Strong phylogeographic co-structure between the anther-smut fungus and its white campion host

Alice Feurtey1, Pierre Gladieux1,2, Michael E. Hood3, Alodie Snirc1, Amandine Cornille1, Lisa Rosenthal3 and Tatiana Giraud1

1Ecologie Systématique Evolutive, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, Orsay 91400, France; 2UMR BGPI, INRA, Montpellier 34398, France; 3Department of Biology, Amherst College, Amherst, MA 01002, USA

Summary

- Although congruence between host and pathogen phylogenies has been extensively investigated, the congruence between host and pathogen genetic structures at the within-species level has received little attention.
- Using an unprecedented and comprehensive collection of associated plant–pathogen samples, we investigated the degree of congruence between the genetic structures across Europe of two evolutionary and ecological model organisms, the anther-smut pathogen Microbotryum lychnidis-dioicae and its host plant Silene latifolia.
- We demonstrated a significant and particularly strong level of host–pathogen co-structure, with three main genetic clusters displaying highly similar spatial ranges in Western Europe, Eastern Europe and Italy, respectively. Correcting for the geographical component of genetic variation, significant correlations were still found between the genetic distances of anther-smut and host populations. Inoculation experiments suggested plant local adaptation, at the cluster level, for resistance to pathogens.
- These findings indicate that the pathogen remained isolated in the same fragmented southern refugia as its host plant during the last glaciation, and that little long-distance dispersal has occurred since the recolonization of Europe for either the plant or the pathogen, despite their known ability to travel across continents. This, together with the inoculation results, suggests that coevolutionary and competitive processes may be drivers of host–pathogen co-structure.

Introduction

Understanding and managing fungal diseases of plants is aided by appreciating the processes that generate and maintain spatial patterns in pathogen diversity (Williams, 2010). Predicting the impacts of coevolving systems, and more generally our ability to manage biological interactions embedded in a complex network of agricultural and natural landscapes, requires knowledge about how spatial heterogeneity, selection and environmental variables contribute to the differentiation of populations (Barrett et al., 2008). The degree of genetic variation in specialized pathogens is particularly influenced by their host’s distribution, abundance and genetic make-up (Wilson et al., 2005). This leads to the expectation that the population structure of pathogens would reflect those of their hosts (Criscione et al., 2005). Under a very similar rationale, host and pathogen phylogenies have long been expected to be congruent (Hafner & Page, 1995), whereas it has come as a recent surprise that host shifts are actually the main diversification processes in the vast majority of host–pathogen associations (de Vienne et al., 2013). Unlike the between-species level, however, few studies have addressed the degrees of congruence between host and pathogen genetic structures within-species, although such studies may also reveal unexpected findings, as did the between-species phylogenetic congruence studies (de Vienne et al., 2013). In addition, within-species coevolution and diversification on different hosts corresponds to different evolutionary mechanisms that can, thus, lead to different patterns, that is, with possibly phylogenetic incongruence at the between-species level (de Vienne et al., 2013) but congruence of population genetic structures at the within-species level.

The few published studies on host–pathogen co-structure have been conducted mostly on parasites of animals (Dybå & Lively, 1996; McCoy et al., 2005; Miura et al., 2006; Nieberding et al., 2008; Rich et al., 2008; Amico & Nickrent, 2009; Nakao et al., 2009), with a few, notable exceptions in plant parasites (Michalakis et al., 1993; Whiteman & Parker, 2005; Wirth et al., 2005; Tsai & Manos, 2010; Magalhaes et al., 2011). The genetic structure of the pathogens has, in these cases, often been found to be congruent with those of their host. Due to shorter generation times and the tendency toward selfing mating systems or partial clonality, pathogen populations have nevertheless often been more subdivided than those of their hosts (Nieberding & Olivieri, 2007; Barrett et al., 2008). This may help in reconstructing the host population history (Nieberding & Olivieri, 2007; Barrett et al., 2008), provided that pathogens have followed the same history as their hosts. Indeed, if pathogens show a stronger
genetic structure than their hosts, analyses of pathogen population subdivision will be more informative about the joint past demographic histories of the two partners than studies conducted using host data alone. In fungal plant pathogens, however, studies of population structure have focused on agents of crop diseases, where the host population is genetically homogeneous and distributed artificially, and the pathogens are dispersed widely, in part by human transport, and are often recently introduced exotic species (Enjalbert et al., 2005; Barres et al., 2008; Fournier & Giraud, 2008; Gladieux et al., 2008; Saleh et al., 2014).

