doi:10.1111/j.1558-5646.2007.00002.x

PHYLOGENETIC EVIDENCE OF HOST-SPECIFIC CRYPTIC SPECIES IN THE ANTHER SMUT FUNGUS

Mickael Le Gac,[1,2,3] Michael E. Hood,[4,5] Elisabeth Fournier,[6,7] and Tatiana Giraud[1,8]

[1]*Ecologie, Systématique et Evolution, UMR 8079 CNRS-UPS-ENGREF, Bâtiment 360, Université Paris-Sud, F-91405 Orsay, France*

[2]*E-mail: legac@zoology.ubc.ca*

[4]*Department of Biology, Amherst College, Amherst, Massachusetts 001002*

[5]*E-mail: MHood@amherst.edu*

[6]*PMDV, INRA, Route de Saint-Cyr, F-78026 Versailles cedex, France*

[7]*E-mail: efournie@versailles.inra.fr*

[8]*E-mail: tatiana.giraud@u-psud.fr*

Received December 13, 2005
Accepted September 22, 2006

Cryptic structure of species complexes confounds an accurate accounting of biological diversity in natural systems. Also, cryptic sibling species often become specialized to different ecological conditions, for instance, with host specialization by cryptic parasite species. The fungus *Microbotryum violaceum* causes anther smut disease in plants of Caryophyllaceae, and the degree of specialization and gene flow between strains on different hosts have been controversial in the literature. We conducted molecular phylogenetic analyses on *M. violaceum* from 23 host species and different geographic origins using three single-copy nuclear genes (β-tub, γ-tub, and Ef1α). Congruence between the phylogenies identified several lineages that evolved independently for a long time. The lineages had overlapping geographic ranges but were highly specialized on different hosts. These results thus suggest that *M. violaceum* is a complex of highly specialized sibling species. Two incongruencies between the individual gene phylogenies and one intragene recombination event were detected at basal nodes, suggesting ancient introgression events or speciation events via hybridizations. However, incongruencies and recombination were not detected among terminal branches, indicating that the potentials for cross-infection and experimental hybridization are often not sufficient to suggest that introgressions would likely persist in nature.

KEY WORDS: **Adaptation, cryptic species, multiple gene genealogies, sibling species, specialization, sympatry, *Ustilago violacea*.**

Species complexes are composed of genetically isolated lineages that are not distinguishable on the basis of purely morphological criteria. Such difficulties have been encountered in almost all taxonomic groups, even the most studied birds and mammals (e.g., Burbidge et al. 2003; Ravaoarimanana et al. 2004), but especially in groups with fewer obvious taxonomic criteria like fungi (e.g., Geiser et al. 1998; Koufopanou et al. 2001; Pringle et al. 2005). Detecting the cryptic structure of species complexes is essential for an accurate accounting of the biological diversity in natural systems, especially as the sibling species often have evolved different specializations. Investigating cryptic genetic structure in relation to specialization is particularly relevant in parasites. For example, many plant enemies considered as broad generalists have recently been recognized as complexes of sibling species, specialized on

[3]Present address: Department of Zoology, University of British Columbia, Vancouver, British Columbia, V6T1Z4, Canada.

different hosts; see for instance reviews by Burnett (2003) for fungi and by Dres and Mallet (2002) for insects.

Several tree-based methods are used to delimit sibling species (Sites and Marshall 2004). Multiple-gene phylogenetic approaches have been particularly useful within the fungal kingdom (e.g., Koufopanou et al. 1997; Geiser et al. 1998; O'Donnell et al. 2000; 2001, Dettman et al. 2003a; Kasuga et al. 2003; Staats et al. 2005; Pringle et al. 2005; Fournier et al. 2005), which use the phylogenetic concordance of multiple unlinked genes to indicate a lack of genetic exchange and the evolutionary independence of lineages. Thus species boundaries can be determined in spite of a lack of other taxonomic characters or experimental crossing possibilities (Avise and Wollenberg 1997; Taylor et al. 2000; Dettman et al. 2003a,b). This approach is conservative, in that ancestral polymorphism may persist in species having recently diverged, preventing the detection of barriers to gene flow (Nichols 2001). In fungi, the species boundaries recognized by this approach have been found to be in good agreement with those identified by mating tests, although the use of mating compatibility can underestimate the number of species compared to the phylogenetic approach (Dettman et al. 2003a,b). A theoretical study has shown that this discrepancy can be caused by specificities in the life cycles of some fungi (Giraud et al. 2006).

