicated to loci, such as microsatellite loci, evolving under the SMM. Under this mutational model, a demographic scenario is likely to influence the distribution of allele sizes (Kimmel *et al.* 1998). In particular, for a population in expansion, the observed distribution of allele sizes should be less variable when compared with a population at demographic equilibrium. We used the imbalance index  $\beta_2$ , which estimates the strength of the deviation between variance in allele size and expected heterozygosity, as it is the most powerful index to reveal population expansion (equation 8 in King *et al.* 2000).  $\ln \beta$  is expected to be negative for a strictly growing population and positive for a population that has recently expanded following a bottleneck (Kimmel *et al.* 1998). Using the R package LOCFIT, 95% confidence intervals were obtained from the distributions of  $\ln \beta_2$  estimates across loci (Loader 1996).

## Results

### Virulence profiles

A total of 1050 individuals of *M. larici-populina* from 23 locations across France (Table 1) were phenotyped for their ability to overcome (or not) the R7 qualitative resistance gene carried by *P. × interamericana* ‘Beaupré’.

The frequency of individuals carrying the corresponding virulence (hereafter referred to as Vir7) ranged from 0 to 1 across sampling sites (Fig. 1), with a mean proportion of 0.53.

The spatial distribution of Vir7 rust individuals was nonrandom. Individuals from the northern half of France exhibited the highest proportion of Vir7 (up to 100% in Charrey-sur-Saône and Roubaix) (Fig. 1). The sole exception to this general trend was Saint-Ay, for which only 40% of individuals were Vir7. However, samples collected at a neighbouring site (Orléans nursery) were almost all Vir7 (91%). Unfortunately, these individuals were not genotyped. A significant difference in the proportion of Vir7 between wild stands and poplar nurseries was found only in Saint-Ay vs. Orléans ( $P < 0.05$ ); the other nurseries (Guémené-Penfao, Amance, and Charrey-sur-Saône) displayed a Vir7 proportion similar to that found in wild poplar stands located nearby (Rennes, Nancy, and Nievroz, respectively) (Fisher’s exact test,  $P = 0.83$ ).

### Summary of genetic and genotypic variability

Of the 514 individuals genotyped, 40 (8%) were excluded from further analyses because they failed to amplify or had too much missing data. These individuals were scattered across sampling sites (Table 1). The remaining 474 individuals exhibited only 2.9% missing data on average. Of the 474 individuals analysed, 388 distinct multilocus genotypes were found. All the repeated multilocus genotypes were found to be significantly over-represented according to the analysis

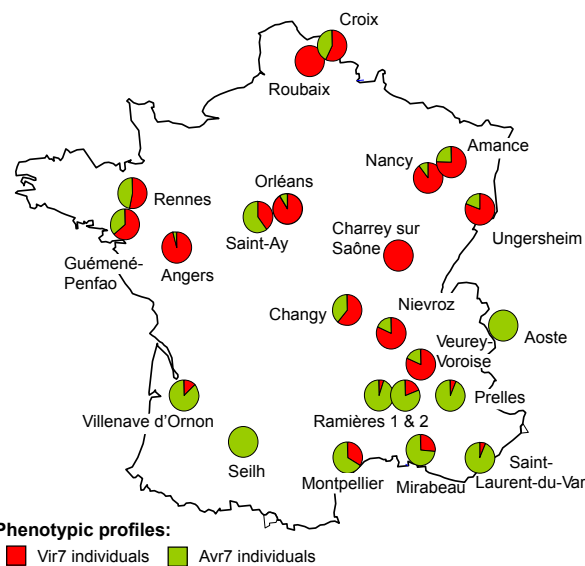


Fig. 1 Geographical distribution of phenotypic profiles showing proportion of Avr7 and Vir7 individuals at each site.



performed using MLG<sub>SIM</sub>. Repetitions of a given multilocus genotype were thus considered to result from clonal reproduction and were further denoted clonal copies. Some clonal copies were sampled from different locations but most were restricted to a single site; only three repeated multilocus genotypes occurred in more than ten clonal copies. Subsequent analyses were performed considering a single copy of each repeated multilocus genotype.

All 25 microsatellite markers were polymorphic, with a number of alleles ranging from 5 to 19 (Table 2). The mean allele number across loci was 9.6 when pooling all samples and 7.0 when considering only the reference population of Prelles, as in the study by Xhaard *et al.* (2009). Gene diversity (i.e. expected heterozygosity,  $H_E$ ) equalled 0.57 when considering all sampling locations and 0.56 for the Prelles sampling site. A highly significant HW deviation was found for almost all loci when considering all samples, which together with the many pairs of loci in significant linkage disequilibrium observed (197/300) suggests some genetic structure (i.e. Wahlund effect). This trend was not observed when considering only the Prelles location, where nearly all loci did not depart significantly from HW equilibrium expectations.

### Genetic structure

GENELAND runs failed to consistently assess the number of genetic groups when taking into account the spatial information: two runs sorted the samples into two genetic groups (93% and 95% of iterations along the

first and second Markov chain, respectively) and the third in three genetic groups (95% of iterations along the third Markov chain). Moreover, when fixing the value of  $K$  (three runs with  $K = 2$  and three runs with  $K = 3$ ), assignment results were incongruent across runs. Conversely, under the nonspatial mode, the three runs gave consistent results, with the three Markov chains converging rapidly on  $K = 3$  (Markov chains stuck at  $K = 3$  for the three runs). Setting a fixed value of  $K = 3$  gave very consistent assignment results. In the following, we thus considered only those results obtained with the nonspatial mode. The membership coefficients  $q$  were averaged over the three runs with a fixed value of  $K = 3$ . Most individuals had high membership in a single genetic group (Fig. 2). Applying an assignment threshold of  $q > 0.8$ , more than 90% of multilocus genotypes (354 out of 390) could be assigned to a genetic group. Re-including multiple copies of the repeated multilocus genotypes, the three genetic groups comprised 192, 170 and 71 individuals, respectively. Some individuals from the same location belonged to different genetic groups (Fig. 2). This gathering of individuals with very different genetic profiles is likely to account for convergence failure of GENELAND when used with the spatial model, as under the spatial mode the same assignment coefficient ( $q$ ) is set for all individuals sampled at a given location. The 43 nonassigned individuals were distributed evenly across the sampling sites, except in Ramières 1, where only three individuals could be confidently assigned to a genetic group (Fig. 2). Most individuals from this location were derived from two clones displaying equal membership

