Host specialization has important consequences for the diversification and ecological interactions of obligate pathogens. The anther-smut disease of natural plant populations, caused by Microbotryum fungi, has been characterized by specialized host-pathogen interactions, which contribute in part to the isolation among these numerous fungal species. This study investigated the molecular variation of Microbotryum pathogens within the geographic and host-specific distributions on wild Dianthus species in southern European Alps. In contrast to prior studies on this pathogen genus, a range of overlapping host specificities was observed for four delineated Microbotryum lineages on Dianthus hosts, and their frequent co-occurrence within single-host populations was quantified at local and regional scales. In addition to potential consequences for direct pathogen competition, the sympatry of Microbotryum lineages led to hybridization between them in many populations, and these admixed genotypes suffered significant meiotic sterility. Therefore, this investigation of the anther-smut fungi reveals how variation in the degrees of host specificity can have major implications for ecological interactions and genetic integrity of differentiated pathogen lineages.

Keywords: generalist, host range, host shift, introgression, Microbotryum violaceum, pathogen sympathy, secondary contact

Received 19 October 2016; revision received 13 January 2017; accepted 9 February 2017

Introduction

Infectious diseases have important impacts on all natural ecosystems, from effects on host individuals to determining community composition (Benítez et al. 2013; Bagchi et al. 2014; Bever et al. 2015), and pathogens themselves represent a major component of biodiversity (Dobson et al. 2008). The evolution of differentiated pathogen lineages is associated with barriers to gene flow that arise through geographic separation and through specialization onto different hosts (Rice 1987; de Vienne 2013; Kohn 2005; Giraud et al. 2006, 2008a, 2010). However, where ecological specialization is not strong, the processes of natural migration or anthropogenic movement can lead to direct interactions between previously isolated pathogen lineages if they come to co-occur in the same host population. There is great concern over such pathogen secondary contact because hybridization and competition may then favour the emergence of new disease properties, such as increased virulence or with new host affinities (Kulda et al. 1999; Brasier 2000; Newcombe et al. 2000; Lin et al. 2001; Giraud et al. 2008a; Gladieux et al. 2011; Lemaire et al. 2016; Leroy et al. 2016; Stukenbrock 2013, 2016).

Examples of altered disease traits resulting from pathogen secondary contact and hybridization are often
reported in the agricultural and human medical literature (Schardl & Craven 2003; King et al. 2015). Hybrids between serotypes of the human pathogen Cryptococcus neoformans can yield more virulent genotypes (Lin et al. 2007). Secondary contact between differentiated pathogen populations of Venturia inaequalis on apples resulted in new virulent hybrids (Lemaire et al. 2016; Leroy et al. 2016). The hybrid rust pathogen Melampsora x columbiana emerged on poplars bred for resistance to the two parental rust species (Newcombe et al. 2000), and a similar emergence of hybrid powdery mildews occurred on cereals (Menardo et al. 2016). Even in the absence of hybridization, competition between co-occurring pathogens can be a driver of pathogen virulence or diversification. A large body of theoretical work describes how competition might, positively or negatively, impact virulence, and there is support for these effects from experimental inoculation studies (reviewed by Alizon et al. 2013; Bull & Lauring 2014). In contrast, secondary contact between previously isolated pathogen lineages may lead to decreased disease severity because the generated hybrids are less fit and/or because competition between lineages lowers disease prevalence (Le Gac et al. 2007b; de Vienne et al. 2009a, Sloan et al. 2008; Gold et al. 2009). Given the current rates of global change, species redistributions are increasingly frequent, fostering ever more secondary contacts between pathogens (Desprez-Loustau et al. 2007; Gladieux et al. 2011, 2015; Lemaire et al. 2016). However, our ecological understanding of the consequences of such interactions is limited by a lack of studies on natural disease systems where multiple related pathogen species co-occur.

The goal of this study was to determine the geographic and host-specific distributions of anther-smut pathogens on wild Dianthus hosts and to assess the consequences of host–pathogen interactions for pathogen sympatry and hybridization. Anther-smut disease, caused by obligately pathogenic fungi in the genus Microbotryum, has served as an important model for disease ecology (Antonovics et al. 2002; Bernasconi et al. 2009), including the conditions that promote pathogen invasion (Fontaine et al. 2013; Gladieux et al. 2015; Feurer et al. 2016), speciation (Le Gac et al. 2007a,b; Réfrégier et al. 2010) and hybridization (de Vienne et al. 2009a; Gladieux et al. 2011). Most of this previous work has been with anther-smut diseases of a related host genus, Silene, where numerous pathogen species are highly specialized to different hosts and there is little hybridization (Le Gac et al. 2007b; Gladieux et al. 2011; Badouin et al. 2017). The monophyletic clade of anther-smut fungi on the host genus Dianthus is less well studied, but recent work suggests there is generally less host specificity compared to the pathogens on Silene hosts (Kemler et al. 2013). Some Microbotryum pathogens on Dianthus can infect overlapping sets of host species, and thus, multiple fungal species can be found on a given Dianthus host species (Réfrégier et al. 2008; Kemler et al. 2013). In such systems where host specificity is low, there should be greater potential for different pathogen lineages to meet and interact directly in the same host population, and possibly to hybridize.

Here, we first address genetic structure and reproductive barriers among anther-smut fungi infecting Dianthus hosts in the southern European Alps using microsatellite variation and in vitro assays, and we show that there are several differentiated Microbotryum lineages. We then describe the geographic and host-specific distributions of these pathogens. Specifically, the goals of the study were as follows: (i) identify genetically differentiated lineages of Microbotryum genotypes and in particular incorporating population-level sampling, (ii) assess sterility of genotypes resulting from hybridization that may indicate reproductive isolation among the pathogen lineages, (iii) determine pathogen distributions at different spatial scales in terms of host specialization and sympathy and (iv) assess the occurrence and distribution of pathogen genotypes resulting from hybridization, and in particular, whether they are rare or frequent, which may also inform on their fitness relative to parental lineages. This study reveals important components of previously undetected pathogen diversity and provides insights into the dynamics of multihost multipathogen systems at the regional scale.

