



HAL
open science

Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts

Elsa Petit, Casey Silver, Amandine Cornille, Pierre Gladieux, Lisa Rosenthal, Emily Bruns, Sarah Yee, Janis Antonovics, Tatiana Giraud, Michael E. Hood

► **To cite this version:**

Elsa Petit, Casey Silver, Amandine Cornille, Pierre Gladieux, Lisa Rosenthal, et al.. Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. *Molecular Ecology*, 2017, 26 (7), pp.1877-1890. 10.1111/mec.14073 . hal-01605903

HAL Id: hal-01605903

<https://hal.science/hal-01605903v1>

Submitted on 25 May 2020

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

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

DR. TATIANA GIRAUD (Orcid ID : 0000-0002-2685-6478)

DR. MICHAEL HOOD (Orcid ID : 0000-0003-4379-0711)

Received Date : 19-Oct-2016

Revised Date : 13-Jan-2017

Accepted Date : 09-Feb-2017

Article type : Special Issue

Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts

Elsa Petit^{1*}, Casey Silver^{1†}, Amandine Cornille[‡], Pierre Gladieux[§], Lisa Rosenthal[†], Emily Bruns[¶], Sarah Yee[†], Janis Antonovics[¶], Tatiana Giraud^{**}, Michael E. Hood[†]

* Stockbridge School of Agriculture, University of Massachusetts, Amherst, MA, USA, [†] Biology Department, Amherst College, Amherst, MA 01002, USA, [‡] Center for Adaptation to a Changing Environment, ETH Zürich, 8092 Zürich, Switzerland, [§] UMR BGPI, INRA, Campus International de Baillarguet, 34398 Montpellier, France, [¶] Department of Biology, University of Virginia, Charlottesville, Virginia 22903, ^{**} Ecologie Systematique Evolution, CNRS, Univ. Paris-Sud, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

¹ Contributed equally to the work.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14073

This article is protected by copyright. All rights reserved.

Comment citer ce document :

Petit, E., Silver, C., Cornille, A., Gladieux, P., Rosenthal, L., Bruns, E., Yee, S., Antonovics, J., Giraud, T., Hood, M. E. (Auteur de correspondance) (2017). Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. *Molecular Ecology*, 26 (7), 1877-1890. . DOI : 10.1111/mec.14073

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

Corresponding author: Michael E. Hood, Biology Department, Amherst College, Amherst, MA 01002, USA, 413-542-2097, mhood@amherst.edu

Running title: Anther-smut pathogens specialized on *Dianthus*

Abstract

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.

This article is protected by copyright. All rights reserved.

Petit, E., Silver, C., Cornille, A., Gladieux, P., Rosenthal, L., Bruns, E., Yee, S., Antonovics, J., Giraud, T., Hood, M. E. (Auteur de correspondance) (2017). Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. *Molecular Ecology*, 26 (7), 1877-1890. . DOI : 10.1111/mec.14073

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 *et al.* 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 favor the emergence of new disease properties, such as increased virulence or with new host affinities (Kuldau *et al.* 1999; Brasier 2000; Newcombe *et al.* 2000; Lin *et al.* 2007; 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

This article is protected by copyright. All rights reserved.

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; Feurtey *et al.* 2016), speciation (Le Gac *et al.* 2007a, 2007b; Refrégier *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 (Refrégier *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.

This article is protected by copyright. All rights reserved.

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: 1) identify genetically differentiated lineages of *Microbotryum* genotypes, and in particular incorporating population-level sampling, 2) assess sterility of genotypes resulting from hybridization that may indicate reproductive isolation among the pathogen lineages, 3) determine pathogen distributions at different spatial scales in terms of host specialization and sympatry, and 4) 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 multi-host multi-pathogen 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 harbor multiple genotypes of the pathogen, but there is frequently competitive exclusion, which is stronger

This article is protected by copyright. All rights reserved.

between genetically more distant lineages (Hood 2003; López-Villavicencio *et al.* 2007; López-Villavicencio *et al.* 2011; Gold *et al.* 2009). 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.* 2013).

This article is protected by copyright. All rights reserved.

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 - 44.5N and longitudes 7.0 - 8.2E (Fig. 2, S1). 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). 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 100m; 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 (std. dev. = 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 if 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) 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).

This article is protected by copyright. All rights reserved.

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). 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. 3x 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-labeled PCR primers. PCR was performed in a final volume of 10 μ L containing 1 \times Phusion HF Buffer (New England Biolabs), 0.2 mmol/L deoxynucleoside-5'-triphosphate (Promega), 0.25 μ mol/L forward and reverse primer, 1 U/ μ L Phusion HF DNA polymerase (New England Biolabs), and 0.8 μ L of template DNA. Separate PCR reactions were started with an initialization step for 30 seconds at 98 $^{\circ}$ C, 50 cycles of 10 seconds at 98 $^{\circ}$ C, 30 seconds at the annealing temperature (Table S2), 30 seconds at 72 $^{\circ}$ C, and 1 final extension cycle of 5 minutes at 72 $^{\circ}$ C. PCR reactions with different primer dyes (Table S2) were multiplexed and a 600 LIZ dye size standard was added. Fragment analysis were run at the UMass Genomics Resources Lab on Applied Biosystems 3130 and 3130xl 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.