Here we use multiple gene phylogenies to investigate the genetic structure within the fungal pathogen *Microbotryum violaceum* (Basidiomycota). The existence of "host races" within *M. violaceum* has been debated since the early 20th century, that is, groups of strains specialized on different hosts with restricted gene flow among them. Diverse approaches have been used to address this issue, including inoculation and hybridization experiments (Goldschmidt 1928; Biere and Honders 1996; Shykoff et al. 1999; Van Putten et al. 2003), morphological variation (Garber et al. 1978), as well as molecular variation for karyotypes (Perlin 1996; Perlin et al. 1997), microsatellite allele frequencies (Shykoff et al. 1999; Bucheli et al. 2000, 2001; Giraud et al. 2002; Van Putten et al. 2005), and DNA sequences (Freeman et al. 2002). Some studies provide evidence for host specialization of *M. violaceum* (Goldschmidt 1928) with limited gene flow among strains on different hosts even when in sympatry (Bucheli et al. 2000, 2001; Van Putten et al. 2005). In contrast, some experimental data highlighted the possibility of cross-infection and the production of viable hybrids among strains from different hosts (Biere and Honders 1996; Shykoff et al. 1999; Van Putten et al. 2005).

Using multiple gene phylogenies, we aimed to answer the following questions: (1) Are there genetically isolated lineages within *M. violaceum* that have not been distinguished based on morphological criteria? (2) Are these lineages specialized on single host species, or are they each found to parasitize several hosts? and (3) Is there a geographical segregation of the different evolutionary lineages?

Material and Methods

STUDY SYSTEM

Microbotryum violaceum causes a sexually transmitted disease, anther smut, of plants in the Caryophyllaceae (Thrall et al. 1993). Diploid teliospores are produced in the anthers of infected plants, replacing the pollen, and female structures are aborted. Deposited by pollinators on a new host, teliospores germinate and undergo meiosis. Sexual conjugation between two haploid cells of opposite mating types is required prior to infection, but selfing is frequent (Delmotte et al. 1999; Giraud 2004; Giraud et al. 2005). *Microbotryum violaceum* has been studied as a model in several fields of genetics, ecology, and evolutionary biology (Garber and Ruddat 2002; Martinez-Espinoza et al. 2002).

FUNGAL ISOLATES, STORAGE, AND CULTURE

For DNA sequencing, 56 strains of *M. violaceum* from 23 host species in Western Europe and America were used (Table 1). Strains were collected between 2000 and 2004 and stored at 4°C under desiccation. Teliospores were grown for ca. one week at 23°C on sterile GMB2 media, that is, GMB1 media (Thomas et al. 2003) modified as follows: instead of glucose 4 g/l and yeast extract 10 g/l, we added glucose 10 g/l, yeast extract 1.5 g/l, and malt extract 3 g/l. Resulting cultures of haploid sporidia were diluted with water and then replated, yielding colonies derived from separate haploid sporidia. These single-sporidial colonies were isolated and cultured for DNA extraction.

DNA EXTRACTION, AMPLIFICATION, AND SEQUENCING

DNA was extracted from single-sporidial colonies using the Chelex (Biorad, Marne-la-coquette, France) protocol (Bucheli et al. 2001). Approximately 600 bp of the β-tubulin (β-tub), 700 bp of the γ-tubulin (γ-tub), and 1000 bp of the Elongation factor 1 α (Ef1α) were amplified by polymerase chain reaction (PCR) using the primers Tub 5-IN/Tub 3-IN, T1/T2 (Freeman et al. 2002) and Ef1u1f (ACGGTCTGACGCATGT CAC)/Ef1u1r (CAAGAACATGATCACTGGTACCTC), respectively. For samples in which the Ef1α gene could not be amplified, internal primers Ef1m4f (GAAGGACTCGACGCA CATC) and Ef1m4r (ACGCTCTCCTCGCCTTCA) were used, yielding 800 bp.

