HAL
open science

Co-occurrence among three divergent plant-castrating fungi in the same Silene host species

Jessica Lee Abbate, Pierre Gladieux, Michael E Hood, Damien de Vienne, Janis Antonovics, Alodie Snirc, Tatiana Giraud

To cite this version:

Jessica Lee Abbate, Pierre Gladieux, Michael E Hood, Damien de Vienne, Janis Antonovics, et al.. Co-occurrence among three divergent plant-castrating fungi in the same Silene host species. Molecular Ecology, 2018, 27 (16), pp.3357 - 3370. 10.1111/mec.14805 . hal-01921830

HAL Id: hal-01921830

https://hal.science/hal-01921830

Submitted on 14 Nov 2018

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.

MOLECULAR ECOLOGY

CO-OCCURRENCE AMONG THREE DIVERGENT PLANT-CASTRATING FUNGI IN THE SAME SILENE HOST SPECIES

Journal:	*Molecular Ecology*
Manuscript ID	MEC-18-0465.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Abbate, Jessica; Institut de recherche pour le developpement France-Sud, MIVEGEC Gladieux, Pierre Hood, Michael de Vienne, Damien Antonovics, Janis Snirc, Alodie; Université Paris-Sud, Ecologie, Systématique et Evolution Giraud, Tatiana; Université Paris-Sud, : Ecologie, Systématique et Evolution
Keywords:	Adaptation, Fungi, Host Parasite Interactions, Hybridization, Phylogeography, Population Genetics - Empirical
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
TG_figure_v2.svg	

SCHOLARONE™
Manuscripts

CO-OCCURRENCE AMONG THREE DIVERGENT PLANT-CASTRATING FUNGI IN THE SAME *SILENE* HOST SPECIES

Jessica L. Abbate[1,2*], Pierre Gladieux[3,4*], Michael E. Hood[5], Damien M. de Vienne[3,6,7], Janis Antonovics[8], Alodie Snirc[3], and Tatiana Giraud[3*]

* These authors contributed equally to the study

[1] ~~IRD,~~ UMR MIVEGEC, ~~CNRS,~~ IRD 224, CNRS, Université de Montpellier, ~~Montpellier,~~ F-34394 Montpellier, France

[2] UMR UMMISCO, IRD 209, UPMC, ~~Bondy,~~ F-93143 Bondy, France

[3] Laboratoire Ecologie Systématique et Evolution, Univ. Paris Sud, CNRS, AgroParisTech, Université Paris Saclay, Orsay, F-91400 France

[4] INRA, UMR BGPI, Bâtiment K; Campus International de Baillarguet, F-34398, Montpellier, France

[5] Biology Department, McGuire Life Sciences Building, Amherst College, Rts 9 & 116, Amherst, MA USA 01002-5000

[6] Laboratoire de Biométrie et Biologie Evolutive, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5558, Université Lyon 1, F-69622 Villeurbanne, France

[7] Université de Lyon, F-69000 Lyon, France;

[8] University of Virginia, Dept. of Biology, Gilmer Hall, Charlottesville, VA 22904, USA

*Corresponding author: Tatiana Giraud, Laboratoire Ecologie, Systématique et Evolution, Bâtiment 360, Université Paris-Sud, 91400 Orsay France

Email address: tatiana.giraud@u-psud.fr

phone: +33 1 69 15 56 69

fax: +33 1 69 15 46 97

Running title: Co-occurrence of three plant-castrating fungi

Formatted: French (France)

Field Code Changed

For Review Only

25

For Review Only

Abstract

The competitive exclusion principle postulates that different species can only coexist in sympatry if they occupy distinct ecological niches. The goal of this study was to understand the geographical distribution of three species of *Microbotryum* anther-smut fungi that are distantly related but infect the same host plants, the sister species *Silene vulgaris* and *S. uniflora,* in western Europe. We used microsatellite markers to investigate pathogen distribution in relation to host specialization and ecological factors. *Microbotryum violaceo-irregulare* was only found on *S. vulgaris* at high elevations in the Alps. *Microbotryum lagerheimii* could be subdivided into two genetically differentiated clusters, one on *S. uniflora* in the UK and the second on *S. vulgaris* in the Alps and Pyrenees. The most abundant pathogen species, *M. silenes-inflatae,* could be subdivided into four genetic clusters, co-occurring in the Alps, the UK and the Pyrenees, and was found on both *S. vulgaris* and *S. uniflora*. All three fungal species had high levels of homozygosity, in agreement with the selfing mating system generally observed in anther-smut fungi. The three pathogen species and genetic clusters had large range overlaps, but occurred at sites with different elevations, temperatures and precipitation levels. The three *Microbotryum* species thus do not appear to be maintained by host specialization or geographic allopatry, but instead may occupy different ecological niches in terms of environmental conditions.

Key words: population structure, endemicity, speciation, hybrid zones, fungi, *Microbotryum violaceum, Silene maritima*, biogeography, altitude

For Review Only

Introduction

The competitive exclusion principle postulates that two species occupying the same niche in sympatry cannot coexist indefinitely (Gause 1934; Hardin 1960). In pathogens, this means that species coexisting in sympatry in association with a common host species are likely to differ in other components of their ecological niches. The coexistence of multiple pathogens on the same crop species is in fact often associated with separation by time, resource use, abiotic conditions or geographical area (Fitt *et al.* 2006; Giraud *et al.* 2017). Closely related pathogens can also specialize on either the early or the late season part of a given plant life cycle (Hamelin *et al.* 2016; Mailleret *et al.* 2012). For example, two eyespot wheat pathogens, *O. yallundae* and *O. acuformis,* differ in temperature optima for growth and fungicide resistance (Fitt *et al.* 2006). In oilseed rape, the pathogens *Leptosphaeria maculans* and *L. biglobosa* differ in their location in stem and leaf tissues and in their maturation rates under low temperatures (Fitt *et al.* 2006). In malaria, variation in host resistance to different *Plasmodium* species has been invoked to explain pathogen coexistence (Snounou & White 2004). Furthermore, some *Plasmodium* species are specialized to attack only the youngest or oldest red blood cells, leading to complex within-host dynamics during co-infection (McQueen & McKenzie 2004a, b). Pathogens with different transmission modes, e.g. sexual versus non-sexual transmission, can also stably coexist (Thrall & Antonovics 1997).

The goal of this study was to delineate and understand the coexistence of different species of anther-smut fungi on two closely related host plants, *Silene vulgaris* and *S. uniflora*. *Microbotryum* is a species complex of basidiomycete fungi responsible for anther-smut disease in many plants in the Caryophyllaceae. *Microbotryum* anther-smut fungi are obligate pathogens that sterilize their hosts. Infected plants produce fungal teliospores in place of pollen, and female structures do not mature; in dioecious or gynodioecious species,

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

infected female plants also develop spore-bearing anthers. Within *Microbotryum*, there are numerous phylogenetically divergent lineages many of which have been given species status (Denchev *et al.* 2009; Kemler *et al.* 2006; Le Gac *et al.* 2007a). These typically show inter-sterility (de Vienne *et al.* 2009b; Le Gac *et al.* 2007b) and strong host specificity (de Vienne *et al.* 2009a). There is usually only one endemic *Microbotryum* species per host plant species (Le Gac *et al.* 2007a; Refrégier *et al.* 2008), although transient host shifts are not uncommon (Antonovics *et al.* 2002; Gladieux *et al.* 2011; Hood *et al.* 2003).

A conspicuous exception to this pattern is the presence of three divergent *Microbotryum* species on *S. vulgaris* and its closely related sister species *S. uniflora* (Figure 1). These are: 1) *M. lagerheimii* (Denchev 2007) (MvSv1 in Le Gac *et al.* 2007a), 2) *M. silenes-inflatae* (MvSv2 in Le Gac *et al.* 2007a), and 3) *M. violaceo-irregulare* (Kemler *et al.* 2006; Lutz *et al.* 2008). These three *Microbotryum* species are not particularly closely related to one-another (Figure 1), and show differences in spore morphology (color and ornamentation; Figure 1; Denchev 2007; Gold *et al.* 2009; Lutz *et al.* 2005; Lutz *et al.* 2008)). *Microbotryum lagerheimii* has lighter spore color, while verrucose spore ornamentation is unique to *M. violaceo-irregulare* (Figure 1; Lutz *et al.* 2005; Vánky 1994). All three species castrate host flowers in the same way, replacing the pollen by their spores and aborting ovaries, and are sexually transmitted (the spores are dispersed by pollinators), suggesting that it is unlikely they can have distinct ecological niches within host individuals. None of the three species has ever been found to persist in a host species other than *S. vulgaris* or *S. uniflora,* indicating that they are maintained endemically on these host plants.

One explanation for coexistence of these three fungal species could be that they are allopatric, although they have been found to co-occur in the same sites (Abbate & Antonovics 2014; Chung *et al.* 2012; Le Gac *et al.* 2007a); however sampling has been limited thus far

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

and it remains possible that contact zones of allopatric distributions were sampled. An alternative hypothesis is that the species display differential specialization towards *S. vulgaris* versus *S. uniflora.* Previous studies further suggest that this hypothesis is unlikely, given that both host species have been found to be parasitized in nature by at least two of the *Microbotryum* species also parasitizing *S. vulgaris*: *M. lagerheimii* and *M. silenes-inflatae* (Chung *et al.* 2012; Smith *et al.* 2017).

A more plausible explanation for the species coexistence could be adaptation to different abiotic conditions. All three species affecting *S. vulgaris* were found almost exclusively in host populations at higher elevations in the French Alps, where overall temperatures were lower and precipitation more consistently higher than conditions experienced by lower-elevation host populations (Abbate & Antonovics 2014). Here we investigated the generality of this pattern with a larger and broader sampling in order to further understand their coexistence. Differentiation in disjunct refugia during the last glaciation in Europe may also have shaped differential adaptation to abiotic conditions and/or generated allopatry. Strong population differentiation has been found in other *Microbotryum* species that reflects footprints of persistence in glaciation refugia, e.g. in *M. lychnidis-dioicae* parasitizing *S. latifolia* (Badouin *et al.* 2017; Feurtey *et al.* 2016; Gladieux *et al.* 2011; Vercken *et al.* 2010) and *M. silenes-acaulis* parasitizing *S. acaulis* (Bueker *et al.* 2016). High selfing rates were inferred for previously studied *Microbotryum* species (Bueker *et al.* 2016; Giraud 2004; Gladieux *et al.* 2011; Hood & Antonovics 2000; Hood & Antonovics 2004), and such a closed mating system may contribute to generate and maintain strong population structure and genetic isolation between species (Gibson *et al.* 2012; Vercken *et al.* 2010).

To assess the wider geographical and elevational distributions of the three different *Microbotryum* species occurring on *S. vulgaris* and *S. uniflora*, we collected anther-smut

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

120 fungi from both host populations in different regions in western Europe, and genotyped them
121 with microsatellite markers. We then examined whether the three *Microbotryum* species 1)
122 had contrasting geographical or elevational distributions, 2) showed different frequencies on
123 the two sibling host species *S. vulgaris* and *S. uniflora*, 3) co-occurred in the same sampling
124 sites, and if so, whether hybrids could be detected, 4) exhibited within-species geographical
125 population subdivision, and 5) displayed high selfing rates like other *Microbotryum* species.