This article is protected by copyright. All rights reserved.

Population genetic subdivision and assessment of hybridization.

Individual-based Bayesian clustering methods, implemented in STRUCTURE 2.3.3 (Pritchard *et al.* 2000), 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 ($2 \leq K \leq 9$), 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), centered 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 1,000 random permutations was performed between the individual

This article is protected by copyright. All rights reserved.

coefficients of relatedness F_{ij} (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 3 of 81 populations or 4 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 (arcsine 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 (A_R) and private allele richness (A_P) calculated with ADZE (Szpiech *et al.* 2008). Allelic richness and private allele richness

This article is protected by copyright. All rights reserved.

were computed using sample sizes of $N=10$ (five individuals x 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 to 30 hours. The four post-meiotic 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 hours 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 post-meiotic cells was compared to non-admixed individuals using the t-test for

This article is protected by copyright. All rights reserved.

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: 1) the internal transcribed spacer (ITS) region of the multi-copy ribosomal RNA gene as used by Kemler *et al.* (2013), and 2) elongation factor 1 α (Ef1 α) 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 Ef1 α 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). 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 multi-copy nature of ITS loci, admixed pathogens between lineages were identified by intra-genomic sequence polymorphisms at particular nucleotide positions in DNA sequencing electrophoregrams (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. Our conclusion on the most relevant number of clusters was based upon a combination of metrics. The peak of ΔK value was found at $K = 3$, and it remained high for $K = 4$ and 5 (Fig. S2). 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 to 0.72, but then dropped markedly to 0.59 and 0.51 for $K = 5$ to 7 (Figs. S2 and S3). $K = 4$ was thus the highest K value defining well-delimited clusters, representing the finest strong genetic subdivision in the dataset, 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), while at $K = 5$ the fifth clusters was 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). 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). Among the 318 cluster-assigned *Microbotryum* samples, 143 samples were assigned to Cluster 1 (Fig 2A, red color), 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

This article is protected by copyright. All rights reserved.

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). Levels of divergence in allelic frequencies between the clusters (F_{ST}) 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 post-meiotic cells compared to non-admixed 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).

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 *M. 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 dataset (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

This article is protected by copyright. All rights reserved.

also delimit these two clusters well. Cluster 4 was more dispersed among Ef1 α variants, while it was well distinguished by ITS. In contrast, Clusters 1 and 2 were less well delimited by ITS than Ef1 α .

Distributions and host affinities of Microbotryum lineages

The four pathogen lineages, identified as differentiated clusters above, were non-randomly distributed with 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, geographic 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 identified as Cluster 4 contributed most strongly to this spatial 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 geographical distance between them approached significance (Mantel test, $r = 0.260$, $P < 0.053$ for 19 valleys) but this measure was not significant for the other pathogen lineages (smallest P-value among other clusters = 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 non-random distribution of the four pathogen lineages among host species was also evident across hierarchical sampling scales, whether analyzing individual 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; Fisher's Exact Test, Monte-Carlo 10,000 iterations, $P < 0.001$ for populations and $P = 0.009$ for valleys). In particular, Cluster 1 was almost exclusively associated with *Dianthus pavonius*; only one out of 143 samples was isolated from another host, *D. deltooides* (Fig. 5). Cluster 3 also was most often associated with *D. pavonius*, though it was also found on *D. furcatus*. The pathogen of Cluster 4 was most frequently associated with *D. seguieri* although also

This article is protected by copyright. All rights reserved.

with three other hosts. The pathogen belonging to Cluster 2 was the most widespread, often on *D. pavonius* while also on four other hosts. The distribution of admixed pathogen samples among hosts is shown in Figure S9A.

Dianthus host species were thus associated with multiple *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 extend down to the level of co-infecting individual host plants. Within three valleys where pathogen sympatry occurred (STA, SBM, and PRA; Table S1) 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.

Sympatry of different diseased host species was markedly less common than sympatry 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, sympatry of diseased host species was not a significant predictor of sympatry for pathogen lineages; only one of the three host-sympatric populations was a location of pathogen sympatry (binomial probability of at least one of three populations having pathogen sympatry, $P = 0.657$), and while each of four valleys with diseased host sympatries was also a location of pathogen sympatry 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).

This article is protected by copyright. All rights reserved.

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 pathogens 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), 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) 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 (Cai *et al.* 2011, de León & Nadler 2010). 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 (Kemper *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

This article is protected by copyright. All rights reserved.

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, inter-lineage pathogen competition.