PCR amplifications were performed using standard techniques (available upon request). PCR products were purified using PEG (Sigma-Aldrich, Lyon, France) and sequenced on ABI 310 (Applied Biosystem, Courtabœuf, France) or CEQ 8000 (Beckman Coulter, Villepinte, France) using dye-terminator

Table 1. Isolates of *Microbotryum violaceum* used in this study: name (including the abbreviated name of the host species), host species, and geographic area of collection.

Sample name	Host species	Region
Dalpinus1	*Dianthus alpinus*	The Alps, Switzerland
Dcarthusianorum_7022	*Dianthus carthusianorum*	The Alps, Switzerland
Dcarthusianorum_7042	*Dianthus carthusianorum*	The Alps, Switzerland
Dcarthusianorum_7515	*Dianthus carthusianorum*	The Alps, Switzerland
Dcarthusianorum_30901	*Dianthus carthusianorum*	The Pyrenees, France
Dcarthusianorum_30902	*Dianthus carthusianorum*	The Pyrenees, France
Dcarthusianorum_31602	*Dianthus carthusianorum*	The Alps, Switzerland
Dgratianopolitanus	*Dianthus gratianopolitanus*	Jura, Switzerland
Dmonspessulanus_12919	*Dianthus monspessulanus*	The Pyrenees, France
Dmonspessulanus_12920	*Dianthus monspessulanus*	The Pyrenees, France
Dneglectus_1	*Dianthus neglectus*	The Alps, Italy
Dneglectus_2	*Dianthus neglectus*	The Alps, Italy
Dsuperbus_8718	*Dianthus superbus*	The Alps, Switzerland
Dsylvestris_6733	*Dianthus sylvestris*	The Alps, Switzerland
Dsylvestris_6740	*Dianthus sylvestris*	The Alps, Switzerland
Dsylvestris_9111	*Dianthus sylvestris*	The Alps, Switzerland
Dsylvestris_9119	*Dianthus sylvestris*	The Alps, Switzerland
Dsylvestris_1	*Dianthus sylvestris*	The Alps, Switzerland
Dsylvestris_us1	*Dianthus sylvestris*	The Alps, Switzerland
Grepens_137	*Gypsophila repens*	The Alps, Italy
Grepens_6	*Gypsophila repens*	The Alps, Italy
Lflos-cuculi_9203	*Lychnis flos-cuculi*	The Alps, Switzerland
Lflos-cuculi_9205	*Lychnis flos-cuculi*	The Alps, Switzerland
Lflos-cuculi_10934	*Lychnis flos-cuculi*	The Alps, Switzerland
Saccaulis_b	*Silene acaulis*	The Alps, Switzerland
Saccaulis_31401	*Silene acaulis*	The Pyrenees, France
Saccaulis_31402	*Silene acaulis*	The Pyrenees, France
Saccaulis_1	*Silene acaulis*	The Alps, Switzerland
Saccaulis_2	*Silene acaulis*	The Alps, Switzerland
Scaroliniana_1	*Silene caroliniana*	USA
Scaroliniana_2	*Silene caroliniana*	USA
Sdioica_7212	*Silene dioica*	The Alps, Switzerland
Sdioica_7237	*Silene dioica*	The Alps, Switzerland
Sdioica_b	*Silene dioica*	Brittany, France
Slatifolia_4001	*Silene latifolia*	Paris region, France
Slatifolia_4106	*Silene latifolia*	Paris region, France
Slatifolia_10002	*Silene latifolia*	The Alps, Switzerland
Slatifolia_10006	*Silene latifolia*	The Alps, Switzerland
Slatifolia_2	*Silene latifolia*	USA
Slemonii	*Silene lemonii*	USA
Smaritima_1	*Silene maritima*	Somerset, U.K.
Smaritima_3	*Silene maritima*	Somerset, U.K.
Snutans_7901	*Silene nutans*	The Alps, Switzerland
Snutans_8742	*Silene nutans*	The Alps, Switzerland
Snutans_lb	*Silene nutans*	Jura, France
Socymoides	*Silene ocymoides*	The Alps, Switzerland
Sofficinalis	*Silene officinalis*	The Alps, Switzerland
Sottites	*Silene ottites*	The Alps, Switzerland
Sparryi	*Silene parryi*	USA