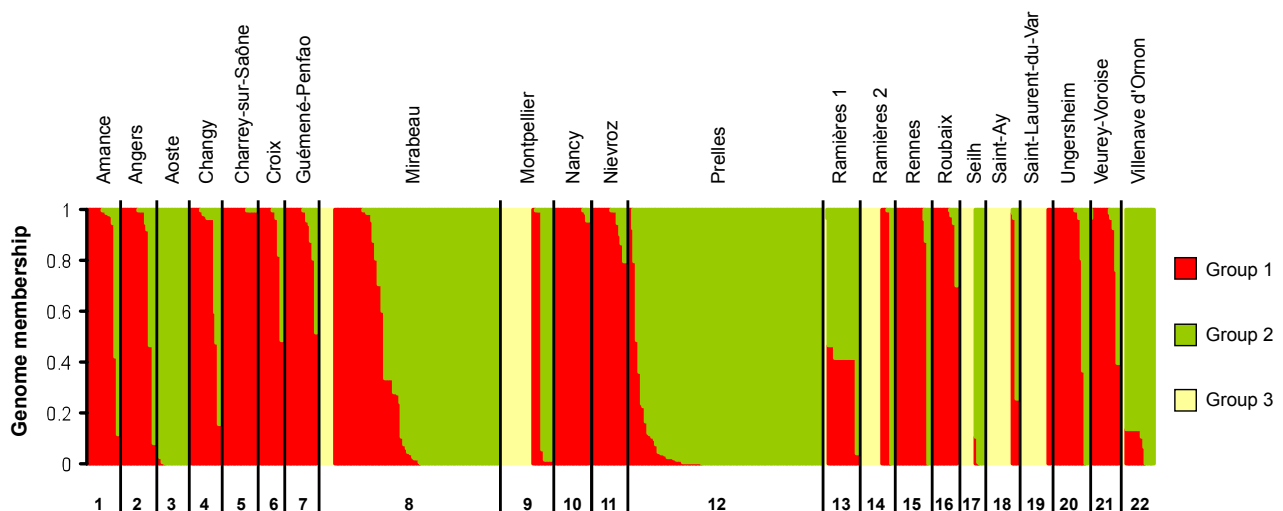
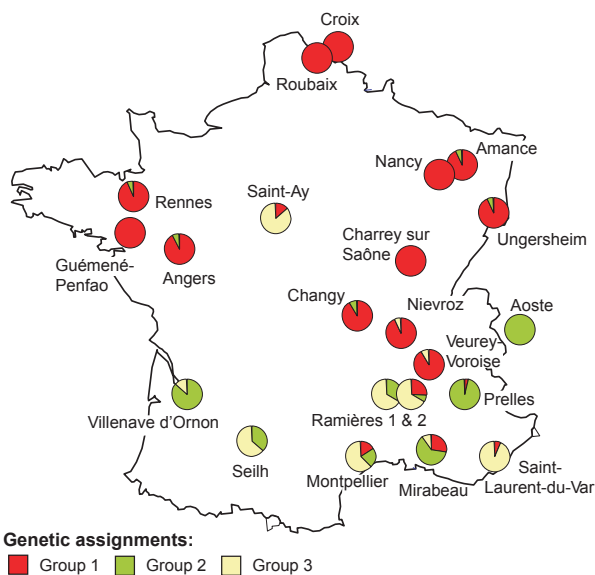


Fig. 2 Assignment results of *Melampsora larici-populina* individuals. Each vertical line represents an individual whose genome is partitioned into  $K$  coloured segments (here  $K = 3$  genetic groups). Segment lengths are proportional to the individual's mean membership coefficients in each genetic group. Vertical bars separate individuals sampled from different locations. Geographical locations are indicated above and population number (Table 1) underneath.

to genetic groups 1 and 2 (Fig. 2). The 43 nonassigned individuals were excluded from subsequent analyses.

Differentiation between the three groups was highly significant ( $P < 0.001$ ), whichever comparison pair was considered, which is in accordance with the clear-cut clustering results. Genetic differentiation between groups 1 and 2 was lower ( $F_{ST} = 0.04$ ) than comparisons including group 3 ( $F_{ST} = 0.24$  and  $0.20$  with groups 1 and 2, respectively).

The spatial distribution of the three genetic groups is illustrated in Fig. 3. Individuals from group 1 occurred in nearly all locations (with the exception of Aoste, Ramières 1, Seilh and Villenave d'Ornon). However, group 1 occurred mostly in the northern part of France (with the exception of Saint-Ay, where this group accounted for only 13% of individuals). The second genetic group (group 2) was more abundant in the southern part of France, but again was unevenly distributed and reached up to 100% of assigned individuals in Aoste, and nearly 97% in Prelles, both Alpine locations. Group 2 was present in all southern locations except Saint-Laurent-du-Var, possibly because of the limited sample size. In the northern part of France, individuals from group 2 were seldom sampled together with group 1. The third genetic group (group 3) was restricted to the southern part of France, with the notable exception of Saint-Ay, where group 3 accounted for 80% of the individuals (Fig. 2). All individuals belonging to this third genetic group were very well assigned ( $q = 1$ , Fig. 2) although they always coexisted with individuals belonging to the other two genetic groups (Fig. 3).



**Fig. 3** Geographical distribution of membership of the three genetic groups, showing proportion of individuals assigned to each genetic group at each site.

#### *Genetic and phenotypic characteristics of the three genetic groups identified*

Group 1 exhibited the highest genotypic richness ( $G/N = 0.96$  with  $N$  the number of individuals genotyped, and  $G$  the number of distinct multilocus genotypes) (Table 3). Only six repeated multilocus genotypes were found in this group, each occurring in two to three clonal copies, all sampled in the same location. Eighty-one percent of the individuals from this group were Vir7, representing 86% of the total amount of Vir7 individuals. A highly significant correlation was found between individual membership coefficients to group 1 and their virulence profile (logistic regression,  $P < 0.001$ ). Pooling all individuals regardless of their location of origin revealed a large heterozygote deficit (HW disequilibrium), and some pairs of loci exhibited significant linkage disequilibrium, suggesting some substructuring. Significant genetic differentiation was found among locations ( $P < 0.001$ , Table 3), but no isolation by distance was detected ( $P = 0.159$ ).