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.* 2007; Kemler *et al.* 2013) tended to confirm the clustering of samples as based on these multi-locus markers. The phylogenetic resolution was most consistent with the separation of Cluster 3 from the others, and further phylogenetic analyses based on genome-wide multi-locus sequence data could aid in clarifying differences among the other lineages and whether additional subdivisions are present.

Considering non-molecular 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 genic 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 post-zygotic isolating mechanism observed here (de Vienne *et al.*

This article is protected by copyright. All rights reserved.

2009a; Le Gac *et al.* 2007a; Gibson *et al.* 2012; B ker *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 datasets 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 (1875) 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 (Kruckeberg 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 favor pathogen specialization. Consistent with this hypothesis, very recently diverged *Silene* host species, e.g. *S. caroliniana* and *S. virginica*, or *S. vulgaris* and *S. uniflora*, also tend to share the same *Microbotryum* species (Freeman *et al.* 2002;

This article is protected by copyright. All rights reserved.

Piątek *et al.* 2013). However, two important considerations should also be noted. First, de Vienne *et al.* (2009b) showed that *D. 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 *D. 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 sympatry 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 harbor 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 sympatry 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 sympatry 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 toward 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

This article is protected by copyright. All rights reserved.

impact their co-existence. 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 *S. 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 sympatry 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 sympatry 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 datasets based on different loci.

This article is protected by copyright. All rights reserved.

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 the 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).

Acknowledgments

This research was supported by NSF awards DEB-1115765, DEB-0747222. 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.

This article is protected by copyright. All rights reserved.

Author contributions

MEH, EP and CS conceived and designed the project with contributions of JA and TG. CS, SY, MEH and JA performed field collections. CS, EP, LR, AC, PG, EB, SY, TG and MEH contributed to molecular analyses and interpretations. MEH, EP and CS drafted the manuscript with contributions by all authors.

Data accessibility

Online Supporting Information: Sampling locations, Supplemental Tables. Dryad: Microsatellite data, Structure input files doi:10.5061/dryad.48qd5. GenBank: DNA sequences accessions KY451038- KY451218.

References

- Alizon S, de Roode JC, Michalakakis Y (2013) Multiple infections and the evolution of virulence. *Ecology Letters* **16**, 556-567.
- Antonovics J, Hood M, Partain J (2002) The ecology and genetics of a host shift: *Microbotryum* as a model system. *The American Naturalist* **160**, S40–S53.
- Bagchi R, Gallery RE, Gripenberg S, Gurr SJ, Narayan L, Addis CE, Freckleton RP, Lewis OT (2014) Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* **506**, 85-88.
- Badouin H, Gladioux P, Gouzy J, Siguenza S, Aguilera G, Snirc A, Le Prieur S, Jeziorski C, Branca A and Giraud T. (2017) Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. *Molecular Ecology* In press
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire génome, populations, interactions, CNRS UMR, 5000, pp.1996-2004.
- Benítez MS, Hersh MH, Vilgalys R, Clark JS (2013) Pathogen regulation of plant diversity via effective specialization. *Trends in Ecology & Evolution* **28**, 705-711.

This article is protected by copyright. All rights reserved.

- Bernasconi G, Antonovics J, Biere A, Charlesworth D, Delph L F, Filatov D, Giraud T, Hood ME, Marais GAB, McCauley D, Pannell J, Shykoff JA, Vyskot B, Wolfe L, Widmer A (2009) *Silene* as a model system in ecology and evolution. *Heredity*. **103**, 5–14.
- Bever JD, Mangan SA, Alexander HM (2015) Maintenance of plant species diversity by pathogens. *Annual Review of Ecology, Evolution, and Systematics* **46**, 305-325.
- Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22**, 148-155.
- Bolton S, Bon C, (2009) *Pharmaceutical statistics: practical and clinical applications*. CRC Press.
- Brasier C (2000) The rise of the hybrid fungi. *Nature* **405**, 134–135.
- Bucheli E, Gautschi B, Shykoff JA (2000) Host-specific differentiation in the anther smut fungus *Microbotryum violaceum* as revealed by microsatellites. *Journal of Evolutionary Biology* **13**, 188-198.
- Bucheli E, Gautschi B, Shykoff J (2001) Differences in population structure of the anther smut fungus *Microbotryum violaceum* on two closely related host species, *Silene latifolia* and *S. dioica*. *Molecular Ecology* **10**, 285–294.
- Bueker B, Eberlein C, Gladieux P, Schaefer A, Snirc A, Bennett DJ, Begerow D, Hood ME, Giraud T (2016) Distribution and population structure of the anther smut *Microbotryum silenes-acaulis* parasitizing an arctic–alpine plant. *Molecular Ecology*. **25**, 811-24.
- Bull JJ, Luring AS, (2014) Theory and empiricism in virulence evolution. *PLoS Pathogen* **10**, e1004387.
- Cai L, Giraud T, Zhang N, Begerow D, Cai G, Shivas RG (2011) The evolution of species concepts and species recognition criteria in plant pathogenic fungi. *Fungal Diversity* **50**, 121-133.
- Darwin C (1859) *On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life*. Murray, London, England, UK
- Day AW (1980) Competition and distribution studies of genetically marked strains of *Ustilago violacea* in the same host plant. *Botanical Gazette* **141**, 313-320.
- de León GPP, Nadler SA (2010). What we don't recognize can hurt us: a plea for awareness about cryptic species. *Journal of Parasitology* **96**, 453-464.
- de Vienne D M, Refrégier G, Hood M E, Guigue A, Devier B, Vercken E, Smadja C Deseille A, Giraud T (2009a) Hybrid sterility and inviability in the parasitic fungal species complex *Microbotryum*. *Journal of Evolutionary Biology* **22**, 683-698.
- de Vienne D M, Hood M E, Giraud T (2009b) Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology* **22**, 2532-2541.
- Denchev CM, Giraud T, Hood ME (2009) Three new species of anthericolous smut fungi on Caryophyllaceae. *Mycologica Balcanica* **6**, 79-84.