Here, we report a study on the co-phylogeography of the casetrating fungal pathogen Microbotryum lychnidis-dioicae (previously named M. violaceum or Ustilago violacea) and its host plant Silene latifolia. Both are model organisms for a variety of topics in ecology and evolution, and, hence, numerous studies are available on their life history, population biology and natural history (Bernasconi et al., 2009; Gladieux et al., 2015). This system is ideal for studying the co-structure between pathogens and their host plants, given that the fungus is an obligate pathogen which is not capable of persisting on its own in the environment (Hood et al., 2010), is highly host-specific in nature (Vercken et al., 2010; Gladieux et al., 2011), the same vectors disperse the fungus and the host pollen (i.e. the pollinators) (Thrall et al., 1995), Silene and Microbotryum species have identical generation times (i.e. one per year) (Thrall & Jarosz, 1994), and there has been no attempt to modify plant distribution or genotypes or to control the disease because its effects are without economic consequences (Bernasconi et al., 2009). A stronger phylogeographic structure is expected in the fungal pathogen than its host plant because M. lychnidis-dioicae is highly selfing (Hood & Antonovics, 2000; Giraud et al., 2005, 2008), whereas S. latifolia is dioecious and shows inbreeding depression (Teixeira et al., 2009), because the fungus only disperses via pollinators, whereas gene flow in the plant also occurs via its seeds in addition to pollen, and because pollinators discriminate against diseased plants (Shykoff & Bucheli, 1995), likely reducing spore transport compared to pollen flow. Note that M. lychnidis-dioicae, unlike most fungi, cannot proliferate by asexual reproduction beyond a few cells after meiosis, and a sex event (i.e. meiosis followed by syngamy and dikaryotic growth) is required for infecting a new plant. However, the mating system is highly automictic – with mostly diplloid selfing by mating among products of the same meiosis (Oudemans et al., 1998; Hood & Antonovics, 2000; Thomas et al., 2003; Badouin et al., 2015). A stronger population genetic structure has been found in M. lychnidis-dioicae than in S. latifolia, although co-structure has been investigated only at a very small spatial scale so far (Delmotte et al., 1999). Local adaptation in this system has likewise been investigated at a very small scale to date, with local plant genotypes being more resistant to their local anther-smut pathogen than to pathogens from other nearby populations (Kalcz et al., 1999). Differences were found in terms of the percentage of plants that became diseased after experimental inoculations, whereas no gene-for-gene interactions have been detected in this system.

Previous studies on the population genetics of M. lychnidis-dioicae provided one of the most clear-cut examples of phylogeographic structure in pathogens (Vercken et al., 2010; Gladieux et al., 2011, 2013; Fontaine et al., 2013). Clustering analyses revealed the existence of three main genetically distinct groups within M. lychnidis-dioicae, with distributions strikingly similar to the major intraspecific lineages identified in Europe for many temperate plant and animal species (Hewitt, 1999). This within-species structure is believed to reflect recolonization from southern refugia after glaciation. The Eastern and Western clusters were each further split into two groups, consistent with colonization from distinct refugia located in the Balkans and further East in Eurasia, and survival of the pathogen in distinct regional refugia in Western Europe.

Phylogeographic analyses of chloroplast genetic data in the host plant, S. latifolia, also revealed patterns of population structure consistent with postglacial expansion from distinct Mediterranean refugia, although these were less clear-cut than for the anther-smut fungus, with large geographical overlaps of maternal lineages (Taylor & Keller, 2007). However, thorough comparisons between host and pathogen structures were prohibited by the facts that the plant and pathogen samples were not collected at the same localities, and that a selective sweep in S. latifolia chloroplast DNA may have obscured historical patterns (Muir & Filatov, 2007). Other studies using microsatellite markers in the host plant S. latifolia were done, but at smaller geographical scales (Delmotte et al., 1999; Barluenga et al., 2011; Magalhaes et al., 2011) or with less dense sampling (Keller et al., 2012). Nevertheless, a clear differentiation was found between Eastern and Western European clusters (Keller et al., 2012).

Thanks to a collection of associated plant/pathogen samples (fungal spores and plant material from the same disease plants) with an unprecedented sampling of the plant in terms of both density and geographical scale in Europe, in the present study we investigated the co-structure of M. lychnidis-dioicae and S. latifolia using nuclear microsatellite markers for both the plant and the fungus, and we used inoculation experiments to test for the existence of local adaptation at the cluster level. More specifically, we addressed the following questions: Can we detect footprints of three glacial refugia in the plant population structure as occur in its fungal pathogen? How congruent are the population structures of the host and the pathogen at the European scale? Are pairwise genetic distances between pathogen individuals and host individuals significantly correlated? In this case, does the correlation still hold when removing the effect of isolation-by-distance: are the host and pathogen population genetic structures more congruent than expected if they were just both affected by the same drift–migration equilibrium after similar postglacial recolonization histories? Is there experimental evidence of local adaptation in the host–pathogen association that may contribute to congruent spatial structure at the cluster level?

Materials and Methods

Sampling, genotyping and species identification

We used the same diseased Silene latifolia plants from which the anther-smut samples had been genotyped previously (Vercken et al., 2010; Gladieux et al., 2011, 2013; Fontaine et al., 2013). Clustering analyses revealed the existence of three main genetically distinct groups within M. lychnidis-dioicae, with distributions strikingly similar to the major intraspecific lineages identified in Europe for many temperate plant and animal species (Hewitt, 1999). This within-species structure is believed to reflect recolonization from southern refugia after glaciation. The Eastern and Western clusters were each further split into two groups, consistent with colonization from distinct refugia located in the Balkans and further East in Eurasia, and survival of the pathogen in distinct regional refugia in Western Europe.