Group 2 contained only 10% Vir7 individuals (Table 3). Differences with the genetic characteristics of group 1 included a lower genotypic richness but a larger genetic diversity ( $H_E$ ) and allelic richness ( $A_r$ ). Substructuring was also more pronounced, as shown by the higher number of loci in linkage disequilibrium and the higher  $F_{IS}$  and global  $F_{ST}$  values. The  $F_{IS}$  decreased markedly when taking into account sampling location (Table 3), although the heterozygote deficit was still significant. For both genetic groups, no pair of loci remained in linkage disequilibrium after removing clonal copies and taking into account sampling location. Of the seven repeated multilocus genotypes, two were more abundant (seven and ten clonal copies) and the five others were found at between two and four clonal copies. As in group 1, all the clonal copies of each repeated genotype were found in a single location.

Group 3 displayed very specific genetic characteristics, with lower allelic and genotypic richness, much more pairs of loci in linkage disequilibrium and larger heterozygote excess, which led to a very negative and significant value of  $F_{IS}$  ( $P < 0.001$ ) (regardless of whether clonal copies were taken into account). The two most repeated genotypes were represented by 17 and 12 clonal copies. The other nine repeated multilocus genotypes were represented by two to four clonal copies. Only three repeated multilocus genotypes were found in different locations (including the most repeated genotype with a maximum resampling distance of 600 km). Group 3 was the only group to have clonal copies sampled at different locations. As with group 2, only a few individuals (9% of the multilocus genotypes) showed a Vir7 profile.

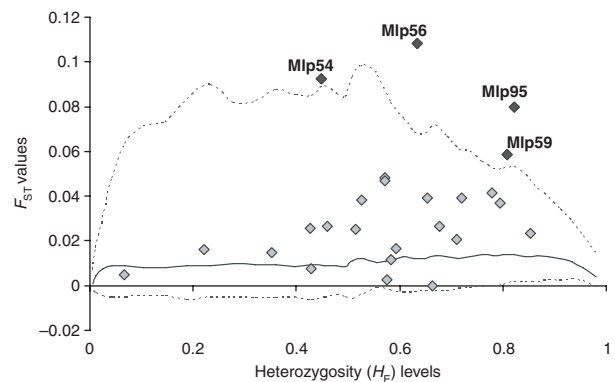
**Table 3** Characteristics of the three genetic groups

	N	G/N	Vir7	$H_E$	$H_O$	Heterozygote		$A_r$	$F_{IS}$	LD	F-statistics	
						Excess	Deficit				$F_{ST}$	$F_{IS}$
Including clonal copies												
Group 1 'virulent-sexual'	192	0.96	0.81	0.536 (0.209)	0.501 (0.199)	ns	***	5.44 (3.17)	0.065 (0.100)	46/300	0.013 *** (0.004)	0.051 *** (0.016)
Group 2 'avirulent-sexual'	170	0.87	0.10	0.578 (0.203)	0.521 (0.181)	ns	***	5.69 (2.88)	0.100 (0.206)	244/300	0.093 *** (0.009)	0.033 ** (0.032)
Group 3 'avirulent-asexual'	71	0.45	0.13	0.378 (0.257)	0.561 (0.416)	***	ns	2.55 (1.16)	-0.486 (0.413)	150/230	0.116 *** (0.027)	-0.666 *** (0.038)
Without clonal copies												
Group 1 'virulent-sexual'	184	-	0.80	0.536 (0.198)	0.500 (0.194)	ns	***	5.48 (2.44)	0.068 (0.103)	5/300	0.009 *** (0.004)	0.057 *** (0.015)
Group 2 'avirulent-sexual'	138	-	0.09	0.578 (0.181)	0.512 (0.176)	ns	***	5.78 (2.36)	0.113 (0.194)	0/300	0.033 *** (0.004)	0.094 *** (0.030)
Group 3 'avirulent-asexual'	32	-	0.09	0.397 (0.248)	0.541 (0.388)	***	ns	2.75 (1.16)	-0.369 (0.381)	118/230	0.071 ** (0.034)	-0.477 *** (0.052)

N, number of individuals; G/N, genotypic richness; Vir7, proportion of virulent individuals;  $H_E$ , unbiased estimate of gene diversity (Nei 1978);  $H_O$ , observed heterozygosity; probability test for heterozygote excess or deficit (\*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; ns, nonsignificant);  $A_r$ , allelic richness estimated for a minimal sample size of 71 (or 32 without clonal copies);  $F_{IS}$  values; LD, number of pairs of significant linkage disequilibrium after correction for multiple tests (for group 3 only 230 comparisons were calculated, because some loci were monomorphic in this group). The same calculations were conducted both with and without clonal copies. F-statistics ( $F_{ST}$  and  $F_{IS}$ ) were calculated according to Weir & Cockerham (1984) taking into account the sampling location (within genetic groups). When available, standard errors across loci are given underneath in brackets.

#### Evidence for positive selection in group 1

Reduced genetic variability and a larger variance in heterozygosity across loci both suggested a possible history of strong natural selection in group 1 (Table 3). We further tested for selection by comparing the observed  $F_{ST}$  between groups 1 and 2 at each locus with the distribution expected under neutrality, conditional on levels of heterozygosity (Fig. 4). Tests were conducted only between genetic groups 1 and 2, as extensive evidence of long-lasting asexual reproduction in group 3 violated the model assumptions. Simulations in  $F_{DIST2}$  were run with the following parameters: two demes, two populations, mean value of  $F_{ST}$  calculated between groups 1 and 2 ( $F_{ST} = 0.04$ ) and the number of individuals without clonal copies, which comprised genetic groups 1 and 2 (322 individuals). Four microsatellite loci deviated from neutral expectations and displayed signs of positive selection (MLP54 and MLP59,  $P < 0.05$  and MLP56 and MLP95  $P < 0.01$ ) (Fig. 4). No locus was detected under balancing selection. The same result was obtained with both mutation models, SMM and IAM. The four loci had lower levels of heterozygosity in group 1 than in group 2, suggesting that group 1 might have been exposed to stronger selective pressures than group 2.