(continued)

Table 1. *(continued)*

Sample name	Host species	Region
Srupestris_10502	*Silene rupestris*	The Alps, Switzerland
Svulgaris_7806	*Silene vulgaris*	The Alps, Switzerland
Svulgaris_7807	*Silene vulgaris*	The Alps, Switzerland
Svulgaris_7903	*Silene vulgaris*	The Alps, Switzerland
Svulgaris_7913	*Silene vulgaris*	The Alps, Switzerland
Svulgaris_30027	*Silene vulgaris*	The Pyrenees, France
Svulgaris_30030	*Silene vulgaris*	The Pyrenees, France
Ccaespitosa	*Calandrinia caespitosa*	Chile

chemistry. The γ-tub and Ef1α genes were sequenced in both directions, whereas the β-tub was sequenced using only the Tub 5-IN primer. The sequences are available in GenBank (accession numbers DQ074479–DQ074635).

PHYLOGENETIC ANALYSES

The three genes used can be considered as unlinked given that (1) they are located on different chromosomes in 2 Basidiomycetes having their genome sequenced (*Ustilago maydis* and *Cryptococcus neoformans*) and (2) the karyotype of *M. violaceum* has been shown to be highly polymorphic, even among strains from the same host (Hood 2002; Hood et al. 2003; Hood and Antonovics 2004), suggesting that few genes may have remained durably linked in *M. violaceum*.

Sequences were aligned in Bioedit v6.0.7 and corrected by hand when necessary. Regions with ambiguous alignments were excluded from all analyses. Phylogenetic trees were reconstructed by maximum parsimony (MP), neighbor joining (NJ), and Bayesian inference. MP and NJ analyses were performed using PAUP version 4.0b10 (Swofford 2003). MP analyses were performed using a heuristic search. For NJ analyses, the ModelTest v3.5 program (Posada and Crandall 1998) with Akaike information criterion (AIC) was used to select the models that best fit our data. Bootstrap confidence values were calculated for 1000 pseudoreplicates (Felsenstein 1985). Bayesian analyses were run using MrBayes version 3.0b5 (Ronquist and Huelsenbeck 2003). Each run consisted of four incrementally heated Markov chains run simultaneously, with heating value set to default (0.2). Priors were constrained according to the ModelTest results. Markov chains were initiated from a random tree and run for 500,000 generations. We used a 50% majority rule consensus tree to obtain the Bayesian posterior probabilities (BPP), considering the trees sampled after the likelihood scores had reached stationary. Details on the phylogenetic analyses are available upon request. Monophylies supported by bootstrap $\geq$ 75% and BPP $\geq$ 0.95 were considered as significant. A node was considered as strongly supported when significant using two of the reconstruction methods.

INTRAGENE RECOMBINATION

To detect intragene recombination we used two phylogenetic methods (RDP by Martin and Rybicki 2000 and Bootscanning by Salminen et al. 1995) and two substitution distribution methods (Maximum χ^2 by Maynard Smith 1992 and Maximum mismatch χ^2 by Posada and Crandall 2001). These four different methods were used as implemented in the software RDP2 (Martin et al. 2005), and details about the parameters used are available upon request. To minimize the risk of false positives, we only considered as reliable potential recombination events detected by more than one method (Posada 2002).