Phylogeographic analyses of chloroplast genetic data in the host plant, S. latifolia, also revealed patterns of population structure consistent with postglacial expansion from distinct Mediterranean refugia, although these were less clear-cut than for the anther-smut fungus, with large geographical overlaps of maternal lineages (Taylor & Keller, 2007). However, thorough comparisons between host and pathogen structures were prohibited by the facts that the plant and pathogen samples were not collected at the same localities, and that a selective sweep in S. latifolia chloroplast DNA may have obscured historical patterns (Muir & Filatov, 2007). Other studies using microsatellite markers in the host plant S. latifolia were done, but at smaller geographical scales (Delmotte et al., 1999; Barluenga et al., 2011; Magalhaes et al., 2011) or with less dense sampling (Keller et al., 2012). Nevertheless, a clear differentiation was found between Eastern and Western European clusters (Keller et al., 2012).

Thanks to a collection of associated plant/pathogen samples (fungal spores and plant material from the same disease plants) with an unprecedented sampling of the plant in terms of both density and geographical scale in Europe, in the present study we investigated the co-structure of M. lychnidis-dioicae and S. latifolia using nuclear microsatellite markers for both the plant and the fungus, and we used inoculation experiments to test for the existence of local adaptation at the cluster level. More specifically, we addressed the following questions: Can we detect footprints of three glacial refugia in the plant population structure as occur in its fungal pathogen? How congruent are the population structures of the host and the pathogen at the European scale? Are pairwise genetic distances between pathogen individuals and host individuals significantly correlated? In this case, does the correlation still hold when removing the effect of isolation-by-distance: are the host and pathogen population genetic structures more congruent than expected if they were just both affected by the same drift–migration equilibrium after similar postglacial recolonization histories? Is there experimental evidence of local adaptation in the host–pathogen association that may contribute to congruent spatial structure at the cluster level?
et al., 2010). Because Microbotryum spores were collected in flower buds of S. latifolia, samples included both materials from the fungus and from the parasitized plants. DNA was extracted using the Nucleospin Plant II kit following manufacturer instructions (Macherey Nagel R). We genotyped the plants using 16 microsatellite markers distributed across the genome (Moccia et al., 2009). Microsatellite markers were amplified by multiplex PCR, with the Multiplex PCR Kit (Qiagen) following manufacturer instructions (for details of the multiplex PCR programs used, see Supporting Information Tables S1 and S2). Genotyping was performed on an ABI Prism X3730XL at the Gentiane Genotyping Platform (INRA, Clermont, France). Alleles were scored with GeneMapper v.4.0 (Applied Biosystems).

For the anther-smut samples, we used a previously published dataset of genotypes produced with a set of 11 microsatellite markers (Vercken et al., 2010). The Microbotryum dataset was rendered haploid by randomly removing one allele in the rare cases of heterozygosity (Vercken et al., 2010). Indeed, individuals are almost completely homozygous because of high selfing rates (Hood & Antonovics, 2000; Giraud, 2004; Giraud et al., 2005, 2008) and this may bias analyses of population subdivision using Bayesian clustering methods assuming random mating. Apparent heterozygosity may, in principle, result from co-infection of the plant by multiple genotypes (Tollenaere et al., 2012), but previous studies have shown that a given S. latifolia bud typically hosts a single M. lychnidis-dioica genotype, even in cases of multiple infections at the plant scale (López-Villavicencio et al., 2007).

For all the analyses, we kept only the individuals (both for the fungus and the plant) successfully genotyped at eight loci or more.

Our plant dataset contained individuals that were identified by the samplers as S. latifolia (n = 592 individuals), S. dioica (n = 74) and putative hybrids between these species (n = 50). The anther-smut fungi dataset initially contained samples that were identified by the samplers as collected on S. latifolia (n = 669), on S. dioica (n = 371) and on hybrids between the two plant species (n = 30). We first ran analyses to ensure that our samples all belonged to the species studied, S. latifolia for the plant and M. lychnidis-dioicae for the fungus. Indeed, the sister plant species S. dioica and S. latifolia have largely overlapping distributions and S. dioica is also affected by anther-smut disease caused by M. silenes-dioicae, the sister species of M. lychnidis-dioicae. Both of the plant and the anther-smut fungi pairs hybridize in nature, and plant and fungal hybrids are difficult to detect morphologically (Kartenberg & Favre, 2008; Vercken et al., 2010; Gladieux et al., 2011). We therefore used Bayesian multilocus genotype assignments to detect potential hybrids in plants and anther-smut fungi (Falush et al., 2003). We identified 578 and 679 genotypes as ‘pure’ representatives of S. latifolia and M. lychnidis-dioicae, respectively (Notes S1; Fig. S1). Only these ‘pure’ individuals were retained for subsequent analyses, from which we could obtain both the genotype of the plant and of its associated anther-smut pathogen for 391 samples.

Population structure

We investigated the population structure within the M. lychnidis-dioicae and S. latifolia species. First, we used the model-based Bayesian clustering methods implemented in TESS v.2.3 (Guillot, 2009). Analyses were performed anew on the genetic marker dataset created for the fungus in our previous study (Vercken et al., 2010) in order to have the same software tools applied on the plant and fungal datasets. We kept only one representative of each genotype in the dataset for clustering analyses and assigned the repeated genotypes to the same cluster as the corresponding representative genotypes. We used the conditional auto-regressive (CAR) Gaussian model of admixture with linear trend surface, setting the spatial interaction parameter (ρ) to 0.6. These parameters (ρ and trend) affect the weight given to spatial distance when clustering genotypes. TESS outputs were processed with CLUMPP v.1.1.2 (Jakobsson & Rosenberg, 2007), to identify distinct clustering solutions in the replicated runs for each K value.