Methods

Study system

Microbotryum is a genus of basidiomycete fungi, in the subphylum Pucciniomycotina, that causes anther-smut disease on plants mostly in the Caryophyllaceae. In this disease, host pollen is replaced by fungal spores which are then spread primarily by insect pollinators (Fig. 1). Host plants are completely sterilized by infection because the anthers of all flowers are diseased and the ovary also fails to fully develop. The fungus persists only on perennial host species (Hood et al. 2010) and results in a systemic and persistent infections that has minimal effect on plant mortality (Thrall & Jarosz 1994). To complete the life cycle, the diploid smut spores from diseased anthers are deposited on a new host, after which they germinate, undergo meiosis and produce a haploid yeast-like stage with two mating types; mating among haploid cells must occur prior to infection of the new host plant (Giraud et al. 2008a). Prior studies on Silene hosts indicate that a single plant may sometimes harbour multiple genotypes of the pathogen, but there is frequently competitive exclusion, which is stronger

© 2017 John Wiley & Sons Ltd
Microbotryum the infection. occurs only during mating on the host surface prior to genetic exchange between pathogen genotypes, as this encounters inside diseased hosts are not a route to infection. The encounters inside diseased hosts are not a route to genetic exchange between pathogen genotypes, as this occurs only during mating on the host surface prior to infection.

Recent studies have aimed to identify species within the Microbotryum complex (Kemler et al. 2006; Denchev, 2007; Lutz et al. 2008; Denchev et al. 2009; Piątek et al. 2012, 2013). While concordance between multiple gene genealogies has been useful in these efforts, this criterion may not be optimal for young species with either incomplete gene coalescence or ongoing hybridization (reviewed in Cai et al. 2011). Nevertheless, several of the pathogens sampled in Europe from Dianthus hosts species have been named as distinct species: Microbotryum dianthorum, Microbotryum carthusianorum, Microbotryum shykoffianum and Microbotryum superbum (Denchev et al. 2009). In these cases, it has been noted that pathogen sampling is sparse, and therefore, taxonomic revision is tentative (Le Gac et al. 2007b; Kemler et al. 2013).

It has been long recognized that Microbotryum from different host species are relatively specialized (Zillig 1921; Goldschmidt 1928). On the host genus Silene, Microbotryum species are generally host-specific, although there are exceptions as transient host shifts have been recorded (Antonovics et al. 2002; Gladieux et al. 2011). In contrast, on Dianthus multiple Microbotryum species have been reported from the same host species. Reciprocally some of the pathogen lineages can infect multiple host species with some apparent degree of overlapping host ranges (Le Gac et al. 2007b; Refrégier et al. 2008; Kemler et al. 2013).

The host genus Dianthus represents a group of broadly distributed, mostly perennial herbs found commonly in habitats that include subalpine and alpine regions of Eurasia and Africa. Rapid diversification in Europe of over 100 Dianthus species is believed to have coincided with regional climate changes in the Pleistocene and the onset of dry Mediterranean summer conditions (Valente et al. 2010).

Field collections

Surveys for anther-smut disease on Dianthus hosts were carried out in the southern European Alps from 2008 to 2011, in a focal region spanning latitudes 44.0–44.5°N and longitudes 7.0–8.2°E (Fig. 2, S1, Supporting information). A total of 443 pathogen samples were collected and genotyped from 81 populations across six Dianthus species, D. deltoides, D. furcatus, D. pavonius, D. seguieri, D. superbus and D. sylvestris (Table S1, Supporting information). Flowering period largely overlaps for these species during the sampling period of early July to mid-August, with D. deltoides and D. sylvestris flowering marginally earlier in the season. Sampling was carried out in 19 valleys, within which populations/communities of a Dianthus species were defined as collections separated by at least 100 m; a community was a collection site with more than one Dianthus species present and diseased. The Dianthus species were readily identified from morphological differences; in only one location were there a few plants of intermediate form, and no pathogen samples from those plants were included. The average number of pathogen samples per population was 5 (SD = 4.3). To avoid potential contamination from pollinator-dispersed spores, Microbotryum was sampled as the spores present in mature unopened flower buds, one sample per diseased plant. These were stored at 4 °C until used for DNA extraction from excised individual anthers or for use in culturing in vitro.

In 2012, additional samples were collected to determine whether sympatric Microbotryum lineages might co-occur at the level of the multiple infections of individual host plants. From valleys (STA, PRA, and SBM; Table S1, Supporting information) where initial studies revealed multiple Microbotryum lineages, multiple flowering stems per individual plant were sampled as separate flower buds. For the STA valley, 22 plants averaging 3.1 stems were sampled; for PRA valley, 27 plants averaging 2.8 stems; and for SBM valley, 53 plants averaging 2.5 stems. Prior studies have shown that during co-infection different Microbotryum lineages almost never occupy the same flower (Day 1980; Hood 2003).

Microsatellite (Simple Sequence Repeat) genotyping

Eleven microsatellite markers were used by the methods described in Giraud (2004) with the isolation of
diploid DNA of Microbotryum samples from Dianthus. Seven markers were from existing libraries (Giraud et al. 2002, 2008b), and four new markers were developed: mvD3, mvD5, mvD11 and mvD23 (Table S2, Supporting information). The new markers were found using the Repeat Finder program (http://www.proweb.org/proweb/Tools/selfblast.html) to search low coverage genome sequence data of Microbotryum isolated from Dianthus (454 sequencing at ca. 3× genome coverage; mean DNA fragments length of ca. 250 bp). PCR primers for 24 loci were designed using PRIMER3 (Rozen & Skaletsky 2000) to flank di- or trinucleotide repeat motifs of at least six iterations. Preliminary assays selected four marker loci showing polymorphism and amplification across all the Microbotryum samples.