OUTGROUP

Sequences from different fungi were used to root the phylogenies. Sequences were chosen as the best hits to our sequences in GenBank. To root the β-tub gene tree, we used *Monascus eremophilus* (GenBank accession number AY498603), *Rhodotorula glutinis* (L47266), *Cercophora sparsa* (AY600253), and *Platygloea pustulata* (AY371532). To root the Ef1α gene tree, we used *Agaricostilbum hyphaenes* (AY879114), *Sporobolomyces syzygii* (AB127096), *Bannoa hahajimensis* (AB127093), *Puccinia graminis* (X73529), and *Rhodotorula mucilaginosa* (AF016239). The γ-tub gene of *M. violaceum* could not be aligned with any published fungal sequences, even those of *U. maydis*. To root the trees, we also sequenced a smut sample infecting a non-Caryophyllaceae host (*Calandrinia caespitosa*) from Chile. The β-tub (DQ074479) and Ef1α (DQ074536) sequences could be obtained in this smut sample but not γ-tub. Given the surprising position of the *C. caespitosa* sample in the trees, DNA extraction, amplifications, and sequencing were performed a second time independently to discard the possibility of a contamination.

CONGRUENCE BETWEEN GENE PHYLOGENIES AND CONCATENATED DATASETS

Congruence between gene phylogenies was estimated using the incongruence length difference (ILD) test (Farris et al. 1994) as implemented in PAUP, the approximately unbiased (AU) test

(Shimodaira 2002) as implemented in CONSEL (for assessing the confidence of phylogenetic free selection) (Shimodaira and Hasegawa 2001), and by visual inspection of topologies and statistical supports (Mason-Gamer and Kellogg 1996). The ILD test is known to be very conservative (Darlu and Lecointre 2002), so the null hypothesis of congruence was rejected only if $P < 0.001$. By visual inspections, incongruence between gene phylogenies was concluded when conflicting nodes were supported by significant statistical values; nodes were considered as congruent when identical and supported by significant statistical values in the phylogenies considered. AU tests were conducted for each incongruency detected by visual inspection, by comparing for each gene the likelihood of the MP topology obtained for the focal gene to the likelihood of the conflicting topology obtained with another gene. Likelihoods were obtained in PAUP using the sequence evolution model selected according to the ModelTest results.

A restricted dataset excluding the strains showing incongruent positions when comparing the single gene phylogenies was used to concatenate the three genes, in order to improve the power of detection of monophyly. Nodes that were neither significantly congruent nor incongruent when comparing the single gene phylogenies but that were found strongly supported in the concatenated analysis were considered as significantly congruent. Only statistical supports and topologies were considered in the combined analyses.

IDENTIFICATION OF INDEPENDENT EVOLUTIONARY UNITS AND THEIR GEOGRAPHICAL DISTRIBUTION

To detect independent evolutionary units within *M. violaceum*, we used the criterion of phylogenetic congruence between different gene phylogenies (Taylor et al. 2000; Dettman et al. 2003a). We thus considered a group of strains as an independent evolutionary lineage when (1) it was significantly supported as monophyletic in at least one gene phylogeny by two of the three reconstruction methods, and (2) this was not contradicted by the other gene phylogenies. In addition, we recorded whether the independent evolutionary lineages geographically overlapped, that is, whether they contained individuals geographically closer to individuals from other lineages than to individuals from its own lineage.

Results

PHYLOGENIES

Information on datasets and phylogenetic trees is shown in Table 2. For each gene, Bayesian, MP, and NJ consensus trees revealed the same relationships between the significantly supported clades. Therefore, only the Bayesian consensus trees of the Ef1α, β-tub, and γ-tub genes are shown in Figures 1, 2, and 3, respectively.

The ILD test indicated that the individual gene phylogenies were at the limit of the incongruence significance, with $P = 0.001$. In agreement, visual inspection of the nodes and supports showed that most of the phylogenetic relationships were congruent among the three genes, such that the same clades were significantly supported in the different gene trees, with a few exceptions outlined below. In particular, all phylogenies strongly supported four deep clades, named A, B, C, and D (Figs. 1, 2, and 3). However, statistical support was generally lower using the β-tub gene for all the reconstruction methods except Bayesian inference (Fig. 2), and with the γ-tub some nodes were better resolved than with the two other genes.