**Fig. 4** Test for selection between genetic groups 1 and 2. Observed  $F_{ST}$  values are plotted against heterozygosity levels for each locus. Solid and dashed lines denote median and 95% confidence interval of the simulated distributions under the stepwise mutation model (SMM), respectively. Four loci were candidates for being under positive selection (black diamonds). The remaining loci were found with neutral expectations (grey diamonds).

#### Demographic history of genetic groups 1 and 2

We used two complementary approaches to reveal the demographic history of *M. larici-populina* populations: a test for expected heterozygosity excess/deficit (Piry *et al.* 1999) and the value of the imbalance index

(Kimmel *et al.* 1998). As a major assumption of these tests is that individuals reproduce sexually, they were performed on groups 1 and 2 only, not taking into account clonal copies. The tests were also performed on the reference population at Prelles located in the original poplar-larch sympatry area and thus expected to have experienced reduced demographic fluctuations. In addition, as loci under selection can bias estimates of demographic parameters, we retained only the 21 loci that did not significantly deviate from neutral expectations in analyses with  $F_{DIST2}$ .

In analyses using BOTTLENECK, the two genetic groups exhibited a significant heterozygosity deficit (Table 4) whatever the mutation model considered (SMM or TPM), suggesting population expansion. The heterozygosity deficit was stronger under the SMM. Considering the reference population of Prelles, mutation–drift disequilibria were observed only under the SMM.

The imbalance index was largely positive in all subsamples considered (Table 4), suggesting that groups 1 and 2 of *M. larici-populina* have recently expanded following a bottleneck (Kimmel *et al.* 1998). On average, the value of the imbalance index was highest (and its density distribution narrowed) for group 1.

## Discussion

### *Coexistence of three differentiated genetic groups on the same host plant species*

In the present study, we investigated the population genetic structure of a plant pathogenic fungus (*M. larici-populina*) on its wild host (*P. nigra*) in France, where

both wild and cultivated poplars are sympatric in many locations. Bayesian clustering analyses revealed that *M. larici-populina* individuals could be clustered into three genetically distinct groups. These groups could coexist locally but predominated in different well-defined geographical areas. The three groups displayed differences in terms of microsatellite variability and prevalence of virulent isolates able to infect cultivated poplar carrying resistance 7. Group 1 consisted mainly of virulent individuals, while the two other genetic groups were mostly represented by avirulent individuals. Group 3 exhibited lower genotypic richness, extensive linkage disequilibrium across loci and a large heterozygote excess suggesting a predominantly asexual reproductive mode (Balloux *et al.* 2003; Halkett *et al.* 2005a). Conversely, genetic characteristics of groups 1 and 2 were consistent with a predominantly sexual reproductive mode. In the following, group 1 is referred to as the ‘virulent-sexual’ group, group 2 as the ‘avirulent-sexual’ group and group 3 as the ‘avirulent-asexual’ group.

Numerous population genetic studies in phytopathogenic fungi highlighted strict host plant specialization with large genetic differences among (not within) host plants (Enjalbert *et al.* 2005; Goyeau *et al.* 2007; Fournier & Giraud 2008; Zaffarano *et al.* 2008; Giraud *et al.* 2010; Gladieux *et al.* 2010b). Unlike these studies, a significant genetic structure was found among *M. larici-populina* samples collected on the same host plant, *P. nigra*, with individuals assigned to distinct genetic groups that can coexist locally. In most fungal plant pathogens, limited dispersal between selection on the host and mating favours host-specific differentiation because mutations for adaptation on a new host affect both fitness on the host and mating patterns (Giraud *et al.* 2006, 2010). However, as discussed by Barrès *et al.* (2008) and Gilabert *et al.* (2009), such processes of host-specific differentiation through host-specialization are not expected in host-alternating species like *M. larici-populina*. Indeed, unless isolating mechanism(s) favouring assortative mating or linkage among loci involved in plant adaptation pre-exist (Bolnick & Fitzpatrick 2007), no possible barriers to gene flow can emerge thereby impeding host-specific differentiation. Gene mixing may be all the more strengthened in *M. larici-populina* as sexual reproduction occurs on the alternate host with a very scattered distribution ensuring extensive genetic reshuffling among individuals with various origins (Gérard *et al.* 2006). The genetic structure revealed by our study would result from the spatial distribution and local aggregation of *M. larici-populina* lineages with three distinct evolutionary histories that came into secondary contact on the same host species.

**Table 4** Tests for mutation–drift equilibrium

	Bottleneck tests		Imbalance index
	(D/E)		
	SMM	TPM	Ln $\beta_2$
Group 1 'virulent-sexual'	20/1****	17/4***	2.83 (1.41–4.26)
Group 2 'avirulent-sexual'	20/1****	15/6**	2.52 (0.90–4.14)
Prelles (reference population)	16/5***	11/10	2.52 (0.79–4.25)

Bottleneck tests: number of loci exhibiting heterozygosity deficit (D) or excess (E) under the two mutation models (SMM, stepwise mutation model, or TPM, two-phase model). The probability of deviation from mutation–drift equilibrium was determined using a two-sided Wilcoxon test (\*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ ). Imbalance index: mean estimates of Ln  $\beta_2$  over loci and 95% confidence intervals (in brackets).

*Strong association among microsatellite and phenotypic markers despite extensive gene flow*

We found evidence of a highly significant relationship between genetic groups and phenotype profiles, as the 'virulent-sexual' group 1 was found to consist mainly of individuals with a Vir7 profile, contrary to the two other groups, which were almost free from Vir7 individuals.