This article is protected by copyright. All rights reserved.

- Desprez-Loustau ML, Robin C, Buee M, Courtecuisse R, Garbaye J, Suffert F, Sache I, Rizzo DM (2007) The fungal dimension of biological invasions. *Trends in Ecology and Evolution*. **30**, 472-80.
- Devier B, Aguilera G, Hood ME, Giraud T (2010) Using phylogenies of pheromone receptor genes in the *Microbotryum violaceum* species complex to investigate possible events of speciation by hybridization. *Mycologia* **102**, 689-696.
- Dobzhansky T (1933) Role of the autosomes in the *Drosophila pseudoobscura* hybrids. *Proceedings of the National Academy of Sciences* **19**, 950-953.
- Dobson A, Lafferty KD, Kuris AM, Hechinger RF, Jetz W (2008) Homage to Linnaeus: How many parasites? How many hosts? *Proceedings of the National Academy of Sciences* **105**, S11482-S11489.
- Fallon SM, Bermingham E, Ricklefs RE (2005) Host specialization and geographic localization of avian malaria parasites: A regional analysis in the Lesser Antilles. *The American Naturalist* **165**, 466-480.
- Feurtey A, Gladieux P, Snirc A, Hood ME, Cornille A, Giraud T (2016) Strong phylogeographic costructure between the anther smut fungus and its white campion host. *New Phytologist* (in press, DOI: 10.1111/nph.14191).
- Fontaine MC, Gladieux P, Hood ME, Giraud T (2013) History of the invasion of the anther smut pathogen on *Silene latifolia* in North America. *New Phytologist* **198**, 946-56.
- Freeman AB, Duong KK, Shi TL, Hughes CF, Perlin MH, (2002) Isolates of *Microbotryum violaceum* from North American host species are phylogenetically distinct from their European host-derived counterparts. *Molecular Phylogenetics and Evolution* **23**, 158-170.
- Giraud T (2004) Patterns of within population dispersal and mating of the fungus *Microbotryum violaceum* parasitizing the plant *Silene latifolia*. *Heredity* **93**, 559-65.
- Giraud T, Fournier E, Vautrin D, Solignac M, Shykoff JA (2002) Isolation of 44 polymorphic microsatellite loci in three host races of the phytopathogenic fungus *Microbotryum violaceum*. *Molecular Ecology Notes* **2**, 142-146.
- Giraud T, Villaréal L, Austerlitz F, Le Gac M, Lavigne C (2006) Importance of the life cycle in sympatric host race formation and speciation of pathogens. *Phytopathology* **96**, 280-287..
- Giraud T, Yockteng R, López-Villavicencio M, Refrégier G, Hood M (2008a) Mating system of the anther smut fungus *Microbotryum violaceum*: selfing under heterothallism. *Eukaryotic Cell* **7**, 765-775.
- Giraud T, Yockteng R, Marthey S, Chiapello H, Jonot O, Lopez-Villavicencio M, De Vienne DM, Hood ME, Refregier G, Gendraud-Jacquemard A, Wincker P (2008b) Permanent genetic resources: isolation of 60 polymorphic microsatellite loci in EST libraries of four sibling species of the phytopathogenic fungus complex *Microbotryum*. *Molecular Ecology Resources* **8**, 387-392.

This article is protected by copyright. All rights reserved.