Amplification of microsatellite loci was conducted using Chelex-extracted DNA (Bucheli et al. 2001) and fluorescent-labelled PCR primers. PCR was performed in a final volume of 10 µL containing 1 × Phusion HF buffer (New England Biolabs), 0.2 mmol/L deoxynucleoside-5′-triphosphate (Promega), 0.25 µmol/L forward and reverse primers, 1 U/µL Phusion HF DNA polymerase (New England Biolabs) and 0.8 µL of template DNA. Separate PCRs were started with an initialization step for 30 s at 98 °C, 50 cycles of 10 s at 98 °C, 30 s at the annealing temperature (Table S2, Supporting information), 30 s at 72 °C and 1 final extension cycle of 5 min at 72 °C. PCRs with different primer dyes (Table S2, Supporting information) were multiplexed, and a 600 LIZ dye size standard was added. Fragment analysis was run at the UMass Genomics Resources Lab on Applied Biosystems 3130 and 3130 xl Genetic Analyzers run with Applied Biosystems five-dye chemistry. Alleles were scored with GENEMAPPER 4.0, and samples scored for the majority of loci (i.e. ≥6) were retained for the following analyses.

Fig. 2 Distributions across southwestern European Alps of the Microbotryum pathogens and Dianthus host populations. Populations are represented by pie graphs, scaled by sample size (ranging from 1 to 28 samples). Locations in the map are approximate, adjusted to allow nonoverlapping representation of all pie graphs. (A) Population structure of 443 pathogen samples was inferred with the Bayesian clustering algorithm implemented in STRUCTURE, where the four genetic clusters are indicated by coloured wedges. Grey wedges indicate samples that did not meet the 0.9 confidence threshold for cluster assignment. (B) Map of the populations of Dianthus host species that were sampled for Microbotryum pathogens. Different host species are indicated by shaded pie graphs, sized in proportion to numbers of individuals sampled.
Population genetic subdivision and assessment of hybridization

Individual-based Bayesian clustering methods, implemented in STRUCTURE 2.3.3 (Pritchard et al. 2010), were used to investigate genetic subdivision of the 2008–2011 pathogen collections. This method uses Markov chain Monte Carlo (MCMC) simulations to infer the assignment of genotypes into different numbers (K) of proposed clusters. The underlying algorithms minimize deviations from Hardy–Weinberg structure and linkage disequilibria within each cluster. Ten independently replicated runs of the analysis were carried out for successive K values (2 ≤ K ≤ 9), with 1 000 000 MCMC iterations after a burn-in of 100 000 steps. Outputs were processed with CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007), for the identification of potentially distinct clusters among replicated runs for each K value. A G′ statistic greater than 80% was used to assign replicates to a common STRUCTURE mode. When multiple STRUCTURE modes were found (K > 3), we compared alternative modes. At each K, the different modes were not substantially different (H′ values > 0.70). We therefore only kept the major mode. To help assess the most relevant value of K, we determined the amount of additional information explained by increasing K using the rate of change in the log probability of data between successive K values (ΔK statistic) (Evanno et al. 2005). Individuals with a membership coefficient >0.90 to a given cluster were assigned to that cluster, while samples with lower values were considered admixed genotypes as the result of hybridization between clusters. In addition to ΔK statistics and the proportion of individual samples assigned to clusters across K-values, we used methods assessing genetic subdivision independently of a population genetic model: factorial correspondence analysis (FCA), centred with respect to clusters, was performed with GENETIX V4.05 (Belkhir et al. 1996), and discriminant analysis of principal component (DAPC) was performed using the adegenet R package with R software (dapc function).

Patterns of isolation by distance were determined by assessing the correlation between matrices of geographic distances and genetic distances within groups of the cluster-assigned pathogens. Mantel test with 1000 random permutations was performed between the individual coefficients of relatedness Fst (Loiselle et al. 1995) and a matrix of the natural logarithm of geographic distances between collection sites. These analyses were performed with SPAGEDi 1.3 (Hardy & Vekemans 2002).

Also, the spatial aggregation of pathogen clusters or host species was assessed using Mantel tests on correlations between the geographic distances matrix between populations or valleys and the matrix of shared presence/absence of the cluster-assigned pathogens or host species among those populations or valleys. Using the most relevant K, as assessed based on STRUCTURE, FCA and sterility analyses, and cluster-assigned samples (i.e. membership coefficient >0.90 to a given cluster), the presence of pathogen clusters was scored as a binary variable using GENAIEX 6.5 (Peakall & Smouse 2012).

We then tested whether the variation of the pathogen cluster distribution was similar among host species at the scales of individual plants, populations and valley; at the latter two spatial scales, the three of 81 populations or four of 19 valleys with multiple host species were excluded. For this analysis, the spss 24 (SPSS Inc.) statistical software was used, conducting Fisher’s exact test with Monte Carlo simulation (10 000 iterations) and GLMz (Poisson probability distribution, log-linear model) for taking into account any confounding effects of host and location. Comparison of the proportions (arc sine-transformed) of samples that were admixed genotypes from populations with or without combinations of cluster-assigned pathogens was performed using a univariate general linear model in SPSS with numbers of samples as a covariate.