We also performed population structure analyses not relying on any population genetic model and therefore not affected by deviations from Hardy–Weinberg assumptions, such as those caused by selfing. Principal component analyses (PCA) were performed using the ade4 R package and discriminant analyses of principal component (DAPC) using the adegenet R package with the R software (find_clusters function).

For both the host and the parasite, we tested the existence of isolation-by-distance (IBD) patterns by assessing the significance of the correlations between matrices of genetic and geographic distances, using a Mantel test (100 000 resamples for the null distribution) as implemented in the ncf R package (Paradis et al., 2004). Genetic distances between individuals were calculated using Nei’s Da estimator as implemented in POPULATIONS v.1.2.32 (Langella, 2002). Geographic distances were computed with TESS (with the option Compute Great Circle Distances). A Mantel test also was performed to evaluate the significance of the correlation between the genetic distances between pathogen individual pairs and plant individual pairs.

In order to investigate the congruence of genetic structure between the host and the pathogen, we performed partial Mantel tests (100 000 resamples for the null distribution) using the ncf R package to assess the significance of the correlation of pairwise genetic distances between pathogen individuals and plant individuals, as above, but removing the effect of the correlation between genetic and geographic distances.

The genetic variation for both plants and pathogens was quantified using the allelic richness, A_r, calculated with ADZE v.1.0 (Szpiech et al., 2008). Expected heterozygosity (H_e) and observed heterozygosity (H_o) and FIS were computed using GENEPOP v.4.3 (Rousset, 2008). The statistical analyses on allelic richness and H_e (correlations and Wilcoxon signed rank tests) were performed with JMP v.9 (SAS Institute Inc., Cary, NC, USA) using the value of the statistics per marker.
Experimental inoculations

In order to generate host material for the inoculation experiment, seeds of *S. latifolia* were produced under glasshouse conditions from plants originating from field-collected seeds. Each location was represented by a pooled set of seeds from at least three crosses between male and female parental plants. Because sufficient parental material was not available from all regions (e.g. the Southern cluster), we could not perform a full design inoculation experiment. For each available seed–fungus population pair, inoculations were performed using a local pathogen strain and a strain chosen at random among available strains from a different cluster (Table S3). Inoculum suspensions were prepared from field-collected teliospores of *M. lychnidis-dioica*, suspended in water at a concentration of 500 spores ml⁻¹. Teliospore germination was confirmed to be >90% by plating on potato dextrose agar and examination after 24 h incubation.

Inoculation procedures followed prior studies (Hood & Antonovics, 2000). Briefly, seeds were germinated in Petri dishes on 0.8% agar medium containing Murashige and Skoog (MS) basal salt. After 10 d of growth, the apical meristems were exposed by expansion of the cotyledons and 4 μl of inoculum suspension was applied. Inoculated plants were maintained at 15°C for 3 d, and then transplanted to soil and maintained under glasshouse conditions until flowering. Infection status was assessed by the presence of *M. lychnidis-dioica* spores in the flowers.

For each source of seeds, experimental treatments included inoculation with *M. lychnidis-dioica* collected from the same location as the source of seeds or with *M. lychnidis-dioica* from one of the other identified clusters (Table S3). Two hundred seeds per location were used, with one hundred seedlings each assigned randomly to one of the two treatments (local vs intercluster inoculum type). Positions of the plants in the Petri dishes and the glasshouse were randomized. A logistic regression was used to test for differences in disease rates between treatments, performed with JMP v.9 (SAS Institute).

Results

Population structure and IBD in the host plant *Silene latifolia*

The population structure of the host plant *S. latifolia* (*n* = 578) was analysed using the TESS Bayesian clustering method. The Deviance Information Criterion (DIC), a statistical measure of the model prediction abilities for TESS, sharply dropped at *K* = 3 and decreased much more slowly with further increase in *K* (Fig. S2). This indicates that the three-cluster structure is the strongest genetic structure. In fact, the haplotypes showed only three clusters even at higher *K* values (Fig. S3). Also consistent with this finding, PCA and DAPC clearly separated the three same genetic clusters (Figs 1a, S4), and no further clear and geographically relevant cluster could be identified at higher *K* using DAPC or looking at further axes in the PCA. A first cluster was found in France, the UK, Belgium and Germany (hereafter called the ‘Western’ cluster; *n* = 319), the second cluster in Italy (hereafter called the ‘Southern’ cluster; *n* = 91), and the third one in Hungary, Ukraine, Russia and the Czech Republic (hereafter called the ‘Eastern’ cluster, *n* = 132) (Fig. 1b,c). We found 36 admixed individuals that could not be assigned to any cluster – genotypes whose highest membership coefficient was <0.8. Half of these admixed individuals came from locations in contact areas among clusters.

The correlation between the matrices of genetic and geographic distances between pairs of *S. latifolia* individuals was significant, at the European scale (*r* = 0.3351, *P* < 0.0001), as well as within each cluster (Mantel tests; Eastern: *r* = 0.1486, *P* < 0.0001; Southern: *r* = 0.2091, *P* < 0.0001; Western *r* = 0.1151, *P* < 0.0001), indicating IBD patterns at both spatial scales.