- Gladieux P, Vercken E, Fontaine MC, Hood ME, Jonot O, Couloux A, Giraud T (2011) Maintenance of fungal pathogen species that are specialized to different hosts: allopatric divergence and introgression through secondary contact. *Molecular Biology and Evolution* **28**, 459–471.
- Gladieux P, Feurtey A, Hood ME, Snirc A, Clavel J, Dutech C, Roy M, Giraud T (2015) The population biology of fungal invasions. *Molecular Ecology* **24**, 1969–1986.
- Gold A, Giraud T, Hood ME (2009) Within-host competitive exclusion among species of the anther-smut pathogen. *BMC Ecology* **9**, 11
- Goldschmidt V (1928) Vererbungsversuche mit den biologischen Arten des Antherenbrandes (*Ustilago violacea* Pers.). *Zeitschrift fuer Botanik* **21**, 1–90.
- Hood ME, Antonovics J (2000) Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity* **85**, 231–241.
- Hood ME (2003) Dynamics of multiple infection and within-host competition by the anther-smut pathogen. *The American Naturalist* **162**, 122–133.
- Hood ME, Mena-Alí JI, Gibson AK, Oxelman B, Giraud T, Yockyeng R, Arroyo MTK, Conti F, Pedersen A, Gladieux P, Antonovics J (2010) Distribution of the anther-smut pathogen *Microbotryum* on species of the Caryophyllaceae. *New Phytologist* **187**, 217–229.
- James SA, O'Kelly MJ, Carter DM, Davey RP, van Oudenaarden A, Roberts IN (2009) Repetitive sequence variation and dynamics in the ribosomal DNA array of *Saccharomyces cerevisiae* as revealed by whole-genome resequencing. *Genome Research* **19**, 626–635.
- Kemler M, Lutz M, Göker M, Oberwinkler F, Begerow D (2009) Hidden diversity in the non-caryophyllaceous plant-parasitic members of *Microbotryum* (Pucciniomycotina: Microbotryales). *Systematics and Biodiversity* **7**, 297–306.
- Kemler M, Martín MP, Telleria MT, Schäfer AM, Yurkov A, Begerow D (2013) Contrasting phylogenetic patterns of anther smuts (Pucciniomycotina: *Microbotryum*) reflect phylogenetic patterns of their caryophyllaceous hosts. *Organisms Diversity and Evolution* **13**, 111–126.
- King KC, Stelkens RB, Webster JP, Smith DF, Brockhurst MA (2015) Hybridization in parasites: consequences for adaptive evolution, pathogenesis, and public health in a changing world. *PLoS Pathogen* **11**, e1005098.
- Koskella B, Giraud T, Hood ME (2006) Pathogen relatedness affects the prevalence of within-host competition. *The American Naturalist* **168**, 121–126
- Kruckeberg AR (1961) Artificial crosses of Western North American *Silenes*. *Brittonia* **13**, 305–333.
- Le Gac M, Hood ME, Giraud T (2007a) Evolution of reproductive isolation within a parasitic fungal species complex. *Evolution* **61**, 1781–1787.
- Le Gac M, Hood M, Fournier E, Giraud T (2007b) Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. *Evolution* **61**, 15–26.

This article is protected by copyright. All rights reserved.

- Lemaire C, De Gracia M, Leroy T, Michalecka M, Lindhard-Pedersen H, Guerin F, Gladieux P, Le Cam B (2016) Emergence of new virulent populations of apple scab from nonagricultural disease reservoirs. *New Phytologist* **209**, 1220–1229.
- Leroy T, Caffier V, Celton JM, Anger N, Durel CE, Lemaire C, Le Cam B (2016). When virulence originates from nonagricultural hosts: evolutionary and epidemiological consequences of introgressions following secondary contacts in *Venturia inaequalis*. *New Phytologist* **210**, 1443-1452.
- Levin SA (1970) Community equilibria and stability, and an extension of the competitive exclusion principle. *The American Naturalist* **104**, 413–423.
- López-Villavicencio M, Jonot O, Coantic A, Hood ME, Enjalbert J, Giraud T (2007) Multiple infections by the anther smut pathogen are frequent and involve related strains. *PLoS Pathogen* **3**, e176.
- López-Villavicencio M, Courjol F, Gibson AK, Hood ME, Jonot O, Shykoff JA, Giraud T. (2011) Competition, cooperation among kin, and virulence in multiple infections. *Evolution* **65**, 1357-66
- Lutz M, Piątek M, Kemler M, Chlebicki A, Oberwinkler F (2008) Anther smuts of Caryophyllaceae: molecular analyses reveal further new species. *Mycological Research*. **112**, 1280-96.
- Menardo F, Praz CR, Wyder S, Ben-David R, Bourras S, Matsumae H, McNally KE, Parlange F, Riba A, Roffler S, Schaefer LK (2016) Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. *Nature Genetics* **48**, 201-5.
- Newcombe G, Stirling B, McDonald S, Bradshaw HD (2000) *Melampsorax columbiana*, a natural hybrid of *M. medusae* and *M. occidentalis*. *Mycological Research* **104**, 261-74.
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.
- Piątek M, Lutz M, Ronikier A, Kemler M, Świdarska-Burek U (2012) *Microbotryum heliospermae*, a new anther smut fungus parasitic on *Heliosperma pusillum* in the mountains of the European Alpine System. *Fungal Biology* **116**, 185-195.
- Piątek M, Lutz M, Kemler M, (2013) *Microbotryum silenes-saxifragae* sp. nov. sporulating in the anthers of *Silene saxifraga* in southern European mountains. *IMA fungus* **4**, 29-40.
- Refrégier G, Le Gac M, Jabbour F, Widmer A, Shykoff JA, Yockteng R, Hood ME, Giraud T (2008) Cophylogeny of the anther smut fungi and their caryophyllaceous hosts: Prevalence of host shifts and importance of delimiting parasite species for inferring cospeciation. *BMC Evolutionary Biology* **8**, 100.
- Refrégier G, Hood M, Giraud T (2010) No Evidence of Reproductive Character displacement between two sister fungal species causing anther smut disease in *Silene*. *International Journal of Plant Sciences* **171**, 847–859.
- Rice W (1987) Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evolutionary Ecology* **1**, 301–314.