The 'avirulent-sexual' group 2 predominated in the Alps (Prelles and Aoste locations), which are included in the native distribution area of *M. larici-populina* (Barres *et al.* 2008), where larches and wild poplars have evolved in sympatry. In addition, the reference population at Prelles, located in the area of poplar-larch sympatry, displayed limited evidence of demographic disequilibrium and did not depart from expectations under panmixia, which is in accordance with a long history of evolution on the wild host *P. nigra*. Despite a certain degree of population substructure, the clustering of these individuals into the same genetic group 2 is consistent with a common ancestral origin and a long-standing coevolution with the wild host.

Several pieces of evidence indicate that individuals from the 'virulent-sexual' group 1 are derived from the breakdown of resistance 7 carried by cultivated poplars such as 'Beaupré'. First, we found that group 1 was genetically less diverse than group 2. This is in accordance with the loss of heterozygosity going hand-in-hand with pathogen evolution on cultivated crops (as in Gladieux *et al.* 2008; Burdon & Thrall 2008), as a result of founder effects because of the limited number of individuals able to infect the overcome cultivar(s) at resistance breakdown (Guérin *et al.* 2007). Second, tests for mutation-drift equilibrium revealed that the virulent population experienced demographic expansion after the initial genetic bottleneck associated with the resistance breakdown. Finally, in addition to demographic changes, the test for neutral evolution revealed that some loci were clearly under positive selection and that these loci exhibited a drop in allelic diversity in group 1. Together with the tremendous population expansion (both revealed indirectly here, and directly observed in Pinon *et al.* 1998), these results are consistent with a very drastic selection pressure exerted by the cultivated host 'Beaupré' on *M. larici-populina*. Both selection by plant resistance and demographic fluctuations (founder event followed by rapid population expansion) caused the observed genetic differentiation.

Some intriguing results arise nonetheless in the analysis of the demographic history of genetic groups 1 and 2: the 'avirulent-sexual' group 2 taken as a whole was not at mutation-drift equilibrium and also displayed extensive (although less pronounced) evidence for pop-

ulation expansion. This might be because of violations of the assumption of the model implemented in BOTTLENECK such as population substructure or the alternation of sexual/asexual reproduction in the life cycle. However, the signal of population expansion might also be because of an increase in the distribution area of the 'avirulent-sexual' group 2 resulting from the extensive plantation of larches as ornamental or wood production trees since the eighteenth century. The fact that this demographic event is older than the one resulting from R7 breakdown might be in accordance with the larger interlocus variance in demographic parameter estimates. Using historical samples of *M. larici-populina* might allow deciphering further the demographic history of these two genetic groups.

The 'avirulent-asexual' group 3 was also found to consist mostly of Avr7 individuals. Such tight link between phenotypic and genetic differentiation has been thoroughly assessed in asexual species of wheat rusts such as *Puccinia triticina* (Goyeau *et al.* 2007) and *Puccinia striiformis* f. sp. *tritici* (Enjalbert *et al.* 2005). The association results from the hitchhiking of neutral marker loci on selected virulence loci thanks to the lack of recombination. Assuming that the clonal lineages of *M. larici-populina* date back to before the breakdown of R7 resistance (1994), the absence of recombination would have prevented these individuals from being introgressed by virulence 7 alleles.

*Virulence introgression dynamics and spatial expansion of Vir7 individuals on the wild host*

The 'virulent-sexual' group 1 was found to be the largest, and to account for nearly 50% of sampled individuals, most of which were gathered in the northern half of France. The spread of Vir7 individuals across poplar cultivation has been thoroughly assessed by population surveys contemporaneous with breakdown of the R7 resistance (Pinon *et al.* 1998). Historically, poplar cultivars carrying the R7 resistance gene, such as the cultivar 'Beaupré', were planted widely on large acreages in northern France, Belgium and the Netherlands between 1980 and 1995. This agrosystem favoured the breakdown of R7 in this area in 1994 (Pinon *et al.* 1998). Virulence 7 emerged in Belgium and in northern France, and rapidly spread southwards. Within 5 years, Vir7 individuals occurred in all French poplar cultivation areas including those, in southern France, where the acreage of R7 cultivars was much lower or null. According to a previous population genetics study, between 36% and 75% of individuals sampled on susceptible poplars in the northern half of France in 2001 were Vir7 (Gérard *et al.* 2006). Here, we confirmed that Vir7 individuals have spread widely on

susceptible plants and that the Vir7 proportion has now reached 54% to 100% of individuals in the northern half of France. This rapid invasion of Vir7 individuals is consistent with the remarkable dispersal abilities of rust spores, as already shown in *M. larici-populina* (Barrès *et al.* 2008). Accordingly, we found no isolation by distance among individuals from the 'virulent-sexual' group 1 sampled across France. Nonetheless, 'virulent-sexual' group 1 prevails only in the northern half of France, and some southern locations (Aoste, Ramières 1, Seilh and Villenave d'Ornon) still remain free of individuals belonging to this group. Interestingly, the spatial distribution of 'virulent-sexual' group 1 individuals matched very well the area where R7 poplar cultivars (like 'Beaupré') were massively planted. This spatial correlation is in accordance with the study by Montarry *et al.* (2008) on *P. infestans*, suggesting that—because of the dispersal ability of plant pathogens—patterns resulting from the adaptation to the most abundant host genotype can only be revealed at large spatial scales.

According to the theoretical work by Lenormand (2002), the combination of strong selection pressure and high gene flow would lead to the (unnecessary) spread of genes conferring adaptation to the environment most likely to be encountered. Our results are consistent with this expectation and suggest that the plantation of R7 cultivars and the subsequent breakdown of the resistance has led to 'gene swamping' (Lenormand 2002) in *M. larici-populina* populations. Given the large genetic signature of selection and the demographic history of 'virulent' group 1, it is quite likely that the spread of virulence 7 into the wild host consisted not only of the introgression of the single virulence gene but rather resulted from an unbalanced gene mixing that led to the replacement of populations from the wild original with the observed virulent genetic group. In support of this hypothesis, we found very few hybrid individuals with mixed membership in genetic groups 1 and 2 (the only consistent exception was found in Ramières 1) and individuals from these two groups only rarely live in sympatry (and never in equal abundance). This large degree of gene mixing is fostered by the peculiarity of the poplar rust life cycle, as sexual reproduction on larch ensures extensive genetic reshuffling between the individuals from different origins that have to gather at these specific habitats.