126

127 **Materials and Methods**

128 ***Teliospore collection***

129 The *Microbotryum* samples analyzed in this study were collected across western Europe
130 (Supplementary Figure S1, Supplementary Table S1) from 51 sites with *S. vulgaris* (n = 407),
131 and 12 sites with *S. uniflora* (n = 59). A sampling site was defined as a location in which
132 samples were collected a few meters apart. Sampling was performed by numerous
133 investigators throughout Europe and was carried out as part of general disease surveys, other
134 field trips, or as part of more formal investigations (Chung *et al.* 2012; Le Gac *et al.* 2007a,
135 and unpublished studies). Within sampling sites, a single diseased flower was collected per
136 individual plant, stored in silica gel, and DNA from the diploid teliospores was extracted for
137 genetic analyses. Multiple infections by different genotypes are not uncommon in *Silene-*
138 *Microbotryum* systems (Bueker *et al.* 2016; Buono *et al.* 2014; Chung *et al.* 2012; Hood
139 2003; López-Villavicencio *et al.* 2011), but teliospores within a single flower originate from a
140 single diploid genotype (Gold *et al.* 2009; Lopez-Villavicencio *et al.* 2007). Teliospores
141 sampled from a given flower were therefore assumed to be from a particular pathogen
142 individual. Because of the difficulty of identifying the different *Microbotryum* species in the

For Review Only

field, sampling was carried out without knowledge of fungal species identity, so our sample corresponds to an unbiased representation of the *Microbotryum* species frequencies in the sampled regions. We obtained permits for collections in national parks and all samples were collected before 2014, so not falling under the Nagoya protocol. Dried spores from each sample are freely available upon request.

Microsatellite, ITS genotyping and spore morphology

DNA was extracted and genotyped using 11 microsatellite loci as described previously (SVG1, SVG2, SVG5, SVG6, SVG8, SVG15, SN2, SN5, SN11, DC2, and DC5; Giraud 2004; Giraud *et al.* 2008). Microsatellite scoring was performed as described previously (Vercken *et al.* 2010). Analyses using GENCLONE (Arnaud-Haond & Belkhir 2007) showed that the eleven markers were sufficient to discriminate multilocus genotypes in the dataset as a plateau was reached in terms of number of genotypes detected (Supplementary Figure S2).

In order to assign the genetic clusters from microsatellite analyses to the three *Microbotryum* species, we sequenced the ribosomal internal transcribed spacer (ITS), spanning ITS1, 5.8S and ITS2 fragments, in 27 individuals using the fungus-specific primers ITS1 and ITS4 (White *et al.* 1990) and matched these to published ITS sequences from *M. silenes-inflatae*, *M. lagerheimii*, *M. violaceo-irregulare* (Kemler *et al.* 2006; Lutz *et al.* 2008). Additionally, we verified in 15 individuals spore color (light versus dark purple) under the binocular and spore ornamentation (i.e. verrucose versus reticulate) under light microscopy. The samples chosen for ITS sequencing and for spore morphology assessment were picked at random within each cluster, and there was complete agreement in the species assignments between ITS, spore morphology and genetic clusters.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

166

167 ***Data analyses***

168 *ITS tree.* A maximum likelihood phylogeny of ITS region sequenced was inferred using
169 PhyML 3.0 (Guindon *et al.* 2010).

Field Code Changed

170

171 *Map.* The sample map (Supplementary Figure 2) was generated using the GGMAP package in
172 R environment (Kahle & Wickham 2013).

Field Code Changed

173

174 *Descriptive statistics.* Microsatellite variability was quantified by the unbiased gene diversity
175 (H_d), allelic richness (A_r), and the fixation index (F_{IS}) using FSTAT 2.9.3.2 (Goudet 2001).

Field Code Changed

176 Observed and unbiased expected heterozygosities (H_O and H_E) were computed using a custom
177 python script. Analyses of variance (ANOVA) and Kruskall-Wallis tests were performed
178 using JMP v7.0 (SAS Institute).

179

180 *Population subdivision.* We assessed population subdivision using Bayesian clustering
181 implemented in STRUCTURE 2.3.4 (Falush *et al.* 2003; Pritchard *et al.* 2000). The method

Field Code Changed

Field Code Changed

182 partitions multilocus genotypes into clusters while minimizing departure from expected
183 frequencies and linkage equilibrium among loci. For this analysis, we used a haploid setting
184 because *Microbotryum* individuals were highly homozygous (Table 1), which could bias
185 structure inferences based on Hardy-Weinberg expectations in a diploid setting. We
186 conducted independent runs with different number of clusters (K = 1 to 15). Each run of
187 400,000 iterations and 100,000 burn-in iterations was repeated ten times, using a model
188 allowing for admixture and correlated allele frequencies. We used CLUMPP 1.1.1 (Jakobsson

Field Code Changed

For Review Only

189 & Rosenberg 2007) to identify potential distinct solutions among the results of independent
190 replicate runs for each *K* ('Greedy' algorithm; 100 random input sequences; G' statistic), and
191 to average individual assignment probabilities (*q*) over replicated runs with identical
192 clustering solution.

193 Because *M. lagerheimii* and *M. violaceo-irregulare* were minority components of the
194 data set, they could not be differentiated in the overall analysis. We therefore also ran
195 STRUCTURE on a sub-dataset encompassing only genotypes assigned to these species based on
196 ITS sequencing. We retained for this sub-dataset the genotypes with total memberships $q \geq 0.8$
197 in the cluster corresponding to both *M. lagerheimii* and *M. violaceo-irregulare* in the analysis
198 on the whole dataset (blue cluster at K=5, see results). The 0.8 threshold is arbitrary but is
199 typically used for assigning individuals, and the genotypes with main assignments below 0.8
200 had assignment probabilities balanced among the different clusters, indicating they likely
201 corresponded to individuals with low assignment power (see results).

202

203 *Hybrid detection*

204 To assess whether individuals with intermediate assignment probability to clusters
205 were high-confidence inter-cluster hybrids or may be simply genotypes with low power of
206 assignment, possibly carrying alleles frequent in all species, we used a two-step procedure. In
207 the first step, we used the membership proportions of individual genomes in different clusters
208 (*q* values) inferred using the admixture model of STRUCTURE to identify 'pure' representatives
209 of the different species or of clusters within species. Based on the STRUCTURE outputs of the
210 full dataset at K=5 (identified as the most relevant subdivision, see results), pure
211 representatives of a given species were identified as genotypes for which the sum of *q* values

For Review Only

in the different clusters assigned to this species was above 0.8; genotypes not meeting this requirement were classified as putative hybrids. Pure representatives of the various clusters within a given species were defined as genotypes displaying *q* values above 0.8 within one of the clusters identified within this species; genotypes not meeting this requirement were considered as putative admixed genotypes among clusters.

In the second step, we used pure representatives of the species or clusters as 'learning samples', allowing estimation of the membership proportions for the putative hybrid samples in 'learning clusters', using STRUCTURE, options USEPOPINFO and UPDATEPFROMPOPFLAGONLY. In assignment tests of putative hybrids between *M. silenes-inflatae* and either *M. lagerheimii* or *M. violaceo-irregulare*, genotypes were considered as hybrids if they had *q* values above 0.2 and below 0.8 in learning clusters. For *M. lagerheimii* and *M. violaceo-irregulare*, genotypes were considered as hybrids if they had a membership proportion *q* above 0.2 and below 0.8 in the two learning clusters corresponding to *M. lagerheimii* and *M. violaceo-irregulare* for differentiating these two species. Because clustering with STRUCTURE can be strongly influenced by unbalanced sample sizes (Kalinowski 2011), and because sample size for *M. silenes-inflatae* was one order of magnitude higher than for the other species, datasets of pure representatives were randomly resampled ten times to the size of the smallest species sample (n=7, *M. violaceo-irregulare*), and STRUCTURE analyses were run on each resampled dataset.

Field Code Changed

In order to check the validity of the genetic subdivision retrieved from the STRUCTURE analysis, using a method not assuming outcrossing or lack of genetic disequilibrium, we also carried out a principal component analysis (PCA) on the microsatellite data using the 'dudi.pca' function in the R-package ADE4 (Dray & Dufour 2007). We then used a discriminant analysis of the principal components (DAPC), as implemented in the R-package

Field Code Changed

For Review Only

236 ADEGENET (Jombart 2008; Jombart *et al.* 2010). The 30 first principal components were
237 retained, representing 90% of total genetic variation, as well as the six first discriminant
238 functions. The membership probabilities of genotypes were computed using DAPC for
239 numbers of clusters *K* ranging from 1 to 25.

Field Code Changed

Field Code Changed

240

241 *Host specificity, geographical and climatic distributions*

242 To assess host specificity, we examined the differential occurrence of the pathogen
243 species and clusters on *S. vulgaris* and *S. uniflora*. We also plotted the geographical
244 distribution of the *Microbotryum* species and their genetic clusters to assess their degree of
245 allopatry. Because there was large variance in sample size at each site, with sites often having
246 only one individual sample, we used analyses based on presence and absence of a species or
247 cluster at a site. Analyses based on raw numbers or relative frequencies often violated model
248 assumptions. The associations between host species and pathogen species or genetic clusters
249 were tested using the 'CATMOD' procedure in the SAS Statistical Computing Environment
250 (SAS Institute Inc. 2013). Associations between species or genetic clusters and geographic
251 regions were tested using Fisher's exact tests with the SAS procedure 'FREQ' due to small
252 sample size for some categories.

Field Code Changed

253 To assess environmental effects, we compared the elevational distributions of the
254 *Microbotryum* species, and of the genetic clusters within *M. silenes-inflatae* and *M.*
255 *lagerheimii*. In addition to elevation, we incorporated in our analysis a suite of bioclimatic
256 variables for each of the sampling locations in order to assess whether components of climatic
257 variation could help explain the observed distributions. Nineteen bioclimatic variables
258 (Supplementary Table S2) were obtained from the WorldClim database (means across years

For Review Only

1950 – 2000) recorded from a network of weather stations and interpolated over a 30-arc second resolution grid (www.worldclim.org) (Hijmans *et al.* 2005). For each sampling site, values of the 19 variables were extracted using nearest-neighbor interpolation with the Spatial Analyst Tool in ArcMap 9.2 software (ESRI (Environmental Systems Resource Institute) 2009). We used a principal components analysis to analyze and reduce the variation to three major axes using the SAS procedure PRINCOMP (based on the correlation matrix). We then carried out a logistic regression to examine the relationship between species presence and the first three components generated by the PCA, using the 'glm' function in the R *stats* package with a quasi-binomial link.

Field Code Changed

Field Code Changed

Field Code Changed

For Review Only

Results

Species identification, species distributions and population subdivision

The STRUCTURE analysis on the whole dataset, modeling K=2 to 15 clusters, revealed several levels of strong genetic differentiation, with well-defined *Microbotryum* clusters appearing up to K=5 (i.e., clusters encompassing multiple genotypes with membership coefficients near the maximum; Figure 2 and Supplementary Figure S3). The genetic clusters were overall broadly distributed among sampled regions, although a few clusters appeared geographically more restricted (Figure 2 and Supplementary Figure S3). Some genetic differentiation was observed among samples collected on *S. uniflora* and *S. vulgaris*, with some genetic clusters being specific to either host species, but also with other genetic clusters occurring on both host species.