Allelic richness (*Ar*), expected heterozygosity (*He*) and inbreeding index *F*ᵢₛ were computed for the tree genetic clusters (Table S4). The Eastern cluster appeared the most genetically diverse but no significant differences were detected between clusters. *F*ᵢₛ values were relatively high for a dioecious plant (*F*ᵢₛ > 0.16).

Population structure and IBD in the fungal pathogen *Microbotryum lychnidis-dioicae*

Genetic structure patterns for *M. lychnidis-dioicae* (*n* = 679) was also investigated using TESS, and gave results very close to the clustering results found in the host plant. Three clusters were differentiated at *K* = 3: one in Western (*n* = 308), one in Southern (mainly Italy, *n* = 160) and one in Eastern Europe (*n* = 178) (Fig. 2a,b). Only 33 individuals could not be assigned to any cluster, half of which came from contact areas between two clusters. The Deviance Information Criterion (DIC) decreased steadily with increasing number of *K*, with no sharp drop (Fig. S2b). The PCA, however, showed that the structure at *K* = 3 was the strongest, with three distinct clusters corresponding to those obtained with the TESS software and a separation of the Western cluster by the first axis, then a subsequent split into Southern and Eastern clusters along the second axis (Fig. 2c).

The correlations between the matrices of genetic and geographic distances between pairs of *M. lychnidis-dioicae* were significant at the European scale as well as within clusters (Mantel tests; European scale, *r* = 0.3937, *P* = 0.0002; Eastern cluster, *r* = 0.3459, *P* = 0.0002; Southern cluster, *r* = 0.2347, *P* = 0.0002). This indicated IBD patterns in *M. lychnidis-dioicae*, both within clusters and at the European scale.

The allelic richness and the expected heterozygosity in *M. lychnidis-dioicae* were much lower than in the plant (Table S5), and no significant differences were detected between clusters. *F*ᵢₛ values were very high, as expected for this highly selfing fungus.

In agreement with the steady decline of the DIC value from TESS, we found a finer population subdivision, with well-defined
At $K=3$, seven geographically circumscribed clusters were thus identified by both model-based and nonmodel-based methods (Figs S5, S6).

**Co-structure between the host *Silene latifolia* and pathogen *Microbotryum lychnidis-dioicae***

At $K=3$, the host and pathogen population structures appeared highly similar (Fig. S3a,b). In particular, the limits of the *M. lychnidis-dioicae* clusters appeared extremely congruent with those found in *S. latifolia*. Of the 80 sampling localities where at least two joint samples of both pathogen and host were available, only five were assigned to different clusters between the host and the pathogen.

Microsatellite data were available for 391 associated pairs of both hosts and pathogens (i.e. the genotypes from the fungus and its host plant individual). Directly associated host and pathogen samples, collected on the same plant individual, predominantly were assigned to the same clusters for the host and the pathogen (i.e. Western, Eastern or Southern cluster). In only 8 (2%) host and pathogen pairs, the plant and fungal individuals were assigned to different clusters. In three host and pathogen pairs, the plant and fungal individuals were both admixed (i.e. 8–9% of the admixed plant or fungal individuals), all of these being located at the suture zones between the Western and Southern clusters.

The correlations between genetic distance matrices for the host and the pathogen were significant at the European scale as well as within clusters (Mantel tests; European scale, $r=0.3505$, $P<0.0001$, Fig. 4; Eastern, $r=0.1389$, $P<0.0001$; Southern, $r=0.1518$, $P=0.0003$; Western, $r=0.0837$, $P=0.0081$).

A partial Mantel test was then performed for testing the congruence between the genetic structures of the host and the pathogen while controlling for the effect of isolation by geographic distance, using the two matrices of genetic distances (host and pathogen) and a matrix of geographic distance between pairs of samples. The partial Mantel test at the European scale was again significant ($r=0.2277$, $P<0.0001$), indicating that the genetically closer two *S. latifolia* individuals were, the genetically closer their anther-smut pathogens were, independently of the shared pattern of IBD. With a threshold of 0.8 for assignation of individuals to a genetic cluster, the partial Mantel test was significant for the Eastern cluster ($r=0.0840$, $P<0.0001$) and close to significance for the two other clusters (Western: $r=0.0450$; $P=0.0903$; Southern: $r=0.0710$, $P=0.0538$). When lowering the threshold of assignation to a cluster to 0.5 (thus increasing the number of sample pairs and adding
Fig. 2 Genetic structure of the anther-smut fungus, *Microbotryum lychnidis-dioicae*, at $K = 3$. (a) Map of mean cluster membership proportions per locality inferred with TESS at $K = 3$. The pie diameter reflects the number of individuals sampled in the corresponding locality: the smallest circles correspond to $n \leq 3$ and the largest circles to $n \geq 16$. (b) Bar plots of membership proportions per individual (as vertical bars) from the TESS analysis for $K = 3$. (c) Principal component analysis on the microsatellite alleles. The dot colour corresponds to the cluster assigned for each individual by the Bayesian computations. The grey dots are the individuals not assigned to any cluster when using a threshold of 0.8.