Descriptive statistics

Genetic polymorphism of the clusters inferred as described above (i.e. clusters inferred by STRUCTURE, FCA and sterility analyses) was quantified as allelic richness (Ar) and private allele richness (Ap) calculated with ADZE (Szpiech et al. 2008). Allelic richness and private allele richness were computed using sample sizes of N = 10 (five individuals × two chromosomes) as the minimum common sample size across clusters for which all microsatellite loci had been successfully genotyped, and clusters were compared using Wilcoxon signed-rank tests. Heterozygosity, Weir and Cockerham’s F-statistics and deviations from Hardy–Weinberg expectations were calculated with GENEPop 4.0 (Raymond and Rousset 1995; Rousset 2008). Population-specific computations excluded admixed genotypes (i.e. retaining only samples with a membership coefficient >0.9 to any cluster), leading to a sample size of N = 318 individuals.

Sterility assays of admixed genotypes

Hybrid sterility, as a measure of reproductive isolation between clusters that were identified with the STRUCTURE and FCA analyses, was investigated by germination assays of Microbotryum spores from the same field-collected flowers as were genotyped. Microbotryum undergoes meiosis following germination of the diploid
spores (teliospores) to produce the tetrad of four haploid yeast-like sporidia (basidiospores) extending from the basidium (as illustrated in Hood & Antonovics 2000). Field-collected diploid spores were germinated on potato dextrose agar and incubated at 22 °C for 24–30 h. The four postmeiotic haploid sporidia were separated by micromanipulation to areas on the agar surface away from other growing cultures, and the proportion of the sporidia that exhibited colony formation (haploid growth) was determined after 48 h at 22 °C.

The sterility assays were carried out for five meiotic tetrads for up to 15 samples per each of four clusters and per each of the six possible types of admixed individuals between them. The putative parental clusters of the admixed genotypes were determined as the two clusters to which STRUCTURE analysis membership coefficients were greatest. Arcsine-transformed data were used to determine 95% confidence intervals for the proportions of inviable meiotic products, which were back-transformed for plotting (Bolton & Bon 2009). The ability of yeast-like gametes to grow from the isolated postmeiotic cells was compared to nonadmixed individuals using the t-test for independence on arcsine-transformed proportions. Previous studies have shown that hybrids between different Microbotryum species may be viable but often suffer from sterility, due to the different karyotypes between species (de Vienne et al. 2009a).

DNA sequence analyses

For comparison with prior studies on Microbotryum from Dianthus, which used DNA sequence-based phylogenetics, two types of nuclear loci were investigated as follows: (i) the internal transcribed spacer (ITS) region of the multicopy ribosomal RNA gene as used by Kemler et al. (2013), and (ii) elongation factor 1α (Ef1α) as used by Le Gac et al. (2007b), which were the two markers providing the best resolution among Microbotryum samples from Dianthus. We conducted bidirectional DNA sequencing for Ef1α and ITS for approximately 100 samples for which STRUCTURE analysis cluster membership coefficients were >0.90 for the most likely K value and which spanned most collection localities and hosts-of-origin (Table S3, Supporting information). DNA sequences were also retrieved from these prior studies for comparison.

In addition, ITS sequencing was performed on the Microbotryum samples collected as multiple stems from individual diseased host plants. Identities of the pathogen were assigned by their ITS sequence variants, and due to the multicopy nature of ITS loci, admixed pathogens between lineages were identified by intragenomic sequence polymorphisms at particular nucleotide positions in DNA sequencing electropherograms (James et al. 2009).

Results

Characterization of Microbotryum lineages

Analysis of microsatellite data for the 443 samples of Microbotryum infecting Dianthus hosts indicated the existence of four main pathogen clusters. Summary statistics on the microsatellite loci are presented in Table S4 (Supporting information). Our conclusion on the most relevant number of clusters was based upon a combination of metrics. The peak of AK value was found at K = 3, and it remained high for K = 4 and 5 (Fig. S2, Supporting information). For K-values of 2, 3 and 4, the proportion of samples assigned to clusters by membership coefficients >0.90 was maximized in the range of 0.74–0.72, but then dropped markedly to 0.59 and 0.51 for K = 5 to 7 (Figs S2 and S3, Supporting information). K = 4 was thus the highest K value defining well-delimited clusters, representing the finest strong genetic subdivision in the data set, which may correspond to lineages isolated from extensive genetic exchange. This was also supported by the factorial correspondence analyses (FCA), which indicated strong separation of clusters at K = 4 (Fig. S5, Supporting information), while at K = 5, the fifth clusters were poorly resolved by FCA. Discriminant analyses of the principal component (DAPC) similarly showed that clusters at K = 4 were well separated by confidence inertia ellipses (Fig. S6, Supporting information). Hybrid sterility analyses (see below) also indicated that the four clusters each contributed to the observed structure through partial meiotic failure of admixed genotypes. Analyses at K = 4 clusters are therefore presented below.

At the cluster assignment at K = 4, 318 of the 443 pathogen samples (72%) had membership coefficients greater than the assignment threshold (>0.90), and 125 were considered to be admixed genotypes (Fig. S3, Supporting information). Among the 318 cluster-assigned Microbotryum samples, 143 samples were assigned to Cluster 1 (Fig 2A, red colour), 69 to Cluster 2 (Fig 2A, blue) and 53 to each of Cluster 3 (Fig 2A, green) and Cluster 4 (Fig 2A, yellow). The four clusters displayed similar significant deficiencies in heterozygotes as is characteristic Microbotryum fungi due to frequent selfing (Fontaine et al. 2013; Bueker et al. 2016) (Table 1). There was also similarity among clusters in patterns of allelic diversity (allele richness and private allele richness), with significant differences for these metrics present only between clusters 1 and 3 (Tables 1 and S8, Supporting information). Levels of divergence in allelic frequencies between the clusters (FST) were strong for all pairwise comparisons, ranging from 0.268 to 0.399 (Table 2).
Differentiation of the four identified pathogen clusters was supported by significant sterility of the field-collected admixed genotypes. Cluster membership coefficients less than the 0.90 threshold significantly predicted the failure of yeast-like gametes to grow from the isolated postmeiotic cells compared to nonadmixed individuals (t-test for independence on transformed proportions, $t(8) = -4.57$, $P = 0.002$). Compared to samples from the four clusters, each of the six pairwise admixed types between them showed more failure of haploid growth for the immediate products of meiosis (Fig. 3, Fig. S7, Supporting information).