In order to assign the genetic clusters to the three previously described *Microbotryum* species, we sequenced ITS in 27 randomly chosen strains and built a maximum likelihood

282 tree, also including published ITS sequences from *M. silenes-inflatae*, *M. lagerheimii*, *M.*
283 *violaceo-irregulare* and other *Microbotryum* species (Supplementary Figure S4; Genbank
284 accession numbers MH491551-MH491577). Out of the 20 sequences in the fully supported
285 clade containing the published *M. silenes-inflatae* ITS sequences, 19 were genotypes assigned
286 to the largest STRUCTURE cluster in analyses assuming *K*=2 clusters (pink cluster;
287 Supplementary Figure S4), indicating that this cluster corresponds to *M. silenes-inflatae*.
288 Analyses at *K*>2 only further subdivided *M. silenes-inflatae* into several clusters (Figure 2).
289 The ITS sequences falling in the fully supported clades containing the published *M.*
290 *lagerheimii* (*n*=5) and *M. violaceo-irregulare* (*n*=2) sequences all belonged to genotypes in
291 the second (blue) cluster of the *K*=2 STRUCTURE model (Supplementary Figure S4).

292 The log-likelihood plateaued at *K*>5 with only a single clustering solution
293 (Supplementary Figure S5A); the additional clusters appearing at K values above 5 mainly
294 introduced additional subdivision within *M. silenes-inflatae,* produced many genotypes with
295 intermediate assignment probability and never clearly separated the genotypes assigned to *M.*
296 *lagerheimii* and *M. violaceo-irregulare* using ITS into distinct clusters. *K*=5 was therefore
297 considered as the most relevant *K* value, with the blue cluster corresponding to both *M.*
298 *lagerheimii* and *M. violaceo-irregulare,* and the four other clusters (green, orange, pink and
299 yellow) to different lineages within *M. silenes-inflatae.*

300 The species *M. lagerheimii* and *M. violaceo-irregulare* could not be separated into
301 distinct clusters at any *K* values using the full dataset, despite their large phylogenetic
302 distance (Figure 1). The lack of ability to distinguish these species is likely due to their small
303 sample sizes compared to that collected for *M. silenes-inflatae.* Such difficulties in
304 recovering species differentiation in cases of unbalanced sample sizes have been shown by
305 simulations (Neophytou 2014). We therefore ran STRUCTURE on a sub-dataset with the

Field Code Changed

For Review Only

genotypes from *M. lagerheimii* and *M. violaceo-irregulare* only. At *K*=3 in this sub-dataset (Figure 3; Supplementary Figures S5B and S6), one cluster (brown) corresponded to *M. violaceo-irregulare* genotypes while the other two clusters (dark and light blue) represented *M. lagerheimii* genotypes, mainly separating isolates collected on *S. uniflora* in the UK and Netherlands (light blue) versus genotypes collected on *S. vulgaris* in the Pyrenees and Alps (dark blue). Spore morphology (color and ornamentation) of 15 individuals, chosen at random within each species, provided further evidence for cluster assignment to species: all five strains assigned to *M. silenes-inflatae* had dark purple spores with reticulate ornamentation, all five strains assigned to *M. lagerheimii* had light purple spores with reticulate ornamentation, and all strains assigned to *M. violaceo-irregulare* had dark purple spores with verrucose ornamentation (Supplementary Table S3).

The principal component analysis (PCA) on the full dataset separated the three species, with *M. lagerheimii* and *M. silenes-inflatae* in separate corners of the parameter space and *M. violaceo-irregulare* appearing intermediate between them (Figure 4A). The ITS sequence variants particular to *M. silenes-inflatae* were not confined to any of its four genetic clusters (Supplementary Figure S3; Figures 4A and 4B), providing evidence that they constitute within-species genetic subdivision. The *M. lagerheimii* and *M. violaceo-irregulare* species and two *M. silenes-inflatae* clusters were found on both *S. vulgaris* and *S. uniflora*, while the two other *M. silenes-inflatae* clusters occurred only on *S. vulgaris* (Figure 2). A PCA run on the sub-dataset of the two less frequent *Microbotryum* species also discriminated them well (Figures 4C and 4D).

Clustering using DAPC supported the clustering patterns found in STRUCTURE analyses (Supplementary Figure S7 and Figure 2). The DAPC Bayesian information criterion reached a minimum value at *K*=28 (Supplementary Figure S8). However, visual inspection of

For Review Only

membership probability patterns across K values indicated that DAPC and STRUCTURE identified the same five main clusters (Supplementary Figure S7 and Figure 2). The DAPC on the full dataset could not distinguish *M. lagerheimii* and *M. violaceo-irregulare* into different clusters either, while the DAPC on the genotypes from these two species separately produced a pattern of membership proportions largely similar to the one obtained with STRUCTURE (Supplementary Figure S9).

Hybridization and admixture checking

We performed assignment tests using the option in STRUCTURE that uses pure reference genotypes as learning samples to determine whether individuals with intermediate assignment probabilities could simply be genotypes with low assignment power instead of hybrids or admixed genotypes. Within *M. silenes-inflatae*, we identified 64, 40, 75 and 79 genotypes representatives of the yellow, orange, pink and green clusters, respectively, that we used as learning samples. In assignment tests for within-species admixture, all individuals with intermediate assignment probabilities had comparable *q* values in multiple learning samples (q values ranging from 0.193 to 0.349, i.e. approximately one divided by the number of clusters), suggesting that all genotypes with intermediate assignment probabilities could be genotypes with low assignment power instead of genuine admixed individuals (Supplementary Table S4). Within *M. lagerheimii*, we identified 24 and 12 pure representatives of the dark and light blue clusters, respectively; the only individual with intermediate assignment probabilities in the *M. lagerheimii* full dataset (q=0.465 in the blue cluster; q=0.535 in the light blue cluster) had high membership in the blue cluster (q=0.731) in the *M. violaceo-irregulare*/*M. lagerheimii* subset analysis.

For Review Only

We identified 7, 38 and 400 pure genotype representatives of *M. violaceo-irregulare*, *M. lagerheimii* and *M. silenes-inflatae*, respectively, that we used as learning samples. In assignment tests for inter-specific hybrid identification, all genotypes had substantial membership in all learning samples considered. Therefore, no genotype could be identified as a high-confidence hybrid, neither between species nor between lineages within *M. silenes-inflatae* (Supplementary Table S5), and they may instead represent genotypes with low assignment power. This inference was supported by the heterozygosity values of the genotypes with intermediate assignment probabilities, whose mean (H_O=0.18) was not higher than that in other samples (Table 1), in contrast to what would be expected for hybrids.

Host specificity, geographical and climatic distributions

Pathogen species and several clusters within *M. silenes-inflatae* and *M. lagerheimii* were differentially distributed among host species. Because pathogens were sampled without prior knowledge of their genotype, their differential distribution was not due to sampling bias. The rarest species, *M. violaceo-irregulare,* was found only on *S. vulgaris* in the western and central Alps (Figure 5H and Supplementary Figure S10). The two other species, *M. lagerheimii* and *M. silenes-inflatae,* were both found in the Alps, the Pyrenees and the UK Figure 5H and Supplementary Figure S10). However, the two genetic clusters within *M. lagerheimii* were well-separated both geographically and with regard to host species, with the light blue cluster occurring only on *S. uniflora* in the UK and the dark blue cluster only on *S. vulgaris* in the Alps and the Pyrenees (Figure 5, F and G). While all *Microbotryum* species were found on *S. vulgaris*, *M. silenes-inflatae* was found in 90% of *S. vulgaris* populations but just 50% of *S. uniflora* populations, the difference in frequency being significant (Fisher's exact test, $p = 0.004$). Clusters within *M. silenes-inflatae* had largely overlapping

For Review Only

377 distributions on *S. vulgaris*, except the pink cluster which occurred only in the Alps and the
378 Pyrenees (Figure 5 A-E). The green cluster of *M. silenes-inflatae* was found commonly on *S.*
379 *uniflora* in the UK (Fig 5 B), while only single records of the yellow and orange clusters were
380 identified there. This bias towards occurrence in *S. vulgaris* populations was statistically
381 significant only for the yellow cluster (Fisher's exact test, $p = 0.025$). In the single site where
382 host plants were identified as *S. vulgaris* var. *prostrata* (Site 22), the high elevation form of
383 the host, all strains were assigned to *M. violaceo-irregulare*.

384 Different species and clusters within species often occurred in the same sampling
385 site. Among the 45 sites with more than one assigned sample, nine displayed sympatry of
386 *Microbotryum* species (*M. silenes-inflatae* and either *M. lagerheimii* or *M.violaceo-*
387 *irregulare*). Within *M. silenes-inflatae*, 11 sampling sites out of 32 displayed sympatry among
388 at least two genetic clusters (Supplementary Figure S10). Genotypes with intermediate
389 assignment probabilities in STRUCTURE were not more frequently found in sites with co-
390 occurring clusters ($\chi^2 = 0.36$, df = 1, $p = 0.54$), supporting the interpretation that they
391 represent genotypes with low assignment power rather than hybrid individuals. Moreover, the
392 genotypes with intermediate assignment probabilities in different species were present as
393 frequently in sites where only a single species was found as in sites where two or more
394 species were found ($\chi^2 = 1.17$, df = 1, $p = 0.26$). Genotypes with intermediate assignment
395 probabilities in different *M. silenes-inflatae* clusters were in 33 out of 40 (82%) sites with a
396 single cluster, and the genotypes with intermediate assignment probabilities in different *M.*
397 *lagerheimii* clusters were in 13 out of 17 (76%) sites with a single cluster (grey genotypes in
398 Supplementary Figure S10).

399 Where our sampling was most extensive (Pyrenees, western Alps and central Alps),
400 we investigated whether the distributions of the species (51 sites) and of clusters within *M.*

For Review Only

silenes-inflatae (37 sites with unambiguous assignments) were associated with local geography, elevation and/or climate. The occurrence of *M. silenes-inflatae* did not differ between the Pyrenees, western Alps or central Alps (Fisher's exact test, $p = 0.168$). The occurrence of *M. violaceo-irregulare* and *M. lagerheimii* differed significantly among regions (Fisher's exact tests, $p = 0.031$ and $p = 0.001$ respectively), with the former being absent in the Pyrenees and the latter absent in the Western Alps (Figure 5 and Supplementary Figure 10). Within *M. silenes-inflatae*, the green cluster showed significant differences in occurrence among regions (Fisher's exact test, $p = 0.025$), being more common in the western Alps. The analysis approached significance for the orange cluster (Fisher's exact test, $p = 0.052$) which was more common in the Pyrenees (Figure 5 and Supplementary Figure 10). The other two clusters showed no significant differences in occurrence among regions (Fisher's exact tests, $p = 0.634$ for pink and $p = 0.408$ for yellow).

The three species occurred at significantly different elevations (Figure 6; ANOVA, df = 2, Sum of squares = 4486478, F-ratio = 8.7, $p = 0.0002$). The elevational distribution was not normal, but ANOVAs are robust to such assumption violations (Lindman 1974) and a non-parametric test also indicated significant differences in elevation among species (Kruskall-Wallis non-parametric test, df = 2, $\chi^2 = 16.7$, $p = 0.002$). The species *M. violaceo-irregulare* was found only at high elevations, *M. lagerheimii* at the lowest elevations, while *M. silenes-inflatae* was found across a greater range of elevations (Figure 6). The two clusters within *M. lagerheimii* were also found at significantly different elevations (Figure 6; ANOVA, df = 1, Sum of squares = 21620792, F-ratio = 167.9, $p < 0.0001$; Kruskall-Wallis test, df = 1, $\chi^2 = 24.1$, $p < 0.0001$), as well as the four clusters within *M. silenes-inflatae* (Figure 6; ANOVA, df = 3, Sum of squares = 10011015, F-ratio = 24.3, $p < 0.0001$; Kruskall-Wallis test, df = 2, $\chi^2 = 56.6$, $p < 0.0001$). In *M. lagerheimii*, the contrast in elevation between

Field Code Changed

For Review Only

its two clusters was associated with the differences in host species, as described above, i.e. coastal *S. uniflora* in the UK versus broad elevational distribution of *S. vulgaris* elsewhere. The Figure 2 also illustrates the differences in elevation between species and clusters within each geographical region: clusters remain well separated when genotypes are sorted by elevation, in particular at K=5.