Visual inspection of topology and supports detected three conflicting nodes between the γ-tub phylogeny and the other gene phylogenies. First, the position of the clade containing fungal strains from *Silene acaulis* (clade D1) was different between the γ-tub phylogeny and the two other phylogenies (incongruency 1 on Figs. 1, 2, and 3). The AU tests were also significant for this node ($P = 0.01$, 0.003, and < 0.001 when enforcing the position of the clade D1 in the β-tub and Ef1α phylogenies as in the γ-tub phylogeny and in the γ-tub phylogeny as in the β-tub and Ef1α phylogenies, respectively). Second, within clade D, the γ-tub phylogeny strongly supported the *Silene vulgaris* and *Lychnis floscuculi* strains as monophyletic, whereas the Ef1α phylogeny strongly supported the *S. vulgaris* and *Silene nutans* strains as monophyletic (incongruency 2, Figs. 1 and 3). The AU tests

Table 2. For the different trees constructed, information on the sequence dataset (number of sequences, including the *Calandrinia* sample, number of aligned sites number of variable sites, and number of parsimony informative sites) and information on maximum parsimony (MP) trees (number of steps and retention index [RI]).

	Number of sequences	Total number of sites	Variable sites	Parsimony informative sites	Number of tree steps	RI
Ef1α	51	754	163 (22%)	111 (15%)	274	0.92
β-tub	57	356	54 (15%)	45 (13%)	119	0.85
γ-tub	48	636	147 (23%)	111 (17%)	107	0.96
Combined	28	1742	279 (16%)	183 (11%)	Not relevant	Not relevant