Fig. 3 Genetic structure of the host plant, *Silene latifolia*, and its pathogen, *Microbotryum lychnidis-dioicae*, from paired samples. (a) Map of mean membership proportions per localities inferred with TESS for the three-cluster structure for *S. latifolia*. (b) Map of mean membership probabilities per locality inferred with TESS for the three-cluster structure for *M. lychnidis-dioicae*. The pie diameter reflects the number of individuals sampled in the corresponding locality and the colours denote different clusters. Only the individuals for which the host–pathogen pair were available are shown here.
Discussion

Footprints of glacial refugia in the population structure of the host Silene latifolia

A sampling of an unprecedented density and spatial scale for Silene latifolia enabled the investigation of genetic structure across Europe for this important model in ecology and evolution (Bernasconi et al., 2009). The strongest pattern of microsatellite variation corresponded to three genetic clusters: a Western cluster, a Southern cluster in Italy and an Eastern cluster. The pattern obtained here using microsatellite markers appears much more clear-cut and revealed further geographical structure (i.e. the Italian clusters) than the one obtained previously with chloroplast markers (Taylor & Keller, 2007), which showed large geographical overlaps between lineages. This discrepancy reinforces the idea of a selective sweep having occurred throughout several clusters in the chloroplast DNA (Ironside & Filatov, 2005), whereas the nuclear gene pools from different southern glacial refugia remained differentiated after recolonization of northern Europe. The population structure of S. latifolia has been investigated previously using microsatellite markers across Europe, but only the East/West structure was highlighted (Keller et al., 2012); nevertheless that previous study suggested that a finer genetic subdivision was present. In fact, re-analysing this independent dataset, that includes different markers and samples from ours, also reveals the Italian cluster at K>2 (not shown).

The biogeographical pattern we found in S. latifolia is typical of the genetic structure originating from contracted ranges in the south of Europe during the last glacial maximum followed by recolonization northward after climate warming at the beginning of the Holocene. The structure thus indicates that, during the last glaciation, S. latifolia likely survived in three main glacial refugia (the Iberian, the Italian and the Balkan refugia), and then recolonized Northern Europe without much large-scale eastern or western dispersal that would have admixed the different clusters. A similar pattern has been found in numerous animal and plant species, these three regions being the main European refugia where climatic conditions stayed mild enough for temperate species during the glacial phases (Hewitt, 1999, 2004).

The identification of a new differentiated gene pool in S. latifolia will be useful for future studies on the different topics for which this plant is a model, from the study of sex chromosomes and mating systems, to speciation and biotic interactions (Marais et al., 2008; Bernasconi et al., 2009; Blavet et al., 2011; Muyle et al., 2012). The knowledge of diversity and population subdivision is essential for such evolutionary and ecological studies; for example, differentiated clusters may display different degrees of reproductive isolation with closely related species or different degrees of degeneration or suppressed recombination on sex chromosomes. In addition, the invasion of S. latifolia in the United States has been studied only with knowledge of two large genetic clusters in Europe (Keller & Taylor, 2008, 2010; Keller et al., 2009, 2012), and it will be highly informative to include the newly identified Italian cluster in future studies investigating

Inoculation experiments

In order to test the existence of local adaptation in the plant and/or the pathogen at the cluster level, seedlings from one location in the western cluster and three locations in the eastern cluster were inoculated with spores coming either from their local population (hereafter, ‘local treatment’) or from a population belonging to a different cluster, randomly assigned (hereafter, ‘inter-cluster treatment’). The probability of a plant being diseased varied significantly with the treatment (i.e. local vs intercluster; Fig. 5); although the seedling origin was not significant as a main effect, the interaction between the seedling origin and the treatment was significant (Table S6). The disease rates were, however, never lower in inoculations by pathogens from a different cluster than by local pathogens (Fig. 5). Altogether, this means that a given plant better resists its local anther-smut fungi than one from other clusters, but not significantly so in all populations. The results thus indicate that local adaptation by the host plants often occurs.

admixed individuals of plants and pathogens), all the partial Mantel tests were significant (Europe: \( r = 0.2277; P < 0.0001 \); Western: \( r = 0.0665; P = 0.0229 \); Eastern: \( r = 0.0889, P < 0.0001 \); Southern: \( r = 0.0987, P = 0.0089 \)).
invasion dynamics and the evolutionary and ecological processes underlying invasions.

Co-structure between the host *Silene latifolia* and pathogen *Microbotryum lychnidis-dioicae*

In *Microbotryum lychnidis-dioicae*, the strongest genetic structure divided the samples into three genetic clusters, as previously reported using the same microsatellite dataset as well as using nuclear gene sequences (Vercken *et al.*, 2010; Glädieux *et al.*, 2011, 2013). These three genetic clusters closely matched those observed in *S. latifolia*, with the same Southern, Western and Eastern clusters. This indicates that the anther-smut pathogen remained during the last glaciation in all of the three refugia of its host (i.e. in the Iberian, Italian and Balkan peninsulas), leaving no major disease-free area.