This article is protected by copyright. All rights reserved.

- Ricklefs RE, Outlaw D C, Svensson-Coelho M, Medeiros M C, Ellis VA, Latta S (2014) Species formation by host shifting in avian malaria parasites. *Proceedings of the National Academy of Sciences* **111**, 14816-14821.
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* **132**, 365–386.
- Stukenbrock EH (2013) Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. *New Phytologist* **199**, 895-907.
- Stukenbrock EH (2016) The role of hybridization in the evolution and emergence of new fungal plant pathogens. *Phytopathology* **106**, 104-12.
- Schardl CL, Craven KD, (2003) Interspecific hybridization in plant-associated fungi and oomycetes: a review. *Molecular Ecology* **12**, 2861-2873.
- Thomas A, Shykoff J, Jonot O, Giraud T (2003) Sex-ratio bias in populations of the phytopathogenic fungus *Microbotryum violaceum* from several host species. *International Journal of Plant Sciences* **164**, 641-647.
- Thrall PH, Jarosz AM (1994) Host-pathogen dynamics in experimental populations of *Silene alba* and *Ustilago violacea*. I. Ecological and genetic determinants of disease spread. *Journal of Ecology*. **82**, 549-559
- Tollenaere C, Susi H, Laine AL, (2016) Evolutionary and epidemiological implications of multiple infection in plants. *Trends in Plant Science* **21**, 80-90.
- Valente L, Savolainen V, Vargas P (2010) Unparalleled rates of species diversification in Europe. *Proceedings of the Royal Society, B* **277**, 1489–1496.
- de Vienne D M, Refrégier G, Hood ME, Guigue A, Devier B, Vercken E, Smadja C Deseille A, Giraud T (2009) Hybrid sterility and inviability in the parasitic fungal species complex *Microbotryum*. *Journal of Evolutionary Biology* **22**, 683–698
- Zillig H (1921) Ueber spezialisierte Formen beim Antherenbrand, *Ustilago violacea* (Pers.) Fuckel. *Zentralblatt fuer Bakteriologie und Parasitenkunde* **53**, 33–73.

Figure legends

Fig 1. *Dianthus pavonius* illustrating symptoms of the anther-smut disease (left) and a healthy flower (right). *Microbotryum* species replace pollen of diseased plants with dark, powdery fungal spores, which can be carried by pollinators to healthy hosts.

Fig 2. Distributions across south-western 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 non-overlapping representation of all pie graphs. A) Population structure of 443 pathogen samples was inferred with the Bayesian clustering algorithm implemented in STRUCTURE,

This article is protected by copyright. All rights reserved.

where the four genetic clusters are indicated by colored 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..

- Fig. 3. Sterility of admixed *Microbotryum* samples (i.e. with <0.9 confidence score for cluster assignment) compared to samples assigned to each of four genetic clusters. The putative parental clusters of the admixed genotypes were determined as the two clusters to which membership coefficients were greatest.
- Fig. 4. DNA sequence-based phylogenies of *Microbotryum* samples from *Dianthus* hosts for comparison with prior studies. Four genetic clusters of *Microbotryum* samples are indicated by colored bars. Approximately 100, cluster-assigned samples were included in both phylogenies. A) Phylogeny of elongation factor 1 α (Ef1 α), with the placement of symbols immediately below the taxa with DNA sequences identical to *Microbotryum* species previously identified in Le Gac *et al.* (2007); namely *M. carthusianorum* NCBI accession DQ074540, *M. shykoffianum* accession DQ074549, and *M. superbum* accession DQ074547. B) Phylogeny of internal transcribed spacer (ITS) region of the multi-copy ribosomal RNA gene, with the symbol referring to a sequence identical to *Microbotryum* clade previously identified in Kemler *et al.* (2013); namely Group 3 accession JQ307901. The numbers of samples differ between the two trees due to a few failures of PCR or sequencing reactions.
- Fig. 5. Distribution of cluster-assigned samples of *Microbotryum* among six host species of *Dianthus*. Four genetic clusters of *Microbotryum* samples are indicated by colored bars.
- Fig. S1. Distributions across south-western European Alps of the *Microbotryum* pathogens and *Dianthus* host populations. Points represent populations shown as pie charts of multiple samples in Fig. 2. Valley names are indicated next to the populations.
- Fig. S2. Measures used in estimating the number of genetic clusters among *Microbotryum* samples from *Dianthus* hosts from INSTRUMENT 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. Each individual is represented by a vertical bar, partitioned into K segments representing proportions of the individual's estimated admixture fractions in the putative number of K clusters, ranging from 2 to 7. The samples were ordered by the majority assignment values for each cluster at $K=4$. Names of clusters across K values are arbitrary.
- Fig. S4. Barplot graphs for the ten replicate runs at $K = 4$. Each individual is represented by a vertical bar, partitioned into K segments representing proportions of the individual's estimated admixture fractions in the putative number of $K = 4$ clusters. The samples were ordered as in Fig. S3.
- Fig. S5 Factorial correspondence analyses (FCA) of *Microbotryum* samples from *Dianthus* hosts. The analysis includes cluster-assigned samples. Four genetic clusters are indicated by colored circles. For perspective in the plot, markers are size-adjusted and approximate drop shadows are shown. Percent contributions of each axis are shown in parentheses.