Analysis of DNA sequence data from the two nuclear loci (i.e. $Ef1\alpha$ and ITS) allowed comparison of our population-based microsatellite study with prior delimitation among $Microbotryum$ pathogens on $Dianthus$ hosts. Identical $Ef1\alpha$ sequences as three species of $Microbotryum$ on $Dianthus$ identified by Le Gac et al. (2007b) were represented in the current study (Fig. 4A). However, two of these three $Ef1\alpha$ sequence variants, for $Microbotryum$ shykoffianum and $M$. carthusianum, only encapsulated a small numbers of samples in the current study, and clusters 1 and 4 of the current study seemed to correspond to lineages distinct from those described previously based on multiple gene genealogies (Le Gac et al. 2007b). Of the four main clades identified based on ITS sequence data in Kemler et al. (2013), only a single one, called Group 3, was represented in our data set (Fig. 4B).

The consistency with which DNA sequence-based variation discriminated microsatellite-based cluster assignments differed between $Ef1\alpha$ and ITS loci. In both the $Ef1\alpha$ and ITS phylogenies, Cluster 3 was most distant from the remaining samples (Fig. 4A,B). $Ef1\alpha$ tended to produce identical sequences for all samples of Cluster 1 and identical sequences for the majority of Cluster 2 and to also delimit these two clusters well. Cluster 4 was more dispersed among $Ef1\alpha$ variants, while it was well distinguished by ITS. In contrast, clusters 1 and 2 were less well delimited by ITS than $Ef1\alpha$.

**Table 1** Genetic polymorphism of differentiation within cluster ($K = 4$) in $Microbotryum$ fungi from $Dianthus$ based on microsatellite variation

<table>
<thead>
<tr>
<th>Clusters</th>
<th>1 (red)</th>
<th>2 (blue)</th>
<th>3 (green)</th>
<th>4 (yellow)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>143</td>
<td>69</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>$H_o$</td>
<td>0.18</td>
<td>0.17</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>$H_E$</td>
<td>0.51</td>
<td>0.63</td>
<td>0.70</td>
<td>0.51</td>
</tr>
<tr>
<td>$F_{IS}$</td>
<td>0.64</td>
<td>0.74</td>
<td>0.76</td>
<td>0.63</td>
</tr>
<tr>
<td>$A_R (n = 10)$</td>
<td>3.59</td>
<td>5.17</td>
<td>5.21</td>
<td>3.34</td>
</tr>
<tr>
<td>$A_P (n = 10)$</td>
<td>0.30</td>
<td>1.09</td>
<td>1.27</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Admixed genotypes were not included in this analysis. $N$: Number of samples, $H_o$ and $H_E$: observed and expected heterozygosities, $F_{IS}$: inbreeding coefficient, $A_R$ and $A_P$: allelic richness and private allele richness, respectively, computed using sample sizes of $n = 10$; colour descriptors correspond to their use to indicate clusters in the figures.

*All $F_{IS}$ values were significant at $P < 0.05$.

© 2017 John Wiley & Sons Ltd

**Table 2** Genetic differentiation ($F_{ST}$) values between $Microbotryum$ genetic clusters ($K = 4$) based on microsatellite variation

<table>
<thead>
<tr>
<th>Clusters</th>
<th>1 (red)</th>
<th>2 (blue)</th>
<th>3 (green)</th>
<th>4 (yellow)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (red)</td>
<td></td>
<td>0.276</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (blue)</td>
<td>0.276</td>
<td></td>
<td>0.262</td>
<td></td>
</tr>
<tr>
<td>3 (green)</td>
<td>0.399</td>
<td>0.262</td>
<td></td>
<td>0.362</td>
</tr>
<tr>
<td>4 (yellow)</td>
<td>0.325</td>
<td>0.268</td>
<td>0.362</td>
<td></td>
</tr>
</tbody>
</table>

Admixed genotypes were not included in this analysis, and thus, analyses were conducted on a total of $N = 318$ individuals; colour descriptors correspond to their use to indicate clusters in the figures.
respect to both location and host species (GMLz Wald-
X^2 for collection valley = 43.4, d.f. = 16, P < 0.001, and
for host species = 27.5, d.f. = 4, P < 0.001). Indeed, geo-
graphic proximity was a significant overall predictor of
sites sharing the same Microbotryum lineage (Mantel
test, r = 0.216, P < 0.001 for 81 populations and
r = 0.164, P = 0.048 for 19 valleys). The lineage identi-
fied as Cluster 4 contributed most strongly to this spa-
tial pattern, being largely restricted to the southeastern
portion of the sampling area (Fig. 2A); this pattern was
not due to the restricted distribution of a particular host
lineage to this region (Fig. 2B). Thus, the probability of
Cluster 4’s presence in any two valleys as a function of
the geographic distance between them approached sig-
nificance (Mantel test, r = 0.260, P < 0.053 for 19 val-
leys), but this measure was not significant for the other
pathogen lineages (smallest P-value among other clus-
ters = 0.167). Within each of the four pathogen lineages,
there was no significant isolation by distance (Mantel
tests with P-values within Cluster 1 = 0.44, Cluster
2 = 0.42, Cluster 3 = 0.46, Cluster 4 = 0.18).