Differences in *Microbotryum* distributions among both elevations and geographic regions could indicate impacts of climate. To test whether species or clusters were associated with different elevations and/or climates among the most heavily sampled regions (*S. vulgaris* populations in the Pyrenees, western Alps, and central Alps), we carried out a two-step analysis to first reduce climatic variables to uncorrelated principal components (PCs), and then test for the relative ability of climate, elevation and geographical location to explain species and cluster distributions. For both species and cluster datasets, the 19 bioclimatic variables (Supplementary Table S2) were reduced to three axes using PCA. All PC axis loadings were very similar for both species and cluster analyses (Supplementary Tables S2 and S6). Axis 1 (explaining 63 and 66% of the variation for species and cluster, respectively) represented decreasing temperature and increasing precipitation; axis 2 (19.3 and 19.4% of the variation) was weighted heavily on increased seasonality of precipitation and more moderate interactions between temperature and rainfall; axis 3 (8.3 and 7.6% of the variation) was weighted positively on more mild winter temperatures and lower temperature seasonality (Supplementary Tables S2 and S6).

Due to the high degree of collinearity among variables, effects of climate, elevation and geographic location were then analyzed in a second PCA that incorporated latitude, longitude, elevation, as well as scores for the first three principal components of the bioclimatic analysis. Three major axes summarized the variation, similarly for species and

For Review Only

449 cluster analyses (Supplementary Table S7). The first axis (accounting for 66.3 and 48.1% of
450 the variation in the species and cluster analyses, respectively) was weighted heavily on
451 geographical position (latitude and longitude) and precipitation seasonality (PC axis 2 in the
452 bioclimate analysis). The second axis (33.0 and 32.6% of the variation) was weighted heavily
453 on elevation and the three climatic PC axes (especially axis 1 for temperature and rainfall)
454 and was almost completely independent of location. Axis 3 (17.0 and 17.1%) was relatively
455 independent of both location and elevation, but weighted strongly on climatic axes
456 representing fewer extremes of cold and dry periods.

457 The logistic regression analysis on species (Supplementary Table S8) showed a
458 significant negative association between the occurrence of *M. lagerheimii* and location and
459 seasonality in precipitation (PC axis 1), suggesting that either climate, location, or both
460 contributed to explain this species' distribution. For the logistic regression analysis on *M.*
461 *silenes-inflatae* genetic clusters (Supplementary Table S8), occurrence of the orange cluster
462 was significantly associated with PC axis 2 (reflecting increased elevation and its
463 characteristic colder and wetter climate), and the yellow cluster approached significant
464 negative association with PC3 (indicating greater chance of occurrence at sites with more
465 extreme cold and dry periods, independent of geography or elevation) (Supplementary Table
466 S8).

467

468 *Levels of genetic variation and mating system*

469 Summary statistics were computed for clusters identified at *K*=5 in analyses of the full dataset
470 for *M. silenes-inflatae*, and at *K*=3 in analyses of the *M. lagerheimii-M. violaceo-irregulare*
471 dataset (Table 1). The genotypes with intermediate assignment probabilities in STRUCTURE
472 outputs were excluded. Low heterozygosity levels were found in all three species, leading to

For Review Only

high levels of inferred selfing rates (mean of 0.67, with range of 0.33-0.92, Table 1). Higher selfing rates were estimated in *M. violaceo-irregulare* (0.92) and *M. lagerheimii* (0.80-0.86) than in *M. silenes-inflatae* (0.33-0.73).

Discussion

This study aimed to understand the coexistence of three distant species of anther-smut fungi, all castrating pathogens of the sister host species, *Silene vulgaris* and *S. uniflora*. The three recognized *Microbotryum* species that are specific to these hosts were found in our sampling, although with contrasting frequencies: *M. silenes-inflatae* was by far the most prevalent, followed by *M. lagerheimii* and then *M. violaceo-irregulare*. These three phylogenetically distant species, previously distinguished based on morphological characters, were unambiguously identified based on microsatellite markers complemented with ITS sequences and spore morphology. We had to analyze a restricted microsatellite dataset to discriminate between *M. lagerheimii* and *M. violaceo-irregulare,* likely because of their low abundances compared to *M. silenes-inflatae* (Neophytou 2014). Despite co-occurring in the same host species and sites, no high-confidence inter-specific hybrids were identified between the pathogen lineages, in agreement with the strong post-zygotic isolation generally found among distant *Microbotryum* species (de Vienne *et al.* 2009b; Le Gac *et al.* 2007b; Petit *et al.* 2017). However, no crossing studies have experimentally investigated post-zygotic isolation among these particular *Microbotryum* species, and we cannot exclude the possibility that our microsatellite markers lacked power to identify hybrids with sufficient confidence, in particular later-generation backcrossed genotypes.

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

For Review Only

We detected genetic subdivision within species, with two clusters in *M. lagerheimii* and four clusters in *M. silenes-inflatae*. Given their weak genetic differentiation (as shown by ITS sequences and distribution in the PCA), the genetic clusters within *M. silenes-inflatae* likely do not represent further cryptic species. Within *M. lagerheimii,* the genetic subdivision corresponded to separation in geography, host and environmental conditions and it remains to be determined if these might be cryptic species. *Silene vulgaris* and *S. uniflora* were allopatric during the last glaciation (Prentice *et al.* 2011), with *S. uniflora* likely persisting in northern refugia from the Baltic/Scandinavian region and *S. vulgaris* in southern European refugia. This suggests that the differentiation between the two *M. lagerheimii* clusters found on *S. vulgaris* and *S. uniflora*, respectively, may result from past separation in disjunct glacial refugia with contrasting climates rather than host adaptation.

Subdivisions within *M. silenes-inflatae* had less obvious biological interpretation as the clusters were broadly overlapping in geography. Indeed, all four *M. silenes-inflatae* clusters were found in the Alps, the UK and the Pyrenees on *S. vulgaris.* However, only the green cluster was found regularly on *S. uniflora* in the UK. Additionally, this green cluster was more frequent in the western Alps than in the Pyrenees, while the opposite was found for the orange cluster, suggesting possible allopatric or host-related differentiation of the *M. silenes-inflatae* genetic clusters. The geographic co-occurrence of the four *M. silenes-inflatae* clusters in the Alps may result from recent movements and secondary contact. Footprints of a glacial refugia in Europe have been found in *M. lychnidis-dioicae* parasitizing *S. latifolia* (Badouin *et al.* 2017; Feurtey *et al.* 2016; Gladieux *et al.* 2011; Vercken *et al.* 2010) and *M. silenes-acaulis* parasitizing *S. acaulis* (Bueker *et al.* 2016). The genetic clusters within *M. silenes-inflatae* have not homogenised despite sympatry, which is likely due to the selfing mating system of these fungi. All three *Microbotryum* species indeed had high estimated

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

selfing levels, being highly homozygous, especially *M. violaceo-irregulare* and *M. lagerheimii*, as has been found for other *Microbotryum* species previously studied (e.g., those on *S. latifolia, S. dioica, S. acaulis, Saponaria officinalis* and *Dianthus* spp.; Bucheli *et al.* 2001; Bueker *et al.* 2016; Delmotte *et al.* 1999; Fortuna *et al.* 2016; Giraud 2004; Gladieux *et al.* 2011; Hood & Antonovics 2000; Hood & Antonovics 2004; Petit *et al.* 2017). Such high selfing rates may be the result of mating among sporidia from one diploid individual (a spore load from a single infected flower). Additionally, it may be the result of mating between sporidia from single teliospores, a process also promoted by intra-tetrad mating that is present in many *Microbotryum* species as a result of ordered segregation of the mating-type loci (Hood & Antonovics 2000; Hood *et al.* 2015). Interestingly, *M. silenes-inflatae* had lower inferred selfing rates than the other species, and it would be ~~interesting~~informative to investigate in future studies if this is due to ecological factors, such as more frequent co-dispersal of spores from different fungal individuals by pollinators, or to intrinsic factors, such as a lower propensity of intra-tetrad mating.

The three *Microbotryum* species had large range overlaps on *S. vulgaris*, with frequent co-occurrence within sampling sites, indicating that allopatry may not be a major factor currently allowing the persistence of these distinct pathogen species on the same host. The distribution of *M. violaceo-irregulare* and *M. lagerheimii* was nevertheless significantly different between the local regions in the Pyrenees and Alps, suggesting they may have a history of past allopatry. However, their current broad range overlap indicates that their coexistence must either be transient or promoted by niche specialization.

The co-existence of the three *Microbotryum* species is clearly not maintained by host specialization, as all three species occurred on *S. vulgaris*. Moreover, it has been shown that resistance of *S. uniflora* to *M. lagerheimii* and *M. silenes-inflatae* was significantly positively

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

correlated among plant families (Chung *et al.* 2012), which indicates that the plant resistance is not specific to any one *Microbotryum* species. This further supports the inference that ~~host~~ specialization on *S. uniflora* versus *S. vulgaris* is not a major factor contributing to the co-occurrence of these *Microbotryum* species. However, because the host was not genotyped in our study and no infection trials were carried out, we cannot ~~be~~ rule~~d~~ out ~~that~~the influence of ~~different~~ host genotypes ~~influence~~on *Microbotryum* species distribution and coexistence.

The situation of the anther-smut species co-occurring on *S. vulgaris* is in some respects similar to that of anther-smut disease on the genus *Dianthus* where multiple *Microbotryum* species co-occur in the same sampling sites on multiple *Dianthus* species (Petit *et al.* 2017). In this case, however, the *Microbotryum* species on *Dianthus* hosts are closely related phylogenetically and hybrids, likely facilitated by the proximal relationships between the *Microbotryum* species, are frequently found (Petit *et al.* 2017). The situation on *Dianthus* spp. and the *S. vulgaris/S. uniflora* sister pair contrasts with the pattern found in all Caryophyllaceae studied to date, where each host species has been found to harbor only one endemic *Microbotryum* species in natural populations (Le Gac *et al.* 2007a; Refrégier *et al.* 2010). *Silene vulgaris* also occasionally harbors *Microbotryum* species as a result of transient cross-species disease transmission from other *Silene* hosts, such as *S. latifolia* and *S. dioica* (Antonovics *et al.* 2002; Hood *et al.* 2003). It is unclear if this is due to a particular susceptibility of *S. vulgaris* to anther smut or to its very wide geographic and environmental ranges that allow contact with the disease on many other *Silene* hosts.