#### *Polymorphism in the reproductive mode of M. larici-populina*

Patterns of genetic variation were consistent with genetic group 3 being composed of individuals that

reproduce mainly asexually (Balloux *et al.* 2003; Halkett *et al.* 2005a). Over-represented multilocus genotypes and heterozygote excess has been used previously to find evidence and quantify the asexual reproductive mode in a wide range of organisms (e.g. fungal pathogens: Burdon & Roelfs 1985; Goyeau *et al.* 2007; aphids: Halkett *et al.* 2005b; Gilabert *et al.* 2009; trees: Stoeckel *et al.* 2006; and nematodes: Villate *et al.* 2010). The numbers of repeated multilocus genotypes and the large heterozygote excess observed in group 3 of *M. larici-populina* were consistent with the reproductive mode being mostly clonal and including only occasionally sexual events (Balloux *et al.* 2003). Persistence of clonal lineages over several years had never been suspected so far in *M. larici-populina*, and repeated multilocus genotypes were always considered as clonal copies resulting from asexual multiplication during a single annual epidemic phase (Gérard *et al.* 2006; Barrès *et al.* 2008). However, despite being the genetic group with the smallest sample size, individuals from group 3 appeared widespread. These long-lasting clonal lineages may have stayed unnoticed so far because of the smaller size and lower resolving power of the marker set formerly used (Barrès *et al.* 2006), which would not have allowed revealing genetic partition of *M. larici-populina* individuals according to their differences in reproduction mode.

Rust fungi (Pucciniales) are well known for displaying potential polymorphism in their reproductive mode (Burdon & Roelfs 1985; Samils *et al.* 2001; Goyeau *et al.* 2007; Mboup *et al.* 2009; Jin *et al.* 2010). In particular, asexuality is the only reproductive mode possible in situations where the alternate host is missing (the plant on which sexual reproduction occurs for a heteroecious rust). This is the case either after migration of the pathogen to a new geographical area where the alternate host is absent (Hovmøller *et al.* 2002; Goyeau *et al.* 2007) or after its eradication (Burdon & Roelfs 1985). Alternatively, the distribution area of host alternating species with obligate sexual phase is constrained by the occurrence of the alternate host. A well-documented example is *M. medusae* f. sp. *deltoidae* with a native distribution in poplar-larch sympatry areas in Canada and annual (recurrent) colonizations of the USA resulting from a dead-end clonal propagation on poplars (Bourassa *et al.* 2007). In species that exhibit both reproductive modes, sexual and asexual lineages generally grow in allopatry. For example, Burdon & Roelfs (1985) documented how *P. graminis* asexual populations evolved in areas where the sexual host was eradicated. In addition, some particular climatic conditions can be more favourable to asexual reproduction (Groth *et al.* 1995; Barrett *et al.* 2008a). The willow rust fungus *Melamps-*

*ora larici-epitea* was shown to exhibit a dual reproductive mode depending on climate, with asexual reproduction in Northern Ireland, where mild winters could favour overwintering of asexual spores (urediniospores), and sexual reproduction on larch in Sweden, where cold winters could favour sexual forms (teliospores) (Samils *et al.* 2001). In the present study, most clonal lineages were found in southern France, where milder winters may favour the asexual survival of *M. larici-populina*, and/or may be less favourable to overwintering of sexual spores. However, we also observed cases of local coexistence of sexual and asexual lineages, which is more puzzling, as competition should lead to the exclusion of one of the two kinds of lineages. Further experiments are required to document differences in the life-history traits of these lineages which can define subtle spatial or temporal niches, as it was found for another host-alternating parasite species (Halkett *et al.* 2006; Gilabert *et al.* 2009).

## Conclusion

This study on the genetic structure of *M. larici-populina* individuals revealed that three genetic groups coexisted on the same host plant species, the wild host *P. nigra*. Our analyses revealed that this structure may result from (i) reproductive mode polymorphism and (ii) the strong selection pressure imposed on the pathogen by the release of resistant domesticated plants. The three genetic groups prevailed in distinct regions. Features of the agricultural landscape may explain the pattern of spatial structure observed in this plant pathogen. Further examination of the possible link between relative cultivated host density and the observed spatial genetic structure may provide more insight into the selective pressures at work. In addition to a finer grained spatial monitoring, it would be worth developing a retrospective approach to retrace the temporal evolution of the genetic signatures of this drastic selective event on the genome of *M. larici-populina*.

## Acknowledgements

We thank Audrey Andanson, Alain Berthelot, Gilles Bossuet, Virginie Diévert, Arnaud Dowkiw, Nicolas Feau, Olivier Forestier, Etienne Klein and François Lefèvre for their help in collecting samples, Adeline Sipieter and Béranger Bertin for their technical help, Cyril Dutech, Jérôme Enjalbert, Frédéric Fabre and Etienne Klein for insightful discussions and Helen Rothnie for her attentive reading and careful English language checking. Richard Hamelin and two anonymous referees provided valuable comments on earlier versions of this manuscript. Constance Xhaard was

supported by a fellowship from the French Ministry of Education and Research (MESR). This research program was supported by grants from INRA (projet innovant EFPA), the Agence Nationale de la Recherche, contract number ANR 07-BDIV-003 (Emerfundis project) and the French Ministry for Agriculture and Fisheries (Programme 142, sous-action 27, CTPS).