There was also a nonrandom distribution of the four
pathogen lineages among host species evident across
hierarchical sampling scales, whether analysing individ-
ual samples (Fig. 5, Fisher’s exact test, Monte Carlo
10 000 iterations, P < 0.001) or treating populations or
valleys as units of observation (Fig. S8, Supporting
information; Fisher’s exact test, Monte Carlo 10 000 iter-
ations, P < 0.001 for populations and P = 0.009 for val-
leys). In particular, Cluster 1 was almost exclusively
associated with Dianthus pavonius; only one out of 143
samples was isolated from another host, Dianthus del-
toides (Fig. 5). Cluster 3 also was most often associated
with Dianthus pavonius, although it was also found on
Dianthus furcatus. The pathogen of Cluster 4 was most
frequently associated with Dianthus seguieri although
also with three other hosts. The pathogen belonging to
Cluster 2 was the most widespread, often on Dianthus
pavonius while also on four other hosts. The distribution
of admixed pathogen samples among hosts is shown in
Fig. S9A (Supporting information).

Dianthus host species were thus associated with mul-
tiple Microbotryum lineages (Fig. 5). Within sampling
locations, different pathogen lineages co-occurred, most
often on the same host species (Fig. 2). Among the 56
populations where more than one cluster-assigned
Microbotryum sample was collected, 17 (30%) contained
multiple pathogen lineages. At the level of valleys, 12
of 19 (63%) contained multiple pathogen lineages.
Pathogen co-occurrence, however, did not frequently

© 2017 John Wiley & Sons Ltd
extend down to the level of co-infecting individual host plants. Within three valleys where pathogen sympathy occurred (STA, SBM, and PRA; Table S1, Supporting information) and where multiple flowering stems were sampled from individual plants, there was only one instance of co-infection by different cluster-assigned pathogen genotypes among 102 plants sampled as multiple stems, averaging three stems per plant.

Sympathy of different diseased host species was markedly less common than sympathy of pathogen lineages; compared to the numbers of sites containing multiple pathogen lineages above, only three local sites sympatric communities of multiple diseased host species and four valleys contained diseased samples from more than one host. Noting that the sample sizes were small, sympathy of diseased host species was not a significant predictor of sympathy for pathogen lineages; only one of the three host-sympatric populations was a location of pathogen sympathy (binomial probability of at least one of three populations having pathogen sympathy, $P = 0.657$), and while each of four valleys with diseased host sympatries was also a location of pathogen sympathy, this was not unexpected (binomial probability, $P = 0.158$).

Regarding the distribution of just the host species among the sampled populations, geographic proximity was a significant predictor of populations sharing the same diseased *Dianthus* host species (Mantel test, $r = 0.167$, $P = 0.010$ for 81 populations), while this host spatial pattern was not significant at the among-valley scale ($r = 0.217$, $P = 0.120$ for 19 valleys) (Fig. 2).

Individuals characterized as admixed genotypes between pathogen lineages were found throughout the sampling region (Fig. 2), in 50 of 81 populations (62%) and in 15 of 19 valleys (79%). Admixed genotypes were found in 90% of the populations with more than five pathogen samples. Somewhat surprisingly, and even controlling for the number of samples collected, the proportions of pathogens that were admixed did not depend significantly upon whether the population had a mixture of assigned pathogen lineages (arcsine-transformed data, $F = 0.972$, $P = 0.329$). However, there was statistical bias among the different *Dianthus* host species for the ratio of admixed pathogens to cluster-assigned pathogens (Fisher’s exact test, Monte Carlo 10 000 iterations, $P = 0.020$; Fig. S9A, Supporting information), driven by somewhat of an excess of admixed genotypes on *D. furcatus*. However, admixed pathogens were not derived from a common ancestral admixture of two other clusters, and the abundances of the six types of admixed pathogens (shown in Fig. S9B, Supporting information) appeared to occur by chance; the types did not differ significantly from the products of relative overall abundances of the four parental pathogen lineages (Fisher’s exact test, Monte Carlo 10 000 iterations, $P = 0.269$).

**Discussion**

Our understanding of pathogen ecology has been greatly influenced by the recognition of cryptic genetic variation, helping to explain disease distributions through species-specific competition, specialization and genetic exchange (de León & Nadler 2010; Cai et al. 2011). There are also important implications for the choices of disease intervention strategies where cryptic pathogen species respond differently to control measures (Bickford et al. 2007). Previously subsumed under a single species name, revisions to the *Microbotryum* genus are ongoing (Le Gac et al. 2007b; Lutz et al. 2008; Denchev et al. 2009; Piątek et al. 2012, 2013), and studies of anther-smut fungi on new host groups or geographic areas often reveal substantial hidden diversity (Kemler et al. 2009). The current study on molecular variation of anther smut on *Dianthus*, combined with studies of the host-of-origin and geographic distributions, contrasts strongly with prior characterizations of anther-smut lineages as highly specialized on a single host and with few hybrids in nature (Le Gac et al. 2007b; Bernasconi et al. 2009; Gladieux et al. 2011; Kemler et al. 2013). Although there was some degree of host specialization, host ranges extended to several *Dianthus* species, and correspondingly any particular host may have several
pathogen lineages. As an important consequence such broad host ranges, sympatry of the differentiated pathogen lineages within host populations was remarkably frequent, providing opportunities for hybridization and, though not studied directly here, intraspecific hybridization.