The three anther-smut species on *S. vulgaris* occurred at significantly different elevations, suggesting that they may occupy different ecological niches in terms of abiotic conditions. This was supported by some association of *M. lagerheimiii* occurrence with climatic variables related to temperature and precipitation. Within *M. silenes-inflatae*, the

Field Code Changed
Formatted: English (U.S.)
Formatted: English (U.S.)
Formatted: English (U.S.)
Formatted: English (U.S.)
Formatted: Font: Italic
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

relative frequency of some of the genetic clusters also varied with elevation and climatic factors. These effects may reflect environmental tolerances of the fungal species and clusters, or ecotypic differences among host populations. It is well known that the host *S. vulgaris* has different ecotypes adapted to contrasting elevations (Marsden-Jones & Turrill 1957), and these may have different anther-smut specificities. In the present study, just one host population was identified as *S. vulgaris* var. *prostrata,* the high elevation form of the host (Site 22, in which all infections were caused by *M. violaceo-irregulare*). However, a previous study of *Microbotryum* on *S. vulgaris* in the small region of the south-eastern French Alps where Site 22 is located showed that none of the three fungal species were confined to any specific host ecotype (Abbate & Antonovics 2014). Furthermore, phenotypic traits including those that mark differences between the ecotypes change gradually over elevation, and the differentiation between populations is relative to their geographic distance (Berardi *et al.* 2016). The much wider sampling in the present study confirms that *Microbotryum* on *S. vulgaris* is largely confined to higher elevations (79% of the samples, 88% of the sites, were found above 1300m), even though populations of the host plant are abundant at lower elevations across its range. More generally, elevation and abiotic conditions have been shown to impact the presence of fungal pathogens in other systems (Cordier *et al.* 2012; Desprez-Loustau *et al.* 1998; Enjalbert *et al.* 2005; Gange *et al.* 2007; Mboup *et al.* 2012; Vacher *et al.* 2008). This study thus begs important questions for future work with *Microbotryum* pathogens investigating how the genetics of adaptation to elevation in fungi interacts with the adaptation to, or of, host plants. For example, one question is whether *Microbotryum* is rare at low elevations because the fungus cannot adapt to low-elevation climatic conditions, or because the costs of resistance in the host to the pathogen decline under those conditions.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

For Review Only

In conclusion, using a molecular ecology approach we showed that we can assign samples to the three *Microbotryum* anther-smut pathogens parasitizing *S. vulgaris* and *S. uniflora* using microsatellite markers. Furthermore, we revealed the existence of four genetic clusters within *M. silenes-inflatae*. We also showed that the different species and genetic clusters had large range overlaps and were all found on *S. vulgaris*. These findings suggest that the coexistence of multiple species and genetic clusters of *Microbotryum* pathogens on *S. vulgaris* and *S. uniflora* cannot be explained by differences in host or geographic distributions alone, although recent secondary contacts cannot be entirely excluded. Instead, we found that the different species and genetic clusters appeared to occupy different ecological niches in terms of abiotic conditions, in particular elevation, temperature and precipitation. This study highlights the importance of including not only species differences and geography but also abiotic conditions when considering the distribution and co-existence of pathogenic fungi on similar hosts.

Acknowledgements

We thank the following people for help with strain collection: Mickael Le Gac, Jacqui Shykoff, Anne Thomas, Michel Bartoli, Alex Widmer, Lucie Salvaudon, Erhart, Marcel Dorken, Marc-André Selosse, Rémy Relli, Jean-Pierre Thoirain, Fabienne van Rossum, Amanda Gibson, Alexis Matamoro, and the French 'Parc National des Pyrénées'. We thank Ricardo Rodriguez de la Vega for help with figures. T.G. acknowledges the grant ANR-12-ADAP-0009 (Gandalf project). M.E.H. acknowledges the grant NSF-DEB 0747222 and 1115765. J.L.A. acknowledges the National Science Foundation award DEB-0910212, ANR grant 10-PDOC- 017-01 'SPATEVOLEPID' awarded to Sébastien Lion, and ANR JC ~~grant~~

For Review Only

'STORY' grant awarded to Benjamin Roche. JA acknowledges NSF grant DEB-1115765 and NIH grant R01GM122061. Sequencing was funded by the ''Consortium National de Recherche en Génomique'' and the ''service de systématique moléculaire'' of the Muséum National d'Histoire Naturelle (CNRS IFR 101). It is part of the agreement n_2005/ between the Genoscope, the Muséum National d'Histoire Naturelle the INRA and the CEA (Genoscope), on the projects ''Macrophylogeny of life'', "SPEED-ID" and "Bibliothèque du Vivant" directed by Guillaume Lecointre and Jean-Yves Rasplus.

Data accessibility

The genotypes and elevation data are available on Dryad : doi:10.5061/dryad.627fm05.

GenBank accession numbers for ITS sequences: MH491551-MH491577.

~~The genotypes and elevation data will be submitted to Dryad upon acceptance.~~

Author contributions

TG, JAn and MEH designed the study; AS, JAb and DdV genotyped samples; PG, JAn, JAb and TG analysed data; PG, JAb and TG wrote the manuscript with contributions by all authors.

REFERENCES

Abbate JL, Antonovics J (2014) Elevational disease distribution in a natural plant-pathogen system: insights from changes across host populations and climate. *Oikos* **123**, 1126-1136.

For Review Only

636 Antonovics J, Hood M, Partain J (2002) The ecology and genetics of a host shift:
637 *Microbotryum* as a model system. *American Naturalist* **160**, S40-S53.

638 Arnaud-Haond S, Belkhir K (2007) GENCLONE: a computer program to analyse genotypic
639 data, test for clonality and describe spatial clonal organization. *Mol. Ecol. Notes.* **7**,
640 15-17.

641 Badouin H, Gladieux P, Gouzy J*, et al.* (2017) Widespread selective sweeps throughout the
642 genome of model plant pathogenic fungi and identification of effector candidates. *Mol*
643 *Ecol* **26**, 2041-2062.

644 Berardi AE, Fields PD, Abbate JL, Taylor DR (2016) Elevational divergence and clinal
645 variation in floral color and leaf chemistry in *Silene vulgaris*. *American Journal of*
646 *Botany* **103**, 1508-1523.

647 Bucheli E, Gautschi B, Shykoff JA (2001) Differences in population structure of the anther
648 smut fungus *Microbotryum violaceum* on two closely related host species, *Silene*
649 *latifolia* and *S. dioica*. *Molecular Ecology* **10**, 285-294.

650 Bueker B, Eberlein C, Gladieux P*, et al.* (2016) Distribution and population structure of the
651 anther smut *Microbotryum silenes-acaulis* parasitizing an arctic-alpine plant.
652 *Molecular Ecology* **25**, 811-824.

653 Buono L, Lopez-Villavicencio M, Shykoff JA, Snirc A, Giraud T (2014) Influence of
654 multiple infection and relatedness on virulence: disease dynamics in an experimental
655 plant population and its castrating parasite. *Plos One* **9**, 6.

656 Chung E, Petit E, Antonovics J, Pedersen AB, Hood ME (2012) Variation in resistance to
657 multiple pathogen species: anther smuts of *Silene uniflora*. *Ecology and Evolution* **2**,
658 2304-2314.

659 Cordier T, Robin C, Capdevielle X*, et al.* (2012) The composition of phyllosphere fungal
660 assemblages of European beech *(Fagus sylvatica)* varies significantly along an
661 elevation gradient. *New Phytologist* **196**, 510-519.

662 de Vienne DM, Hood ME, Giraud T (2009a) Phylogenetic determinants of potential host
663 shifts in fungal pathogens. *J Evol Biol* **22** 2532-2541.

664 de Vienne DM, Refrégier G, Hood M*, et al.* (2009b) Hybrid sterility and inviability in the
665 parasitic fungal species complex *Microbotryum*. *Journal of Evolutionary Biology* **22**,
666 683–698.

667 Delmotte F, Bucheli E, Shykoff JA (1999) Host and parasite population structure in a natural
668 plant-pathogen system. *Heredity* **82**, 300-308.

669 Denchev C, Giraud T, Hood M (2009) Three new species of anthericolous smut fungi on
670 Caryophyllaceae. *Mycologia Balcanica* **6**, 79-84.

671 Denchev CM (2007) *Microbotryum lagerheimii* sp. nov. (Microbotryaceae). *Mycologia*
672 *Balcanica* **4**, 61-67.

673 Desprez-Loustau ML, Capron G, Dupuis F (1998) Relating germination dynamics of
674 *Melampsora pinitorqua* teliospores to temperature and rainfall during overwintering.
675 *European Journal of Forest Pathology* **28**, 335-347.

676 Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for
677 ecologists. *Journal of Statistical Software* **22**, 1-20.

For Review Only

Enjalbert J, Duan X, Leconte M, Hovmoller MS, De Vallavieille-Pope C (2005) Genetic evidence of local adaptation of wheat yellow rust *(Puccinia striiformis f. sp tritici)* within France. *Molecular Ecology* **14**, 2065-2073.

ESRI (Environmental Systems Resource Institute) (2009) ArcMap 9.2. . *ESRI, Redlands, California*.

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

Feurtey A, Gladieux P, Hood ME*, et al.* (2016) Strong phylogeographic co-structure between the anther-smut fungus and its white campion host. *New Phytologist* **212**, 668-679.

Fitt B, Huang Y-J, van den Bosch F, West JS (2006) Coexistence of related pathogen species on arable crops in space and time. *Ann Rev Phytopathol* **44**, 163-182.

Fortuna TM, Snirc A, Badouin H*, et al.* (2016) Polymorphic microsatellite markers for the tetrapolar anther-smut fungus *Microbotryum saponariae* based on genome sequencing. *PLoS One* **11**, e0165656.

Gange AC, Gange EG, Sparks TH, Boddy L (2007) Rapid and recent changes in fungal fruiting patterns. *Science* **316**, 71-71.

Gause GF (1934) *The Struggle For Existence (1st ed.)* Williams & Wilkins, Baltimore.

Gibson AK, Hood ME, Giraud T (2012) Sibling competition arena: selfing and a competition arena can combine to constitute a barrier to gene flow in sympatry. *Evolution* **66**, 1917-1930.

Giraud T (2004) Patterns of within population dispersion and mating of the fungus *Microbotryum violaceum* parasitising the plant *Silene latifolia*. *Heredity* **93**, 559-565.

Giraud T, Koskella B, Laine AL (2017) Introduction: microbial local adaptation: insights from natural populations, genomics and experimental evolution. *Molecular Ecology* **26**, 1703-1710.

Giraud T, Yockteng R, Marthey S*, et al.* (2008) Isolation of 60 polymorphic microsatellite loci in EST libraries of four sibling species of the phytopathogenic fungal complex *Microbotryum*. *Molecular Ecology Resources* **8**, 387-392.

Gladieux P, Vercken E, Fontaine M*, et al.* (2011) Maintenance of fungal pathogen species that are specialized to different hosts: allopatric divergence and introgression through secondary contact. *Mol Biol Evol* **28**, 459-471.

Gold A, Giraud T, Hood ME (2009) Within-host competitive exclusion among species of the anther smut pathogen. *BMC Ecology* **9**, 11.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/popgen/softwares/fstat.htm.

Guindon S, Dufayard JF, Lefort V*, et al.* (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**, 307-321.

Hamelin FM, Bisson A, Desprez-Loustau ML, Fabre F, Mailleret L (2016) Temporal niche differentiation of parasites sharing the same plant host: oak powdery mildew as a case study. *Ecosphere* **7**, 11.

Field Code Changed

For Review Only

Hardin G (1960) The competitive exclusion principle. *Science* **131**, 1292-1297.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965-1978.

Hood ME (2003) Dynamics of multiple infection and within-host competition by the anther-smut pathogen. *American Naturalist* **162**, 122-133.

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, Antonovics J, Heishman H (2003) Karyotypic similarity identifies multiple host-shifts of a pathogenic fungus in natural populations. *Infection Genetics and Evolution* **2**, 167-172.

Hood ME, Antonovics JA (2004) Mating within the meiotic tetrad and the maintenance of genomic heterozygosity. *Genetics* **166**, 1751-1759.

Hood ME, Scott M, Hwang M (2015) Breaking linkage between mating compatibility factors: Tetrapolarity in *Microbotryum*. *Evolution* **69**, 2561-2572.

Jakobsson M, Rosenberg N (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**, 94.