## References

- Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, **19**, 535–544.
- Balloux F, Lehmann L, de Meeus T (2003) The population genetics of clonal and partially clonal diploids. *Genetics*, **164**, 1635–1644.
- Barrès B, Dutech C, Andrieux A, Caron H, Pinon J, Frey P (2006) Isolation and characterization of 15 microsatellite loci in the poplar rust fungus, *Melampsora larici-populina*, and cross-amplification in related species. *Molecular Ecology Notes*, **6**, 60–64.
- Barrès B, Halkett F, Dutech C, Andrieux A, Pinon J, Frey P (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: Evidence for isolation by distance in Europe and recent founder effects overseas. *Infection, Genetics and Evolution*, **8**, 577–587.
- Barrett LG, Thrall PH, Burdon JJ, Nicotra AB, Linde CC (2008a) Population structure and diversity in sexual and asexual populations of the pathogenic fungus *Melampsora lini*. *Molecular Ecology*, **17**, 3401–3415.
- Barrett LG, Thrall PH, Burdon JJ, Linde CC (2008b) Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology & Evolution*, **23**, 678–685.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B-Biological Sciences*, **263**, 1619–1626.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, Logiciel Sous Windows TM Pour La Génétique Des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, **29**, 1165–1188.
- Bolnick DI, Fitzpatrick BM (2007) Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution and Systematics*, **38**, 459–487.
- Bourassa M, Bernier L, Hamelin RC (2007) Genetic diversity in poplar leaf rust (*Melampsora medusae* f. sp. *deltoidae*) in the zones of host sympatry and allopatry. *Phytopathology*, **97**, 603–610.
- Bucheli E, Gautschi B, Shykoff JA (2001) Differences in population structure of the anther smut fungus *Microbotryum violaceum* on two closely related host species, *Silene latifolia* and *S. dioica*. *Molecular Ecology*, **10**, 285–294.
- Burdon JJ (1993) The structure of pathogen populations in natural plant-communities. *Annual Review of Phytopathology*, **31**, 305–323.

- Burdon JJ, Chilvers GA (1982) Host density as a factor in plant disease ecology. *Annual Review of Phytopathology*, **20**, 143–166.
- Burdon JJ, Roelfs AP (1985) The effect of sexual and asexual reproduction on the isozyme structure of populations of *Puccinia graminis*. *Phytopathology*, **75**, 1068–1073.
- Burdon JJ, Thrall PH (2008) Pathogen evolution across the agro-ecological interface: implications for disease management. *Evolutionary Applications*, **1**, 57–65.
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.
- Couch BC, Fudal I, Lebrun MH *et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, **170**, 613–630.
- Daszak P, Cunningham AA, Hyatt AD (2000) Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science*, **287**, 443–449.
- Desprez-Loustau ML, Robin C, Buee M *et al.* (2007) The fungal dimension of biological invasions. *Trends in Ecology & Evolution*, **22**, 472–480.
- Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature*, **418**, 700–707.
- Enjalbert J, Duan X, Leconte M, Hovmoller MS, De Vallavieille-Pope C (2005) Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. *Molecular Ecology*, **14**, 2065–2073.
- Fournier E, Giraud T (2008) Sympatric genetic differentiation of a generalist pathogenic fungus, *Botrytis cinerea*, on two different host plants, grapevine and bramble. *Journal of Evolutionary Biology*, **21**, 122–132.
- Frey P, Gérard PR, Feau N, Husson C, Pinon J (2005) Variability and population biology of *Melampsora* rusts on poplars. In: *Rust Diseases of Willow and Poplars* (eds Pei MH, McCracken AR). pp. 63–72, CAB International, Wallingford, UK.
- Gérard PR, Husson C, Pinon J, Frey P (2006) Comparison of genetic and virulence diversity of *Melampsora larici-populina* populations on wild and cultivated poplar and influence of the alternate host. *Phytopathology*, **96**, 1027–1036.
- Gilbert A, Simon JC, Mieuxet L *et al.* (2009) Climate and agricultural context shape reproductive mode variation in an aphid crop pest. *Molecular Ecology*, **18**, 3050–3061.
- Gilbert GS (2002) Evolutionary ecology of plant diseases in natural ecosystems. *Annual Review of Phytopathology*, **40**, 13–43.
- Giraud T, Villareal L, 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.
- Giraud T, Gladieux P, Gavrillets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution*, **25**, 387–395.
- Gladieux P, Zhang XG, Afoufa-Bastien D, Sanhueza RMV, Sbaghi M, Le Cam B (2008) On the origin and spread of the scab disease of apple: out of Central Asia. *PLoS ONE*, **3**, e1455.
- Gladieux P, Zhang XG, Roldan-Ruiz I *et al.* (2010a) Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Molecular Ecology*, **19**, 658–674.
- Gladieux P, Caffier V, Devaux M, Le Cam B (2010b) Host-specific differentiation among populations of *Venturia inaequalis* causing scab on apple, pyracantha and loquat. *Fungal Genetics and Biology*, **47**, 511–521.
- Gómez-Alpizar L, Carbone I, Ristaino JB (2007) An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. *Proceedings of the National Academy of Sciences*, **104**, 3306–3311.
- Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.
- Goyeau H, Halkett F, Zapater MF, Carlier J, Lannou C (2007) Clonality and host selection in the wheat pathogenic fungus *Puccinia triticina*. *Fungal Genetics and Biology*, **44**, 474–483.
- Groth JV, McCain JW, Roelfs AP (1995) Virulence and isozyme diversity of sexual versus asexual collections of *Uromyces-Appendiculatus* (Bean rust fungus). *Heredity*, **75**, 234–242.
- Guérin F, Gladieux P, Le Cam B (2007) Origin and colonization history of newly virulent strains of the phytopathogenic fungus *Venturia inaequalis*. *Fungal Genetics and Biology*, **44**, 284–292.
- Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation—the correlated allele frequencies model revisited. *Bioinformatics*, **24**, 2222–2228.
- Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.
- Guillot G, Mortier F, Estoup A (2005b) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.
- Halkett F, Simon JC, Balloux F (2005a) Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, **20**, 194–201.
- Halkett F, Plantegenest M, Prunier-Leterme N, Mieuxet L, Delmotte F, Simon JC (2005b) Admixed sexual and facultatively asexual aphid lineages at mating sites. *Molecular Ecology*, **14**, 325–336.
- Halkett F, Kindlmann P, Plantegenest M, Sunnucks P, Simon JC (2006) Temporal differentiation and spatial coexistence of sexual and facultative asexual lineages of an aphid species at mating sites. *Journal of Evolutionary Biology*, **19**, 809–815.
- Hansen EM (1987) Speciation in plant pathogenic fungi – the influence of agricultural practice. *Canadian Journal of Plant Pathology*, **9**, 403–410.
- Hovmøller MS, Justesen AF, Brown JKM (2002) Clonality and long-distance migration of *Puccinia striiformis* f. sp. *tritici* in north-west Europe. *Plant Pathology*, **51**, 24–32.
- Imbert E, Lefevre F (2003) Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river system. *Journal of Ecology*, **91**, 447–456.
- Jin Y, Szabo L, Carson M (2010) Century-old mystery of *Puccinia striiformis* life history solved with the identification of *Berberis* spp. as an alternate host. *Phytopathology*, **100**, 432–435.
- Kimmel M, Chakraborty R, King JP, Bamshad M, Watkins WS, Jorde LB (1998) Signatures of population expansion in microsatellite repeat data. *Genetics*, **148**, 1921–1930.