The level of differentiation among the Microbotryum lineages was informed by several lines of molecular and developmental or ecological evidence. The resolution of four lineages based on microsatellite data was supported by the high ΔK statistic, high proportion of samples assigned to clusters, and cluster separation in the FCA and DAPC analyses. Furthermore, DNA sequence variation at loci previously used for Microbotryum species-level phylogenetics (Le Gac et al. 2007b; Kemler et al. 2013) tended to confirm the clustering of samples as based on these multilocus markers. The phylogenetic resolution was most consistent with the separation of Cluster 3 from the others, and further phylogenetic analyses based on genome-wide multilocus sequence data could aid in clarifying differences among the other lineages and whether additional subdivisions are present.

Considering nonmolecular criteria, the distinction of at least these four identified lineages is supported by their different distributions among host species and the significant sterility of the admixed genotypes, demonstrated by the failure of the latter to grow as haploid meiotic products in the yeast stage. Such sterility suggests substantial divergence in genetic or chromosomal structures that prevent normal meiotic segregation (e.g. Dobzhansky 1933). A range of sterility values was observed within cluster or admixed sample types, and while some such variation might reflect admixed genotypes not being F1 hybrids, the sterility of a few cluster-assigned samples at the 50% level may be due to a phenomenon of mating-type bias that is known to be an infrequent polymorphism causing such haploid viability in Microbotryum species (Thomas et al. 2003). Previous studies have shown that experimental hybrids between Microbotryum species from Silene do suffer particularly from the type of postzygotic isolating mechanism observed here (de Vienne et al. 2009a; Le Gac et al. 2007a; Gibson et al. 2012; Bükker et al. 2013). Microbotryum species often have different chromosome sizes and gametes of F1 hybrids display abnormal DNA content, higher than either parental species (de Vienne et al. 2009a), indicating that meiotic failure may arise from karyotype differences.

Although the sampling in this study was more thorough at the population level than in previous ones, the sampling region was restricted and the number of host species was small relative to the known Dianthus distribution, even within Europe (Valente et al. 2010). Because only some of the DNA sequence variation for Microbotryum from Dianthus was overlapping with that reported by Le Gac et al. (2007b) and Kemler et al. (2013), the concordance of these data sets remains difficult to fully ascertain. It is highly probable that a substantially larger number of Microbotryum lineages may exist on Dianthus hosts elsewhere, and more studies that integrate molecular and ecological characterizations seem essential in advance of taxonomic delimitation to resolve species-level identities of Microbotryum on Dianthus.

The broader host ranges of anther-smut pathogens on Dianthus compared to those on the genus Silene is of particular interest, as noted previously (Le Gac et al. 2007b; Kemler et al. 2013) and quantified here. Dianthus-Silene comparisons are quite old and were also made by Darwin (1859) where he remarked, ‘In the same family there may be a genus, as Dianthus, in which very many species can most readily be crossed; and another genus, as Silene, in which the most persevering efforts have failed to produce between extremely close species a single hybrid’. While we now know that a few closely related Silene species will indeed cross (Krucekberg 1961), the statement reflects a real difference between the genera in the relative timing of radiation, with morphologically diverse European Dianthus species radiating much more recently than most Silene species (Valente et al. 2010; see also Kemler et al. 2013). This more recent divergence among Dianthus species may contribute to the ability of Microbotryum pathogens to shift between them as there may be insufficient physiological differences to favour pathogen specialization. Consistent with this hypothesis, very recently diverged Silene host species, for example S. caroliniana and S. virginica, or S. vulgaris and S. uniflora, also tend to share the same Microbotryum species (Freeman et al. 2002; Piątek et al. 2013). However, two important considerations should also be noted. First, de Vienne et al. (2009b) showed that Dianthus carthusianorum was unusually susceptible to cross-inoculation of Microbotryum pathogens even when they originated from Silene or Saponaria species. Therefore, we cannot exclude the possibility that the lack of specialized affinities of the Microbotryum-Dianthus system may be due to broad susceptibility as a Dianthus ‘host trait’ separate from the recent radiation. Second, among the four Microbotryum lineages studied here, there was one almost entirely restricted to Dianthus pavonius suggesting that it has strong host specialization. Clearly, broader sampling across host species and DNA-based phylogenetic studies should be complemented by experimental inoculation to help resolve the causes of variation in host–pathogen affinities.

A major consequence of lacking one-to-one host–pathogen affinities is that Dianthus species were infected by multiple Microbotryum lineages, very often in sympathy at the local population scale. However, this pathogen sympatry was unlikely to have been due to...
transient host shifts because the co-occurrence of different diseased host species was rare. Instead, the pathogen distributions suggest that *Dianthus* hosts may harbour endemic populations of quite distinct anther-smut fungi; in a few cases, a local host population even contained three *Microbotryum* lineages. One outcome of this sympathy is expected to be direct competition among these pathogen lineages. *Microbotryum* on *Silene* has been well studied as a model of multiple infections of individual hosts (Koskella et al. 2006; López-Villavicencio et al. 2007; Tollenaere et al. 2016), and it has been shown that resident infections tend to prevent colonization by subsequent pathogens. On *Silene* hosts, this within-host exclusion was stronger between *Microbotryum* species than among genotypes within species (Gold et al. 2009). This is consistent with the rarity of co-infection on *Dianthus* observed here given that spores of sympatric pathogen lineages might be deposited together on the host surface prior to infection. However, the principle of competitive exclusion (e.g. Levin 1970) indicates that sympathy in an obligate pathogen like *Microbotryum* would be an unstable condition if it was the only factor affecting coexistence. Therefore, further research should be directed towards the nature of the interactions between co-occurring *Microbotryum* lineages, including the products of hybridization between them, and other factors such as disease transmission and resistance dynamics that might impact their coexistence. This issue is not completely unique to *Microbotryum* on *Dianthus*, as co-occurrence of distinct anther-smut species on the same host species, while very rare in *Silene*, does have a counterpart in *Silene vulgaris*, where three distinct *Microbotryum* species are endemic and specialized to this host (Le Gac et al. 2007b; Piątek et al. 2012).