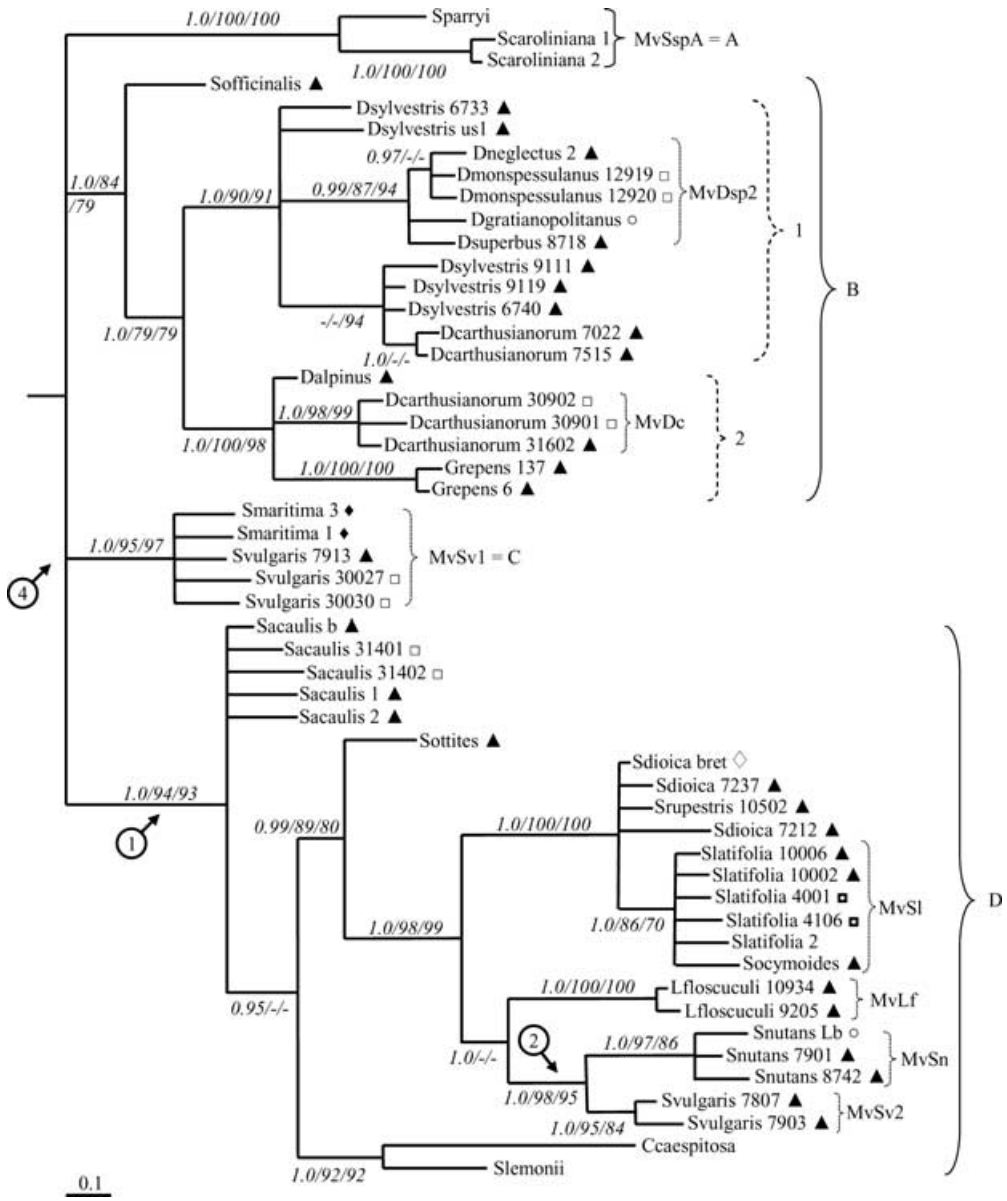

Figure 1. Bayesian 50% majority-rule consensus tree based on the Ef1α gene (mean lnL = −2968). Nodes not strongly supported are represented as unresolved. Statistical supports indicate Bayesian posterior probabilities (BPP)/maximum parsimony bootstraps/neighbor joining bootstraps. Only statistical supports higher than 0.9/70/70 are indicated. Numbered arrows indicate conflicting nodes between the different gene phylogenies or intragene recombination. Taxa labels correspond to the host plant on which fungal strains were collected. Brackets indicate clades and evolutionary units identified (see text). □ indicates strains from The Pyrenees, ▲ strains from The Alps, ◘ strains from Paris region, ◇ strains from Brittany, ○ strains from Jura, ♦ strains from Somerset, and no symbol indicates strains from America.

were also significant for this node ($P = 0.03$ and < 0.001 when enforcing the monophyly of the *S. vulgaris* and *S. nutans* strains in the γ-tub phylogeny and of the *S. vulgaris* and *L. floscuculi* strains in the Ef1α phylogeny, respectively). Third, the *Dianthus sylvestris* strain 1 and *D. sylvestris* strain 6740 were strongly supported as monophyletic in the γ-tub phylogeny, whereas in the β-tub phylogeny the *D. sylvestris* strain 6740 was not in the same strongly supported clade as the *D. sylvestris* strain 1 (incongruency 3, Figs. 2 and 3). The AU tests were also significant for this node ($P < 0.001$ and $P = 0.005$ when enforcing the *D. sylvestris* strain 1 and *D. sylvestris* strain 6740 as monophyletic in the β-tub phylogeny and as belonging to different clades in the γ-tub phylogeny, respectively).

When removing these strains showing incongruencies among individual phylogenies, the three gene tree topologies appeared visually identical, but for the significance of statistical support for