This article is protected by copyright. All rights reserved.

Fig. S6. Discriminant analysis of principal component (DAPC) on the *Microbotryum* from *Dianthus* hosts, plotted on the first two linear discriminants. Dot colors correspond to the cluster-assigned samples, and ellipses are inertia limits calculated under default setting as representing 67% of the variance on both pc-axes. Variance percentages captured by the first and second axes are indicated on each axis. Principal components retained (black highlight) and their cumulative variance explained are in the inset graph.

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. The percent of viable among 20 meiotic products per sample are shown plotted against the cluster assignment score. The putative parental clusters of the admixed genotypes were determined as the two clusters to which membership coefficients were greatest. Four genetic clusters and admixed samples are indicated by colored circles.

Fig. S8. Distribution of cluster-assigned samples of *Microbotryum* among six host species of *Dianthus*. The plots are the same as in Fig 3, but representing units of observations as A) populations, or B) valleys. Four genetic clusters of *Microbotryum* samples are indicated by colored bars. Host names are abbreviated.

Fig. S9. Distribution and abundance of cluster-assigned and admixed *Microbotryum* samples, using <0.9 confidence score for cluster assignment. A) Numbers of samples are shown for each host species; host names are abbreviated. B) Number of observed admixed *Microbotryum* samples for each pairwise type, and the numbers expected based upon the relative abundances of samples assigned to the four genetic clusters. The putative parental clusters were determined as the two clusters to which membership coefficients of the admixed genotypes were greatest.

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

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

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; color descriptors correspond to their use to indicate clusters in the figures. *All *F_{IS}* values were significant at *P* < 0.05.

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

	Clusters			
	1 (red)	2 (blue)	3 (green)	4 (yellow)
1 (red)				
2 (blue)	0.276			
3 (green)	0.399	0.262		
4 (yellow)	0.325	0.268	0.362	

Admixed genotypes were not included in this analysis and thus analyses were conducted on a total of N = 318 individuals; color descriptors correspond to their use to indicate clusters in the figures.

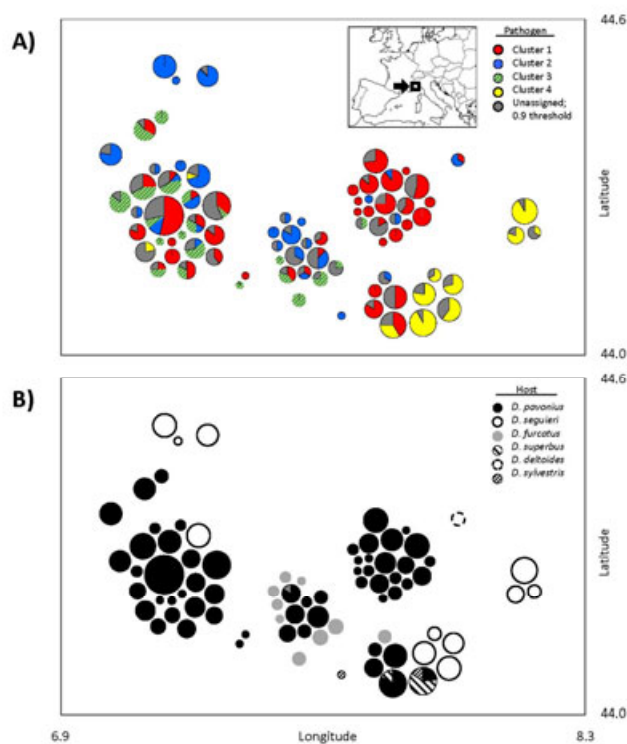
Fig. 1.



This article is protected by copyright. All rights reserved.