Kahle D, Wickham H (2013) ggmap: Spatial visualization with ggplot2. *The R Journal* **5**, 144-161.

Kalinowski S (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity* **106**, 625-632.

Kemler M, Goker M, Oberwinkler F, Begerow D (2006) Implications of molecular characters for the phylogeny of the Microbotryaceae (Basidiomycota : Urediniomycetes). *BMC Evolutionary Biology* **6**, 35.

Le Gac M, Hood ME, Fournier E, Giraud T (2007a) Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. *Evolution* **61**, 15–26.

Le Gac M, Hood ME, Giraud T (2007b) Evolution of reproductive isolation within a parasitic fungal complex. *Evolution* **61**, 1781–1787.

Lindman HR (1974) *Analysis of variance in complex experimental designs* W. H. Freeman & Co, San Francisco.

López-Villavicencio M, Courjol F, Gibson A*, et al.* (2011) Competition, cooperation among kin and virulence in multiple infections. *Evolution* **65**, 1357-1366.

For Review Only

759 Lopez-Villavicencio M, Jonot O, Coantic A, *et al.* (2007) Multiple infections by the anther
760 smut pathogen are frequent and involve related strains. *PLoS Pathogens* **3**, e176.

761 Lutz M, Goker M, Piatek M, *et al.* (2005) Anther smuts of Caryophyllaceae: molecular
762 characters indicate host-dependent species delimitation. *Mycological Progress* **4**, 225-
763 238.

764 Lutz M, Pia Tek M, Kemler M, Chlebicki A, Oberwinckler F (2008) Anther smuts of
765 Caryophyllaceae: Molecular analyses reveal further new species. *Mycological*
766 *Research* **112** 1280-1296.

767 Mailleret L, Castel M, Montarry J, Hamelin FM (2012) From elaborate to compact seasonal
768 plant epidemic models and back: is competitive exclusion in the details? *Theoretical*
769 *Ecology* **5**, 311-324.

770 Marsden-Jones EM, Turrill WB (1957) *The Bladder Campions (Silene maritima and S.*
771 *vulgaris)* Ray Society, London.

772 Mboup M, Bahri B, Leconte M, *et al.* (2012) Genetic structure and local adaptation of
773 European wheat yellow rust populations: the role of temperature-specific adaptation.
774 *Evolutionary Applications* **5**, 341-352.

775 McQueen PG, McKenzie FE (2004a) Age-structured red blood cell susceptibility and the
776 dynamics of malaria infections. *Proceedings of the National Academy of Sciences of*
777 *the United States of America* **101**, 9161-9166.

778 McQueen PG, McKenzie FE (2004b) Theoretical studies of dual-species *Plasmodium*
779 infections: effects of competition for red blood cells. *Molecular Biology of the Cell* **15**,
780 465A-465A.

781 Neophytou C (2014) Bayesian clustering analyses for genetic assignment and study of
782 hybridization in oaks: effects of asymmetric phylogenies and asymmetric sampling
783 schemes. *Tree Genetics & Genomes* **10**, 273-285.

784 Petit E, Silver C, Cornille A, *et al.* (2017) Co-occurrence and hybridization of anther-smut
785 pathogens specialized on *Dianthus* hosts. *Mol Ecol.* **26**, 1877-1890.

786 Prentice HC, Andersson S, Mansby E (2011) Mosaic variation in allozyme and plastid DNA
787 markers in the European ranges of *Silene vulgaris* and its partially sympatric relative
788 *S. uniflora* (Caryophyllaceae). *Botanical Journal of the Linnean Society* **166**, 127-148.

789 Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using
790 multilocus genotype data. *Genetics* **155**, 945-959.

791 Refrégier G, Hood ME, Giraud T (2010) No evidence of reproductive character displacement
792 between two sister fungal species causing anther smut disease in *Silene*. *Int J Plant*
793 *Sciences* **171**, 847-859.

794 Refrégier G, Le Gac M, Jabbour F, *et al.* (2008) Cophylogeny of the anther smut fungi and
795 their Caryophyllaceous hosts: Prevalence of host shifts and Importance of delimiting
796 parasite species. *BMC Evol Biol* **8**, 100.

797 SAS Institute Inc. (2013). SAS system for Windows, Cary, NC, USA.

798 Smith PA, Lutz M, Ziegler R, Piątek M (2017) Anther smuts of *Silene acaulis* and *S. uniflora*
799 in the Outer Hebrides, including an assessment of ITS genotypes of *Microbotryum*
800 *silenes-acaulis*. *IMA Fungus* **8**, 107–116

For Review Only

Snounou G, White NJ (2004) hooThe co-existence of *Plasmodium*: sidelights from *falciparum* and *vivax* malaria in Thailand. *Trends in Parasitology* **20**, 333-339.

Thrall PH, Antonovics J (1997) Polymorphism in sexual vs. non-sexual transmission. *Proceedings of the Royal Society of London Series B* **264**, 581-587.

Vacher C, Vile D, Helion E, Piou D, Desprez-Loustau ML (2008) Distribution of parasitic fungal species richness: influence of climate versus host species diversity. *Diversity and Distributions* **14**, 786-798.

Vánky K (1994) *European smut fungi* Gustav Fischer Verlag, Stuttgart.

Vercken E, Fontaine MC, Gladieux P*, et al.* (2010) Glacial refugia in pathogens: European genetic structure of anther smut pathogens on *Silene latifolia* and *Silene dioica*. *PloS Pathogens* **6**, e1001229.

White TJ, Bruns TD, Lee S, Taylor J (1990) Analysis of phylogenetic relationships by amplification and direct sequencing of ribosomal RNA genes. In: *PCR Protocols: a Guide to Methods and Applications* (eds. Innis MA, Gelfand DH, Sninsky JJ, White TJ), pp. 315–322. Academic Press, New York.

For Review Only

818 **Table 1.** Summary statistics of genetic variation and estimates of selfing rates in the identified
819 *Microbotryum* species and clusters.

Species/Cluster	N	H_d	A_r	H_O/H_E	F_{IS}	s
Microbotryum silenes-inflatae						
Green	79	0.346	1.72	0.24/0.33	0.355	0.52
Yellow	64	0.468	2.01	0.18/0.44	0.577	0.73
Pink	75	0.243	1.49	0.20/0.23	0.199	0.33
Orange	40	0.178	1.35	0.10/0.17	0.387	0.56
Microbotryum violaceo-irregulare						
Species-wide	7	0.594	2.07	0.08/0.44	0.852	0.92
Microbotryum lagerheimii						
Blue	24	0.492	2.03	0.12/0.45	0.760	0.86
Light blue	12	0.491	1.98	0.13/0.43	0.662	0.80

820 N, number of pure representatives of the species/clusters, as determined using Bayesian
821 clustering analyses; H_d, gene diversity; A_r, allelic richness, corrected for sample size
822 using rarefaction (standardized sample size of two diploid individuals); H_O/H_E,
823 observed/expected heterozygosity; F_{IS}, fixation index; s, selfing rate estimated as
824 $2F_{IS}/(1+F_{IS})$.

825
826
827
828

For Review Only

Figure Legends

Figure 1. Maximum-likelihood tree showing the phylogenetic placement of *Microbotryum lagerheimii* (HQ832090), *M. violaceo-irregulare* (AY588104), and *M. silenes-inflatae* (JN223404), all parasitizing *Silene vulgaris*, among other *Microbotryum* species, based on internal transcribed spacer (ITS) sequences. *Microbotryum* from *Polygonum bistorta* was used as an outgroup. Bootstraps are indicated. Pictures show differences in teliospore color between the dark purple of *M. silenes-inflatae* and *M. violaceo-irregulare* versus the lighter lavender of *M. lagerheimii*, as well as differences in teliospore ornamentation (verrucose in *M. violaceo-irregulare*, fully reticulated in *M. silenes-inflatae* and *M. lagerheimii*) under 1000x light-microscope magnification.

Figure 2. Proportions of ancestry in K (from 2 to 10) clusters of *Microbotryum spp.* genotypes inferred with the STRUCTURE program. Each genotype is represented by a vertical bar, partitioned into K segments representing the amount of ancestry of its genome in K clusters. When several clustering solutions (‘‘modes’’) were found within replicate runs, only the major mode is shown with its corresponding proportion of runs. IS: Iceland, NL: Netherlands, UK: United Kingdom, CH: Switzerland, IT: Italy, F: France. For each region, genotypes are ordered by sampling elevation (represented at bottom). See Supplementary Figure S3 for a sorting by membership coefficient.

Figure 3. Proportions of ancestry in *K* (from 2 to 5) clusters of the *Microbotryum violaceo-irregulare* and *M. lagerheimii* genotypes inferred with the STRUCTURE program, using a sub-dataset excluding *M. silenes-inflatae*. In the sub-dataset we kept only the genotypes assigned to the blue cluster at *K*=5 in the output of the Structure analysis on the whole dataset (Figure 4). A single clustering solution (‘‘mode’’) was found among replicate runs. Each genotype is represented by a vertical bar, partitioned into *K* segments representing the amount

For Review Only

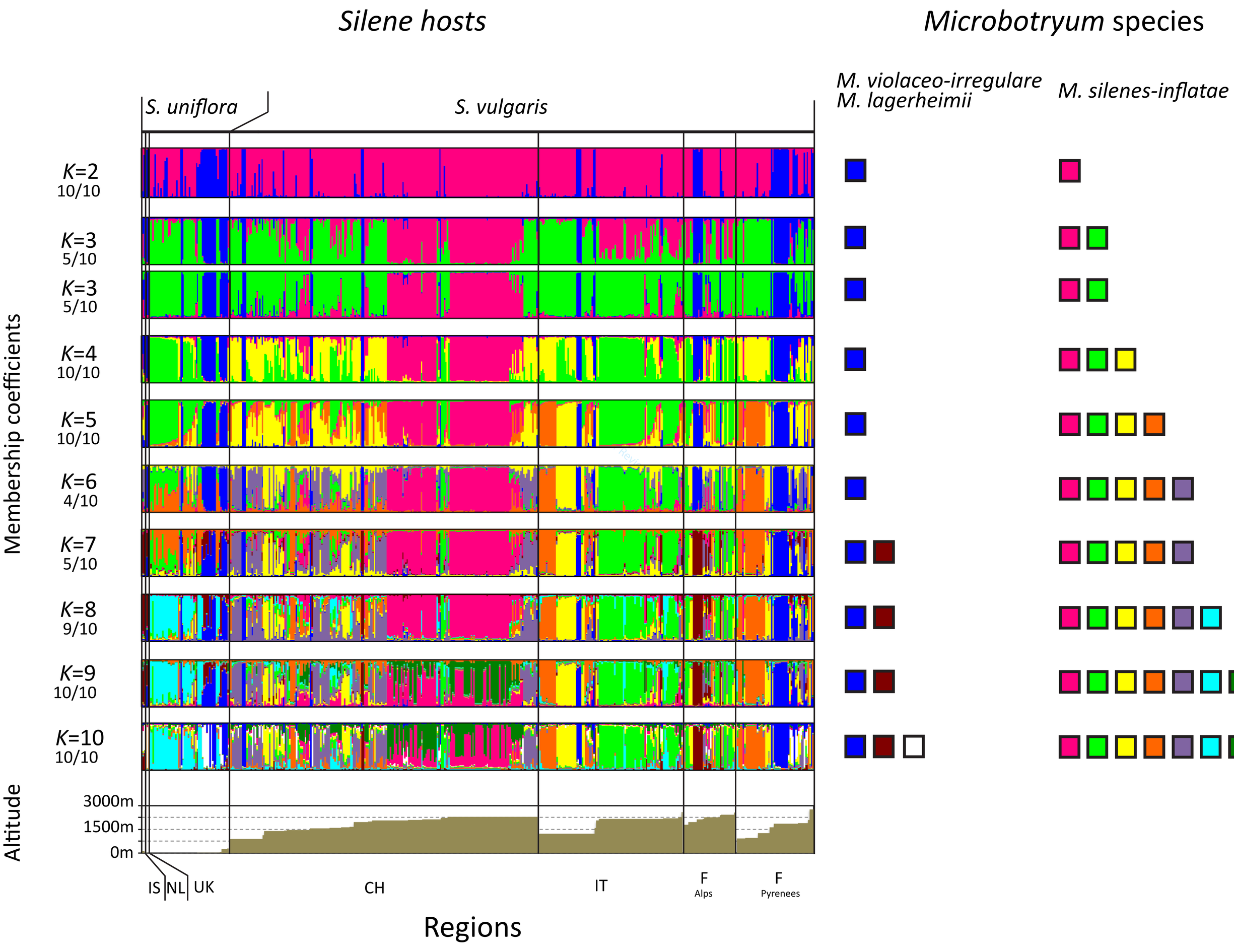
of ancestry of its genome in *K* clusters. NL: Netherlands, UK: United Kingdom, CH: Switzerland, IT: Italy, F: France. For each region, samples are ordered by sampling elevation (represented at bottom; see Figure S5 for a sorting by assignment coefficient within regions).