Our extensive sampling of host–pathogen pairs from the same individual plants allowed comparison of the population structures of the two interacting species at a very fine level across Europe. The population structures were strikingly congruent between the pathogen and its host. It is noteworthy that the population structures of both *S. latifolia* and its pathogen indicates the rarity of recent effective long-distance dispersal events, whereas both are known to be able to disperse across large distances. For instance, both have invaded large ranges of North America within the last centuries (Keller *et al.*, 2012; Fontaine *et al.*, 2013; Glädieux *et al.*, 2015). In Europe, recolonization after glaciation may also have occurred through long-distance dispersal events (Bialozyt *et al.*, 2006). Yet the three main clusters identified in Europe in the plant and the pathogen were clearly circumscribed geographically, suggesting that they remained isolated one from each other for thousands of generations.

One of the factors possibly explaining these discrepancies is local adaptation to the host. A previous study at a smaller geographical scale showed that the plant was adapted locally to its anther-smut fungus: thus, *S. latifolia* individuals were more resistant to their local *M. lychnidis-dioicae* strains than to geographically more distant strains (Kaltz *et al.*, 1999; Kaltz & Shykoff, 2002). The results presented here suggest that this pattern of local adaptation also holds at the cluster level. Lower resistance to foreign pathogens may cause a form of immigrant sterility by preventing *S. latifolia* genotypes to establish in regions where local *M. lychnidis-dioicae* populations belong to a cluster that is different from the cluster present in their region of origin, as the immigrant plant genotypes would be sterilized more often by local pathogens. The pattern of local adaptation supports suggestions of ongoing cycles of coevolution, with the plant running faster in the arms race thanks to its higher levels of interpopulation migration levels (Kaltz & Shykoff, 2002).

Local adaptation of *S. latifolia* to their anther-smut pathogens could, by contrast, favour long-distance establishment of the fungus, as our results suggest that immunity against foreign pathogens is lower than immunity against local ones. However, another factor may limit pathogen migration and contribute to the relative isolation among geographically-circumscribed genetic clusters of the anther-smut fungus. A competitive exclusion of distantly related genotypes by resident genotypes has been shown previously to occur in *M. lychnidis-dioicae* (Hood, 2003; López-Villavicencio *et al.*, 2007, 2011; Buono *et al.*, 2014). This relatedness-dependent competitive exclusion may prevent long-distance migrants to establish in genetically different clusters. Moreover, for both the host and the pathogen, local adaptation to climatic conditions may also play a role in the low effective dispersal across Europe. If the different genetic clusters are locally adapted to abiotic conditions, migrants may be prevented from establishment by being poorer competitors in abiotic conditions to which they are ill-suited.

The significant correlations between genetic distance matrices for the host and the pathogen further revealed that their population structures were congruent even at the within-cluster scale.

Fig. 5 Disease rates obtained in inoculation experiments of *Silene latifolia* by *Microbotryum lychnidis-dioicae*. The percentage of diseased plants is represented for each treatment, the total number of plants having flowered being indicated at the top of the bars. Seedlings from one location in the Western cluster (left) and three localities within the Eastern cluster (right) have been inoculated with spores coming either from their local populations (black bars) or from a population belonging to a different cluster grey bars); see Supporting Information Table S3 for a description of the inoculation design.
Such highly congruent population structures are also likely due, in part, to the anther-smut fungus being an obligate pathogen (i.e. being unable to grow outside its host plant). In addition, the pathogen spores are dispersed by the same insect vectors serving as pollinators for the plant (Jennersten, 1988), increasing the probability that the host plant and its anther-smut fungus followed the same recolonization pathways at fine scales. It would be interesting in future studies to also assess the genetic co-structure between the plants the fungus and the pollinators. In fact, a previous work at a smaller geographical scale found a correlation between among-population genetic distances of *S. latifolia* and *Hadena biscuris*, a pollinator and seed-eater of the plant (Magalhaes *et al.*, 2011).

The congruence of population structures between *S. latifolia* and its *M. lychnidis-dioicae* pathogen appeared even stronger than what could be expected by geography and dispersal alone. Indeed, the partial Mantel tests revealed that the correlations between genetic distances of pairs of individuals between the host and the pathogen remained significant or close to significant when controlling for isolation by distance effect. This suggests either that similar dispersal pathways have been taken by the host and the pathogen, that were not linear with geographical distance, and/or that coevolution has played a role in the congruence of the population structures. Indeed, the anther-smut fungus may more frequently disperse to host populations that carry similar host genotypes as its population-of-origin on these particular host genotypes. As explained above, genotype-by-genotype effects on disease probability have actually been reported in experiments in the *S. latifolia–M. lychnidis-dioicae* system, showing mostly local adaptation in the host more than in the pathogen (Kaltz *et al.*, 1999; Kaltz & Shykoff, 2002).

**Stronger genetic subdivision in the pathogen *Microbotryum lychnidis-dioicae* than in the host *Silene latifolia***

The strongest genetic structure in *M. lychnidis-dioicae* was the three-cluster structure across Europe, with the Eastern, Northern and Southern clusters. However, TESS detected further subdivisions in the fungus. As suggested previously (Vercken *et al.*, 2010), these subdivisions may be a clue to a more complex history than the classic model of the Iberian, Italian and Balkan refugia. The split in the Eastern cluster is consistent with recolonization from two distinct refugia, one in the Balkans and the other further East. The lack of such subdivision in the host *S. latifolia*, despite more markers being used than in the fungus, and despite the markers being more polymorphic, might suggest that the host plant distribution was larger than that of its anther-smut pathogen during the last glaciation, spanning across both the Balkans and further East, whereas the distribution of *M. lychnidis-dioicae* would have been more fragmented. This would be consistent with contemporaneous observations that large areas colonized by *S. latifolia* are disease-free (T. Giraud, pers. obs.).