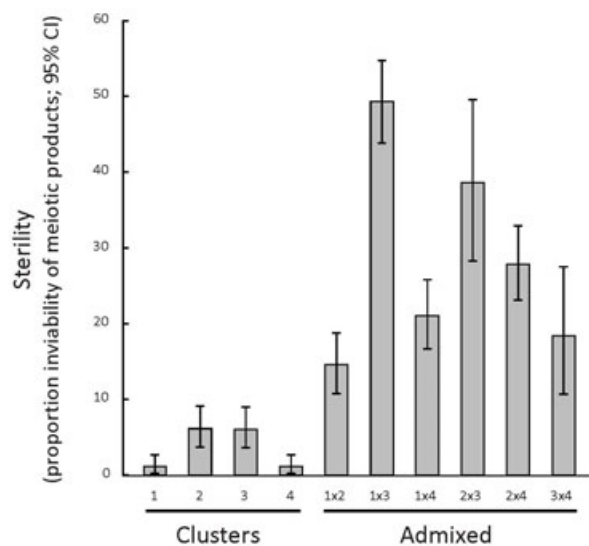
Petit, E., Silver, C., Cornille, A., Gladioux, P., Rosenthal, L., Bruns, E., Yee, S., Antonovics, J., Giraud, T., Hood, M. E. (Auteur de correspondance) (2017). Co-occurrence and hybridization of anther-smut pathogens specialized on Dianthus hosts. *Molecular Ecology*, 26 (7), 1877-1890. . DOI : 10.1111/mec.14073

Fig. 2.



This article is protected by copyright. All rights reserved.

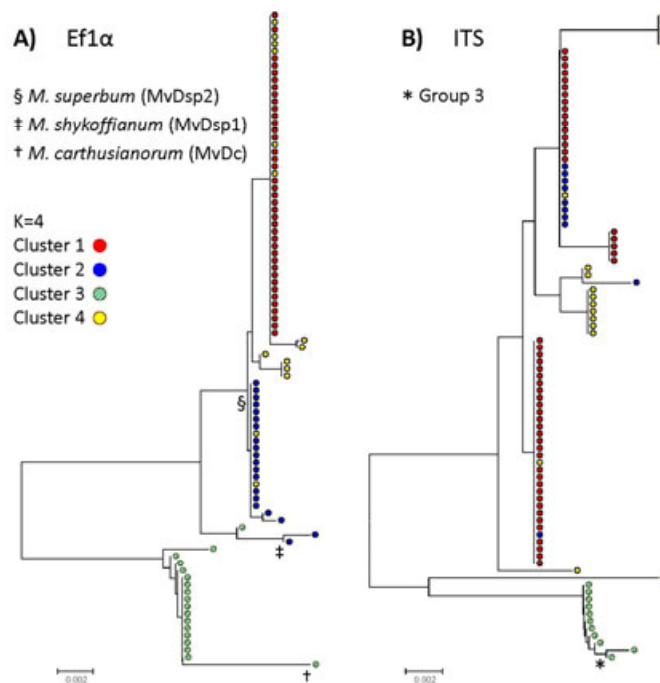
Fig. 3.



This article is protected by copyright. All rights reserved.

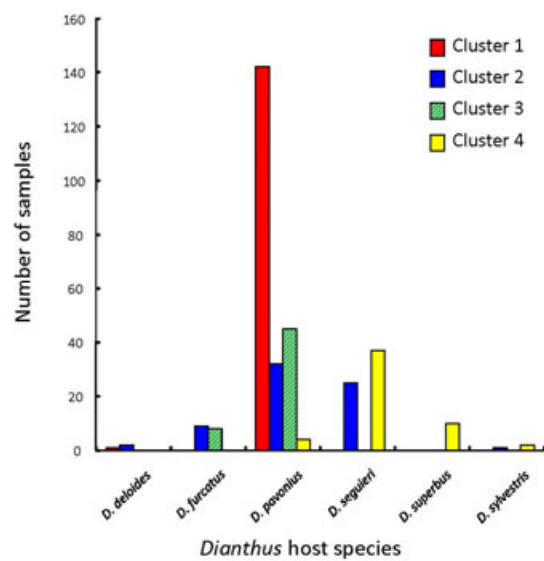
Petit, E., Silver, C., Cornille, A., Gladioux, P., Rosenthal, L., Bruns, E., Yee, S., Antonovics, J., Giraud, T., Hood, M. E. (Auteur de correspondance) (2017). Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. *Molecular Ecology*, 26 (7), 1877-1890. DOI : 10.1111/mec.14073

Fig. 4.



This article is protected by copyright. All rights reserved.

Fig. 5.



This article is protected by copyright. All rights reserved.

Petit, E., Silver, C., Cornille, A., Gladioux, P., Rosenthal, L., Bruns, E., Yee, S., Antonovics, J., Giraud, T., Hood, M. E. (Auteur de correspondance) (2017). Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. *Molecular Ecology*, 26 (7), 1877-1890. DOI : 10.1111/mec.14073