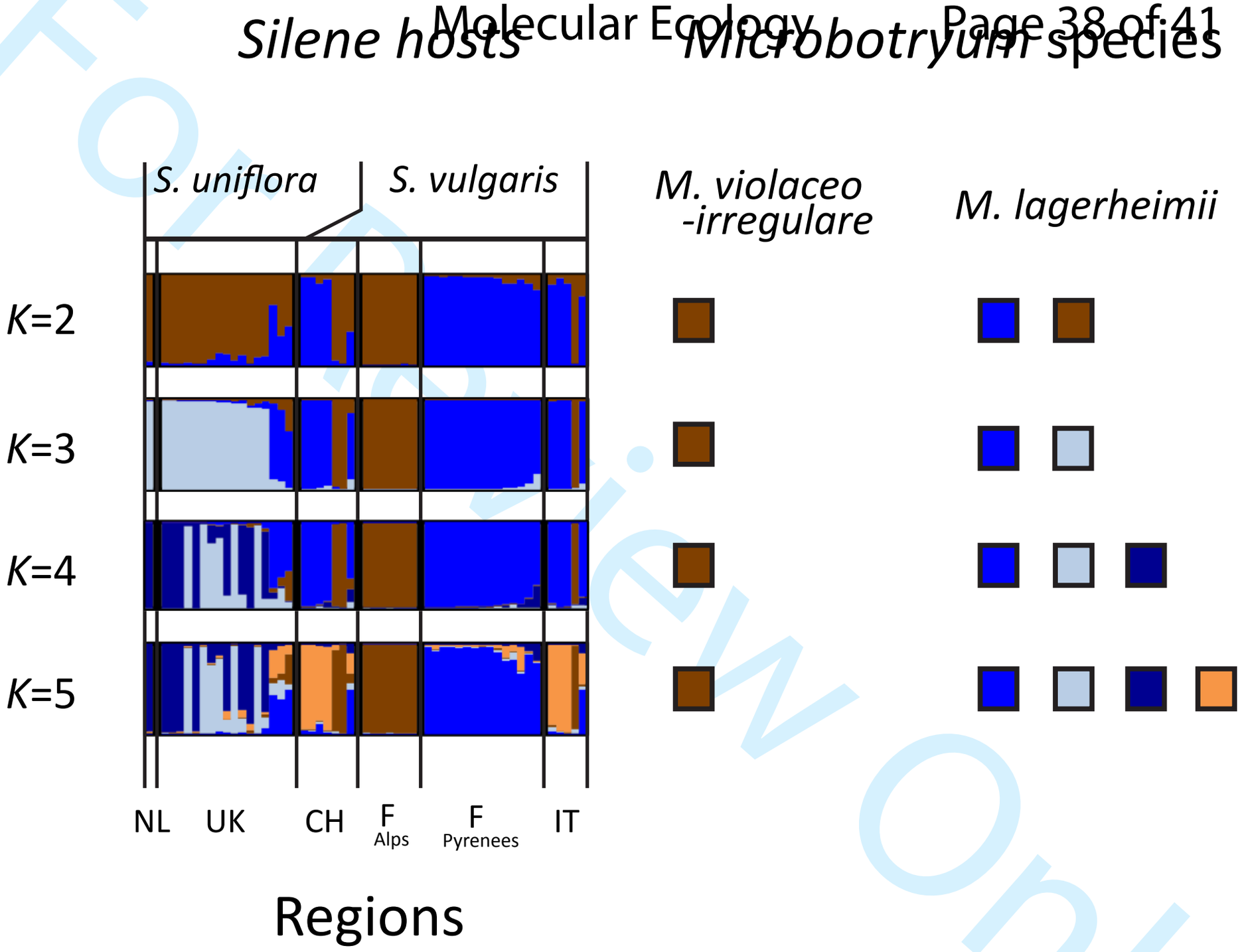
Figure 4. Principal component analysis (PCA) on multilocus microsatellite genotypes for the dataset including the three *Microbotryum* species (A: principal component 1 vs principal component 2; B: principal component 1 vs principal component 3) for the sub-dataset including only *M. lagerheimii* and *M. violaceo-irregulare* (C: principal component 1 vs principal component 2; D: principal component 1 vs principal component 3). Scatterplots for the first three principal components are shown using a colour labelling of genotypes defined according to the assignment of multilocus genotypes to three species and six clusters within species using Bayesian clustering analyses. Colors indicate clusters as in Figures 2 and 3. Black outlines indicate the samples for which the ITS region has been sequenced.

Figure 5. Distribution maps of the different genetic clusters and species of anther-smut *Microbotryum* fungi on *Silene vulgaris* and *S. uniflora*, and maps of their co-occurrence. Pie charts are proportional to the number of samples. Colors correspond to those used in other figures.

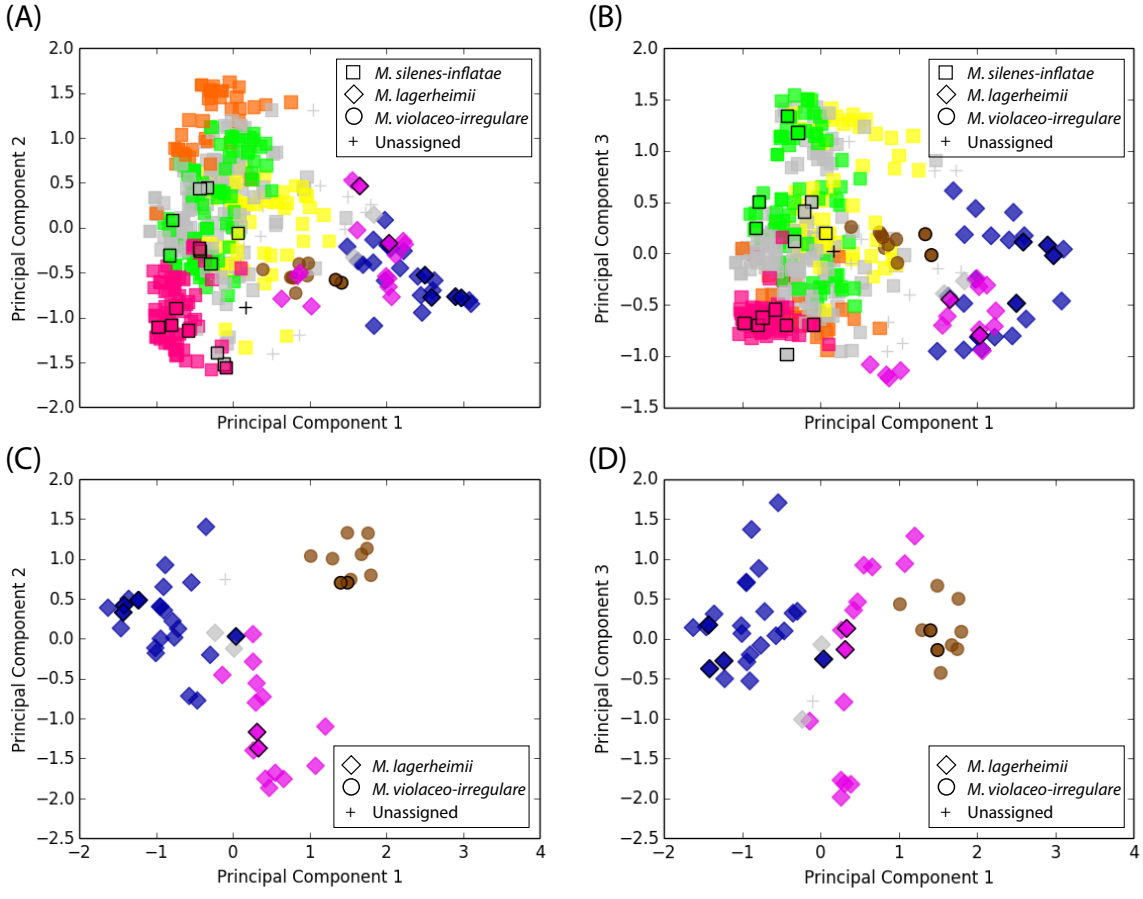
Figure 6. Violin plots showing elevational distribution of species and clusters within species. All samples belonging to *Microbotryum silenes-inflatae* at K=5 (including all genotypes with low assignment power to the green, orange, pink and yellow *M. silenes-inflatae* clusters) are included in the leftmost *M. silenes-inflatae* column. Otherwise, only samples with high-confidence cluster assignment are shown. Samples belonging to *M. lagerheimii* and *M. violaceo-irregulare* were determined using the K=3 output from the sub-dataset excluding *M. silenes-inflatae*.