The second important outcome of co-occurring *Microbotryum* lineages on *Dianthus* hosts was the very high prevalence of admixed pathogen genotypes, assuming the cluster-assignment threshold of 0.9. This group of fungi on *Silene* hosts has a mating system that is predominantly selfing, via automixis (Giraud et al. 2008a), but hybridization can occur and has played an historic role in the past evolution of *Microbotryum* species (Devier et al. 2010). In extant populations, Gladieux et al. (2011) showed that genetic exchange between *Microbotryum* species was predicated on sympathy between two *Silene* host species, each carrying their own specialized pathogen species. Therefore, in general, divergence in the habitat occupied by two host species can contribute to the allopatry of their respective pathogens. The lack of ecological specialization by the pathogen may then be especially important to the disease on *Dianthus*, where the lack of strict affinities of *Microbotryum* lineages to different host species contributes to direct contact, in sympathy on the same host, and the resulting frequent hybridization. Although a failure in this study to detect both parental lineages in some populations may be due to small sample sizes, the absence of populations that consisted of only admixed genotypes, without any cluster-assigned samples is consistent with our finding that hybridization leads to reduced fitness in the form of significant sterility. Because mating occurs on the plant surface prior to infection, and thus prior to competitive exclusion, the mixing of spores during pollinator movement between plants might frequently give rise hybridization even if fertility of the resulting genotypes is low. The extent to which crosses between lineages might persist through early generations or the extent to which they serve as a route for introgression between the pathogen lineages remains unclear and could contribute to the observed range of cluster-assignment scores and to incongruence among data sets based on different loci.

In summary, this study provides insights into the complex distribution of *Microbotryum* pathogens specialized to natural populations of *Dianthus* hosts. Importantly, several *Microbotryum* lineages on *Dianthus* appear to lack narrow and strict host affinities, and sympatric contact between them likely drives both direct competition and hybridization. The full consequences of these processes for pathogen dynamics and evolution remain to be investigated, as these conditions may be quite general. Similar contact between closely related pathogens is becoming recognized in a range of other systems, such as for the avian malaria parasites (Fallon et al. 2005; Ricklefs et al. 2014), suggesting *Microbotryum* might be a tractable model for a broadly occurring phenomenon in disease ecology. Furthermore, the common occurrence of admixed genotypes, contrasting sharply with the general patterns in observed *Microbotryum* on *Silene* hosts (Gladieux et al. 2011), adds to recent calls for investigating the circumstances that promote hybridization in pathogen populations and the study of their potential evolutionary consequences (King et al. 2015). Indeed, although secondary contacts have been extensively studied in animals and plants, they have been little explored in pathogens, where the few recent examples show that they can have especially dramatic consequences on disease emergence (Lemaire et al. 2016; Leroy et al. 2016; Stukenbrock 2016).

**Acknowledgements**

This research was supported by NSF awards DEB-1115765, DEB-0747222, and by Agence Nationale de Recherche MULTI-VIR ANR-13-BSV7-0008. We thank Valentina Carasso, Ivan Pace and Bruno Gallino at the Parco naturale del Marguareis for help with field work and Felix Horns, Daniel Peterson, Molly Scott and Paul Tyler for research assistance.


M.E.H., E.P. and C.S. conceived and designed the project with contributions of J.A. and T.G. C.S., S.Y., M.E.H. and J.A. performed field collections. C.S., E.P., L.R., A.C., P.G., E.B., S.Y., T.G. and M.E.H. contributed to molecular analyses and interpretations. M.E.H., E.P. and C.S. drafted the manuscript with contributions by all authors.

Data accessibility

Supporting information
Additional supporting information may be found in the online version of this article.

Fig. S1 Distributions across south-western European Alps of the Microbotryum pathogens and Dianthus host populations.

Fig. S2 Measures used in estimating the number of genetic clusters among Microbotryum samples from Dianthus hosts from STRUCTURE analyses using the ΔK and the proportions of assigned samples at the 0.9 confidence threshold for cluster assignment.

Fig. S3 Population structure in Microbotryum (N = 443, 81 sites across the Alps), inferred with the Bayesian clustering algorithm implemented in STRUCTURE.

Fig. S4 Barplot graphs for the ten replicate runs at K = 4.

Fig. S5 Factorial correspondence analyses (FCA) of Microbotryum samples from Dianthus hosts.

Fig. S6 Discriminant analysis of principal component (DAPC) measures used in estimating the number of genetic clusters among Microbotryum samples from Dianthus hosts.

Fig. S7 Sterility of admixed Microbotryum samples (with <0.9 confidence score for cluster assignment) compared to samples assigned to each of four genetic clusters.

Fig. S8 Distribution of cluster-assigned samples of Microbotryum among six host species of Dianthus.

Fig. S9 Distribution and abundance of cluster-assigned and admixed Microbotryum samples, using <0.9 confidence score for cluster assignment.

Table S1 Population details for 2008 sampling Microbotryum from Dianthus hosts. Valleys represented by three-letter code.

Table S2 PCR primer sequences and annealing temperatures for newly developed microsatellite loci in Microbotryum from Dianthus hosts.

Table S3 Information on Microbotryum samples included in the phylogenetic analysis of DNA sequence data.

Table S4 Summary statistics of microsatellite loci used to analyze Microbotryum samples from Dianthus hosts.

Table S5 P-values from Wilcoxon signed-rank tests for among cluster comparisons of allelic richness and private allele richness using Wilcoxon signed-rank tests.

© 2017 John Wiley & Sons Ltd