- King JP, Kimmel M, Chakraborty R (2000) A power analysis of microsatellite-based statistics for inferring past population growth. *Molecular Biology and Evolution*, **17**, 1859–1868.
- Kiros-Meles A, Gomez D, McDonald BA, Yahyaoui A, Linde CC (2011) Invasion of *Rhynchosporium commune* onto wild barley in the Middle East. *Biological Invasions*, **13**, 321–330.
- Lê Van A, Durel C, Le Cam B, Caffier V (2011) The threat of wild habitat to scab resistant apple cultivars. *Plant Pathology*, doi: 10.1111/j.1365-3059.2011.02437.x
- Lebarbenchon C, Brown SP, Poulin R, Gauthier-Clerc M, Thomas F (2008) Evolution of pathogens in a man-made world. *Molecular Ecology*, **17**, 475–484.
- Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.
- Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.
- Loader CR (1996) Local likelihood density estimation. *Annals of Statistics*, **24**, 1602–1618.
- Mboup M, Leconte M, Gautier A *et al.* (2009) Evidence of genetic recombination in wheat yellow rust populations of a Chinese overwintering area. *Fungal Genetics and Biology*, **46**, 299–307.
- McDonald B (2010) How can we achieve durable disease resistance in agricultural ecosystems? *New Phytologist*, **185**, 13.
- Montarry J, Glais I, Corbière R, Andrivon D (2008) Adaptation to the most abundant host genotype in an agricultural plant-pathogen system – potato late blight. *Journal of Evolutionary Biology*, **21**, 1397–1407.
- Munkacsí AB, Stoxen S, May G (2008) *Ustilago maydis* populations tracked maize through domestication and cultivation in the Americas. *Proceedings of the Royal Society B-Biological Sciences*, **275**, 1037–1046.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- Parker IM, Gilbert GS (2004) The evolutionary ecology of novel plant-pathogen interactions. *Annual Review of Ecology Evolution and Systematics*, **35**, 675–700.
- 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**, 472–479.
- Pinon J, Frey P (2005) Interactions between poplar clones and *Melampsora larici-populina* populations and their implications for breeding for durable resistance. In: *Rust diseases of Willow and Poplar* (eds Pei MH, McCracken AR). pp. 139–154, CAB International, Wallingford, UK.
- Pinon J, Frey P, Husson C, Schipfer A (1998) Poplar rust (*Melampsora larici-populina*): the development of E4 pathotypes in France since 1994. In: *Proceedings of the First IUFRO Rusts of Forest Trees Conference, Saariselkä, Finland* (eds Jalkanen R, Crane PE, Walla JA, Aalto T), pp. 57–64. Finnish Forest Research Institute, Research Paper 712, Saarijärvi.
- Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.
- Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature*, **457**, 843–848.
- Rizzo DM, Garbelotto M, Hansen EM (2005) *Phytophthora ramorum*: integrative research and management of an emerging pathogen in California and Oregon forests. *Annual Review of Phytopathology*, **43**, 309–335.
- Rousset F (1997) Genetic Differentiation and estimation of gene flow from *F*-Statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*, **13**, 58–62.
- Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
- Samils B, Stepien V, Lagercrantz U, Lascoux M, Gullberg U (2001) Genetic diversity in relation to sexual and asexual reproduction in populations of *Melampsora larici-epitea*. *European Journal of Plant Pathology*, **107**, 871–881.
- Stenberg P, Lundmark M, Saura A (2003) MLGsim: a program for detecting clones using a simulation approach. *Molecular Ecology Notes*, **3**, 329–331.
- Stoeckel S, Grange J, Fernandez-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S (2006) Heterozygote excess in a self-incompatible and partially clonal forest tree species - *Prunus avium* L. *Molecular Ecology*, **15**, 2109–2118.
- Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9440–9445.
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75–100.
- Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution*, **24**, 398–411.
- Villate L, Esmenjaud D, Van Helden M, Stoeckel S, Plantard O (2010) Genetic signature of amphimixis allows for the detection and fine scale localization of sexual reproduction events in a mainly parthenogenetic nematode. *Molecular Ecology*, **19**, 856–873.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Wisler GC, Norris RE (2005) Interactions between weeds and cultivated plants as related to management of plant pathogens. *Weed Sciences*, **53**, 914–917.
- Xhaard C, Andrieux A, Halkett F, Frey P (2009) Characterization of 41 microsatellite loci developed from the genome sequence of the poplar rust fungus, *Melampsora larici-populina*. *Conservation Genetics Resources*, **1**, 21–25.
- Yarwood CE (1970) Man-made plant diseases. *Science*, **168**, 218–220.
- Zaffarano PL, McDonald BA, Linde CC (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution*, **62**, 1418–1436.

---

This study forms part of the PhD research of C.X., which addresses the gene flow between fungal populations of *Melampsora larici-populina* and the genetic consequences of dispersal. F.H. is a population biologist interested in the population genetics and evolutionary biology of fungal plant

pathogens, with a confirmed experience in population genetics of clonal and partially clonal organisms. P.F. has a long-standing knowledge of the biology and genetics of rust fungi. B.B. is a population geneticist of fungal plant pathogens, who did his PhD on the inference of dispersal processes in *M. larici-populina*. P.G. is a population geneticist particularly interested in speciation, adaptive divergence and the evolutionary response of pathogens to the domestication of their hosts. A.A. is in charge of the molecular techniques used for fungus genotyping,

and B.F. is responsible for the phenotypic characterization of rust individuals.

---

### Data accessibility

Microsatellite data (474 individuals genotyped): DRYAD entry doi:10.5061/dryad.022 dv.