For Review Only

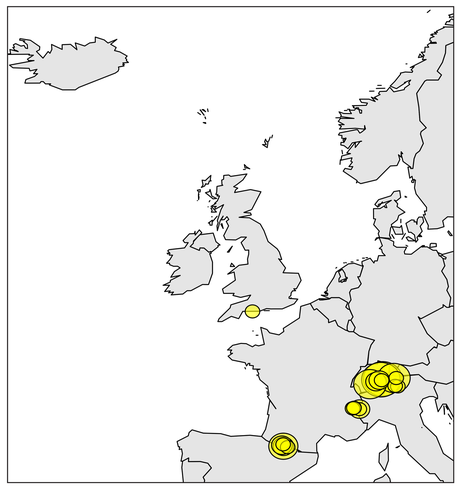

Silene hosts
Microbotryum species
S. uniflora
S. vulgaris
M. violaceo-irregulare
M. lagerheimii
M. silenes-inflatae
K=2
10/10
K=3
5/10
K=3
5/10
K=4
10/10
K=5
10/10
K=6
4/10
K=7
5/10
K=8
9/10
K=9
10/10
K=10
10/10
Membership coefficients
Altitude
3000m
1500m
0m
IS
NL
UK
CH
IT
F
Alps
F
Pyrenees
Regions



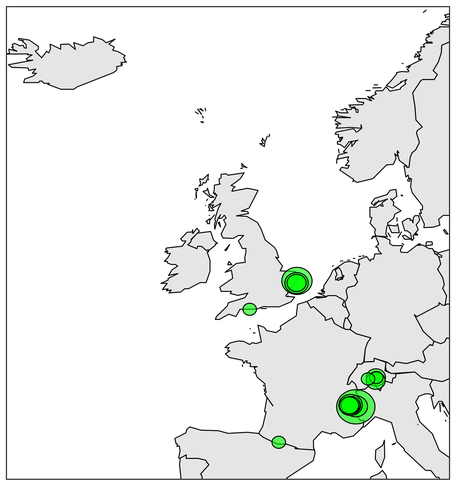
For Review Only


(A)
(B)
(C)
(D)
M. silenes-inflatae
M. lagerheimii
M. violaceo-irregulare
Unassigned
Principal Component 1
Principal Component 2
Principal Component 3

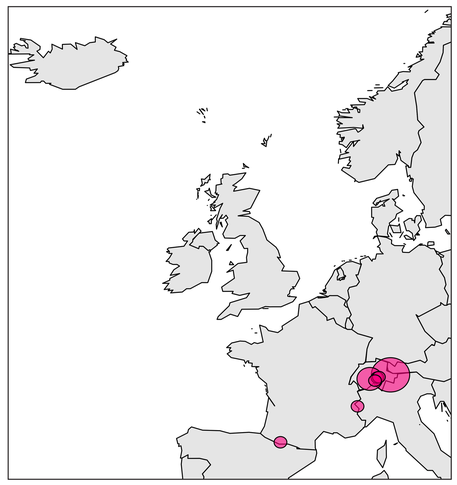
(A) Yellow cluster of *M. silenes-inflatae*



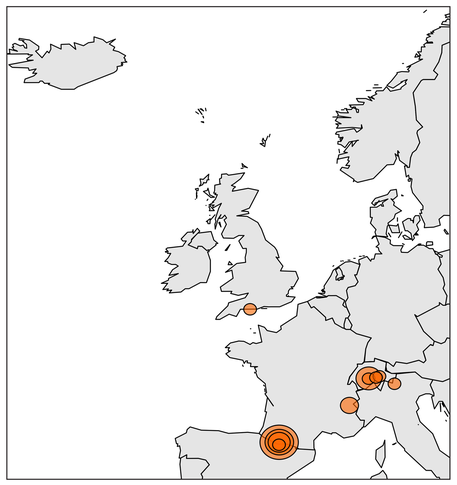
(B) Green cluster of *M. silenes-inflatae*



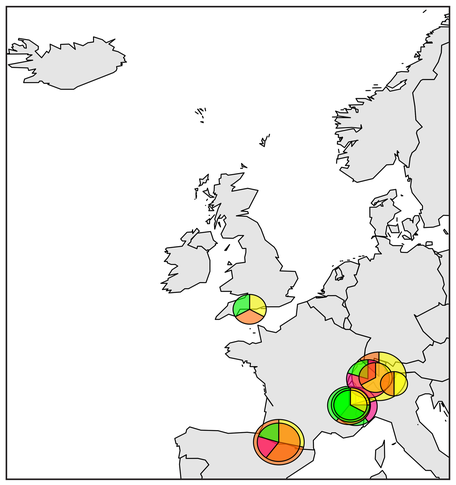
(C) Pink cluster of *M. silenes-inflatae*



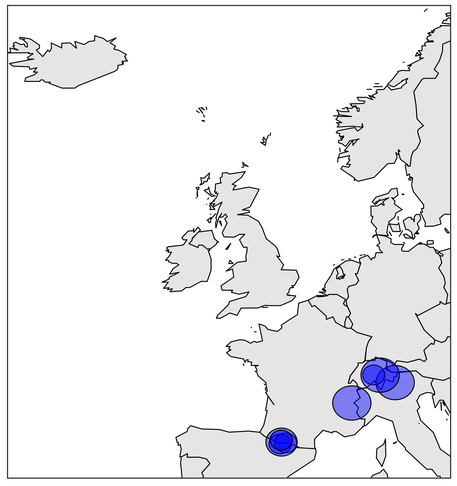
(D) Orange cluster of *M. silenes-inflatae*



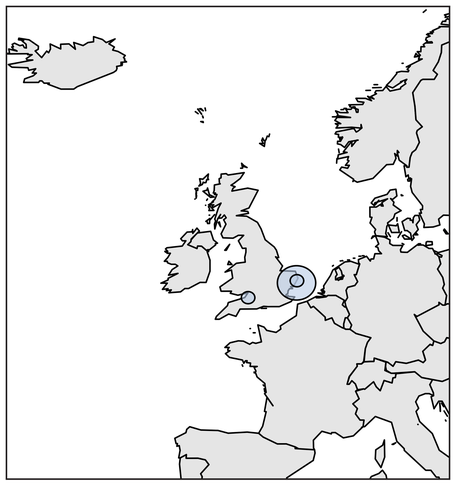
(E) Coexistence of *M. silenes-inflatae* clusters



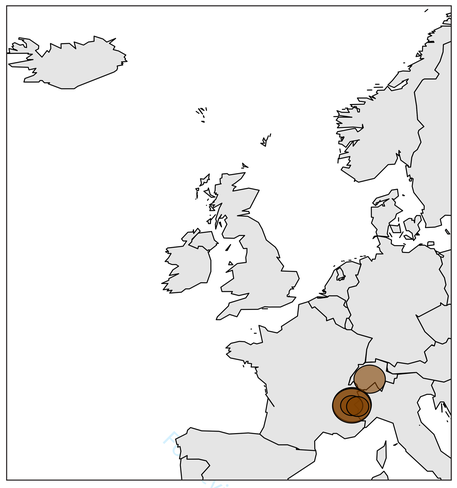
(E) Blue cluster of *M. lagerheimii*



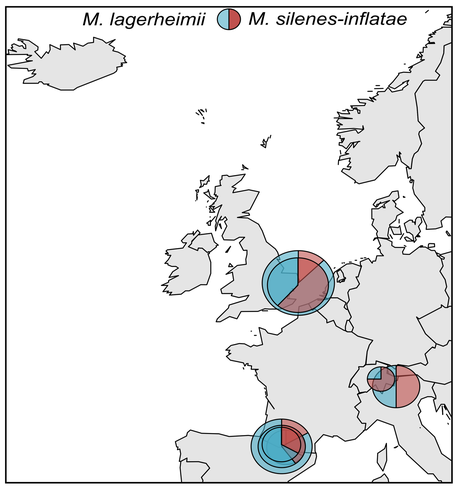
(F) Light blue cluster of *M. lagerheimii*



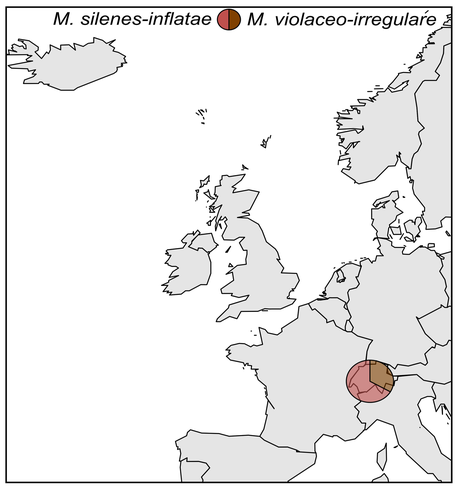
(G) *M. violaceo-irregulare*



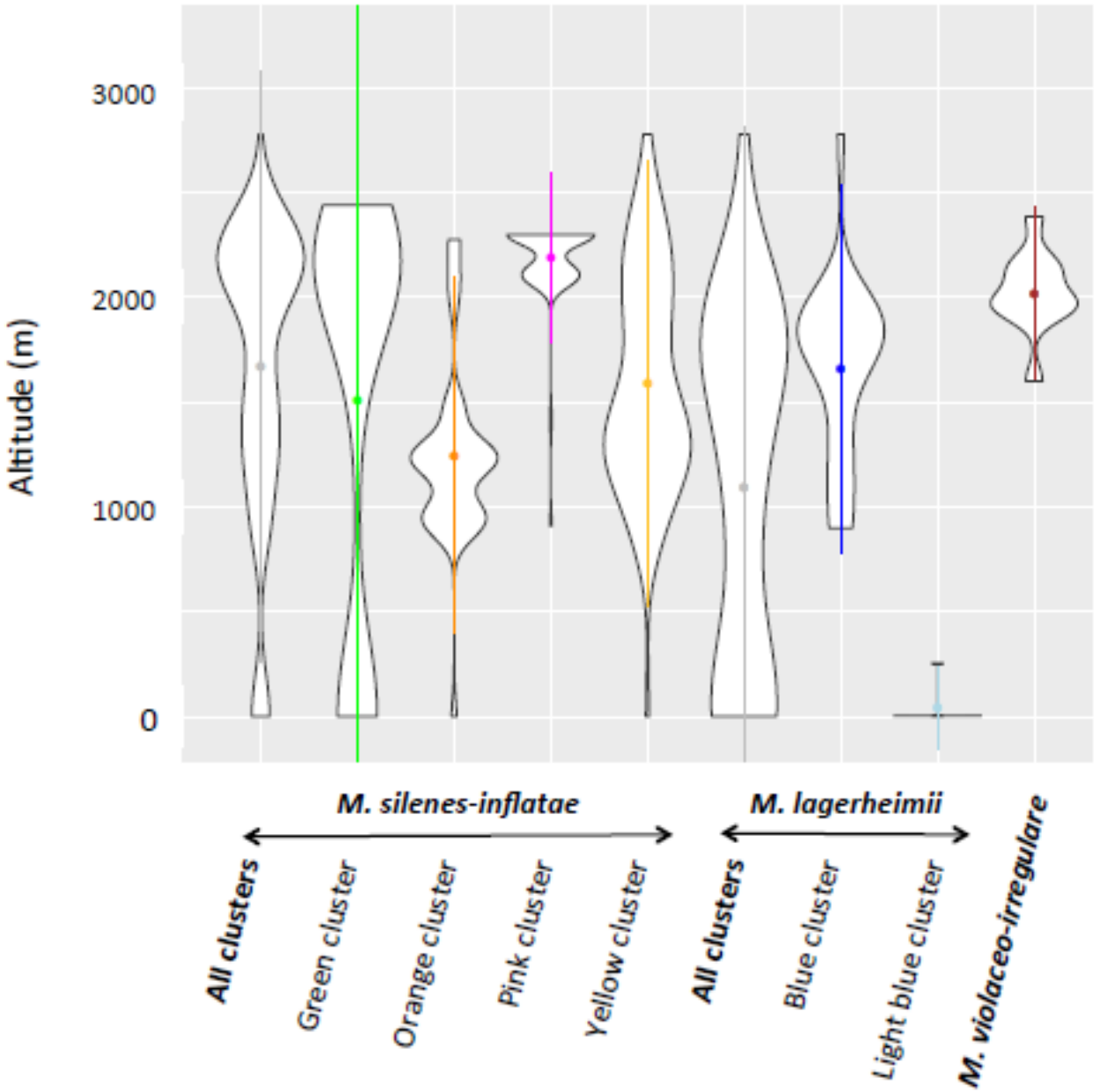
(H) Coexistence of *M. silenes-inflatae* and *M. lagerheimii*



(I) Coexistence of *M. silenes-inflatae* and *M. violaceo-irregulare*



Review Only


Altitude (m)
3000
2000
1000
0
M. silenes-inflatae
M. lagerheimii
All clusters
Green cluster
Orange cluster
Pink cluster
Yellow cluster
All clusters
Blue cluster
Light blue cluster
M. violaceo-irregulare