An alternative, although not exclusive, interpretation is that several life-history traits of both host and pathogen may contribute to a higher genetic subdivision in the fungus than in the plant, even if their distribution ranges have remained quite similar in the southern refugia. First, *S. latifolia* gene flow is mediated both by dispersal of its seed and pollination, whereas *M. lychnidis-dioicae* gene flow occurs only via spore transport by pollinators as seeds do not carry the fungus (Baker, 1947). In addition, the pollinators discriminate against the infected flowers (Shykoff & Bucheli, 1995) which may further reduce the extent of pathogen gene flow compared to its host. Finally, the mating systems are important life-history traits that may explain lower levels of gene flow among populations in the fungus than in the plant. Although *F_{IS}* values obtained here and previously suggest that some inbreeding occurs in this dioecious plant (Delomote *et al.*, 1999; Magalhaes *et al.*, 2011; Keller *et al.*, 2012) despite dioecy and inbreeding depression (Teixeira *et al.*, 2009), the fungus has a much more closed mating system, being highly selfing (Delomote *et al.*, 1999; Hood & Antonovics, 2000; Giraud, 2004; Giraud *et al.*, 2008). Therefore, it may well be that the stronger subdivision in the anther-smut fungus reflects fragmented ranges in southern refugia and/or more northern refugia, but that the higher level of gene flow in the plant has erased their footprints by homogenizing allele frequencies.

**Conclusion**

In fungal plant pathogens, studies of genetic structures at continental scales have been performed almost exclusively for crop diseases, where the host population is often genetically homogeneous across the landscape and pathogen dispersal is impacted heavily by anthropogenic introductions and regional transport (Enjalbert *et al.*, 2005; Barres *et al.*, 2008; Fournier & Giraud, 2008; Gladieux *et al.*, 2008; Saleh *et al.*, 2014). The pathogen population structures in these cases were either very weak at a continental scale (Barres *et al.*, 2008; Fournier & Giraud, 2008; Saleh *et al.*, 2014) or reflected temperature adaptation (Enjalbert *et al.*, 2005; Mboup *et al.*, 2012). The present study therefore brings novel insights into the co-structure and coevolutionary patterns of plants and their pathogens, in a natural system, revealing strikingly strong congruent subdivisions. This study is one of the first to be conducted at this spatial scale with such comprehensive and paired sampling, associated with inoculation experiments, and thus provides a benchmark to develop future studies. Studies on a variety of other systems are needed to assess whether congruence population structure between host and pathogen is a general phenomenon. The knowledge of the co-structure is essential for understanding process of co-local adaptation between hosts and pathogens (Gandon *et al.*, 1996).

**Acknowledgements**

We thank the Plateforme de Génotypage GENTYANE INRA UMR1095 and William Fontanaud for assistance with genotyping. We thank all those who helped with sampling seeds, plant material and pathogens (Vercken *et al.*, 2010; Gladieux *et al.*, 2015). This work was supported by the European Research
Council (starting grant GenomeFun 309403 to T.G.), the Agence Nationale de la Recherche (grants GANDALF ANR-12-ADAP-0009 and Emerfundis 07-BDIV-003), and the National Science Foundation (grant DEB-1115765 to M.E.H.).

Author contributions

T.G. and M.E.H. designed and supervised the study; A.F. performed the population genetics analyses with the help of T.G., P.G. and A.C.; M.E.H. and L.R. performed the inoculation experiments; A.F. and A.S. genotyped the plants; T.G. and M.E.H. contributed to the funding of the study.

References


Hood ME. 2003. Dynamics of multiple infection and within-host competition by the anther-smut pathogen. American Naturalist 162: 122–133.


Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Identification of species.

Fig. S2 Statistics indicating the number of clusters K corresponding to the strongest genetic structure.

Fig. S3 Membership proportions per Silene latifolia individual from the TESS results, for K = 2 to K = 6.
Fig. S4 Discriminant analyses of principal component (DAPC) on the *Silene latifolia* dataset.

Fig. S5 Membership proportions per *Microbotryum lychnidis-dioicae* individual inferred with TESS from $K = 2$ to $K = 6$.

Fig. S6 Discriminant analyses of principal component (DAPC) on the *Microbotryum lychnidis-dioicae* dataset.

Table S1 Details of the multiplex PCR used to amplify microsatellite markers in *Silene latifolia*

Table S2 Details of the PCR program used to amplify microsatellite markers in *Silene latifolia*

Table S3 Cross-inoculation design, origin of the *Microbotryum lychnidis-dioicae* strains and *Silene latifolia* seeds

Table S4 Summary statistics for *Silene latifolia* for each of the three TESS genetic clusters

Table S5 Summary statistics for *Microbotryum lychnidis-dioicae* for each of the three TESS genetic clusters

Table S6 Logistic regression testing local adaptation at the cluster level in the system *Microbotryum lychnidis-dioicae–Silene latifolia*

Notes S1 Identification of species.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.