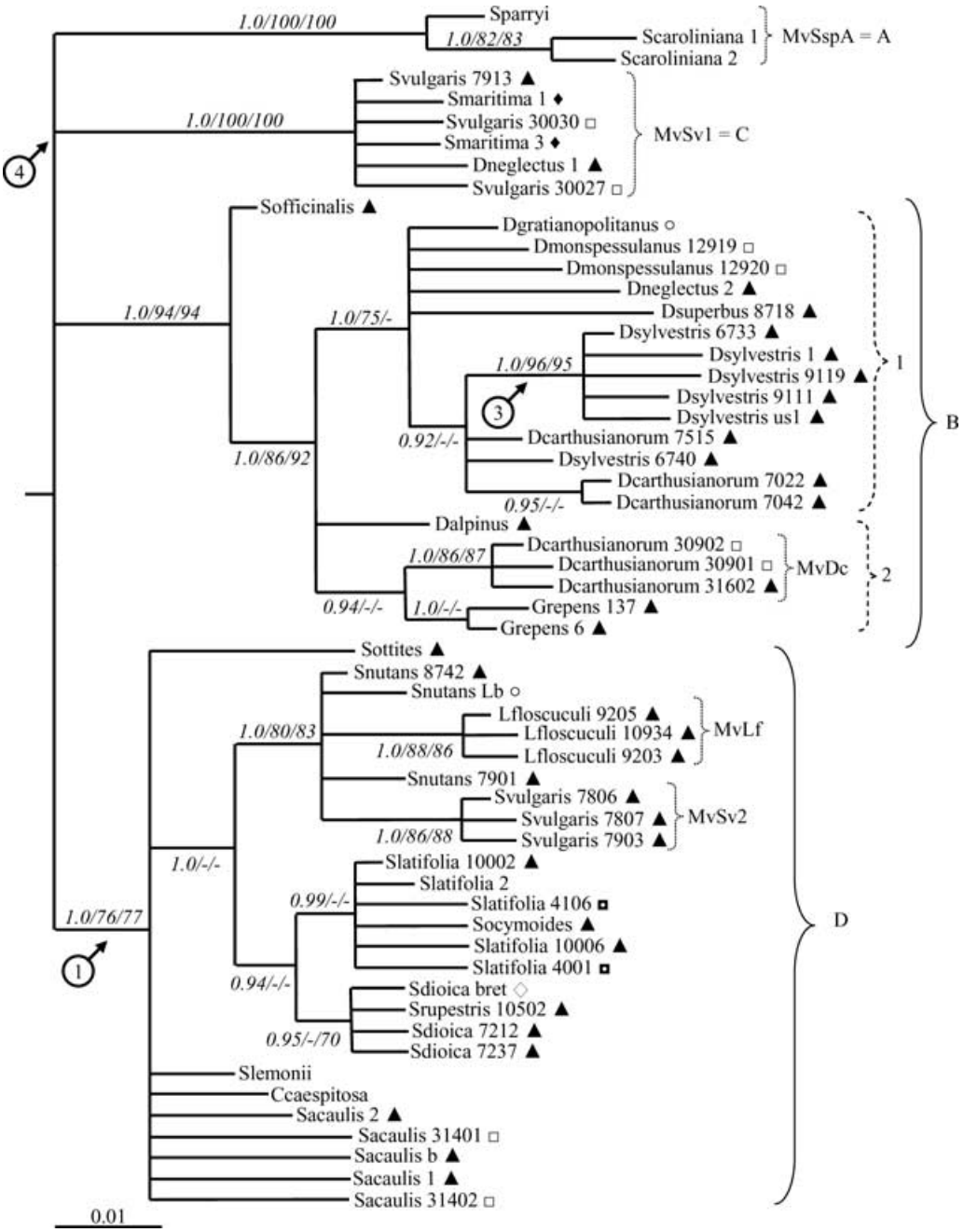

Figure 2. **Bayesian 50% majority-rule consensus tree based on the β-tub gene (mean lnL $= -1039$). See Figure 1 for the legend of the bootstraps, arrows, and taxa labels.**

various nodes. The ILD test confirmed that the gene phylogenies are then more congruent ($P = 0.06$, NS). The phylogenetic relationships inferred from the concatenated dataset were the same as for the individual gene trees, but with higher statistical supports (Fig. 4).

Phylogenetic analyses of the β-tub and Ef1α genes were first performed using several fungal sequences as outgroups; the γ-tub tree remained unrooted. The root positions found using the β-tub and Ef1α genes were identical (Figs. 1 and 2) and fully concordant with the results obtained by Freeman et al. (2002) using ITS, and can therefore be considered as reliable. Surprisingly, the three reconstruction methods strongly supported the branching of the strain from the non-Caryophyllaceae host *C. caespitosa* within *M. violaceum* clade D in both the β-tub and Ef1α phylogenies (Figs. 1 and 2). The other outgroups strongly supported the monophyly of all the *M. violaceum* strains plus the strain infecting *C. caespitosa* for both the β-tub and Ef1α trees. Outgroup sequences were removed from the datasets for subsequent analyses, except

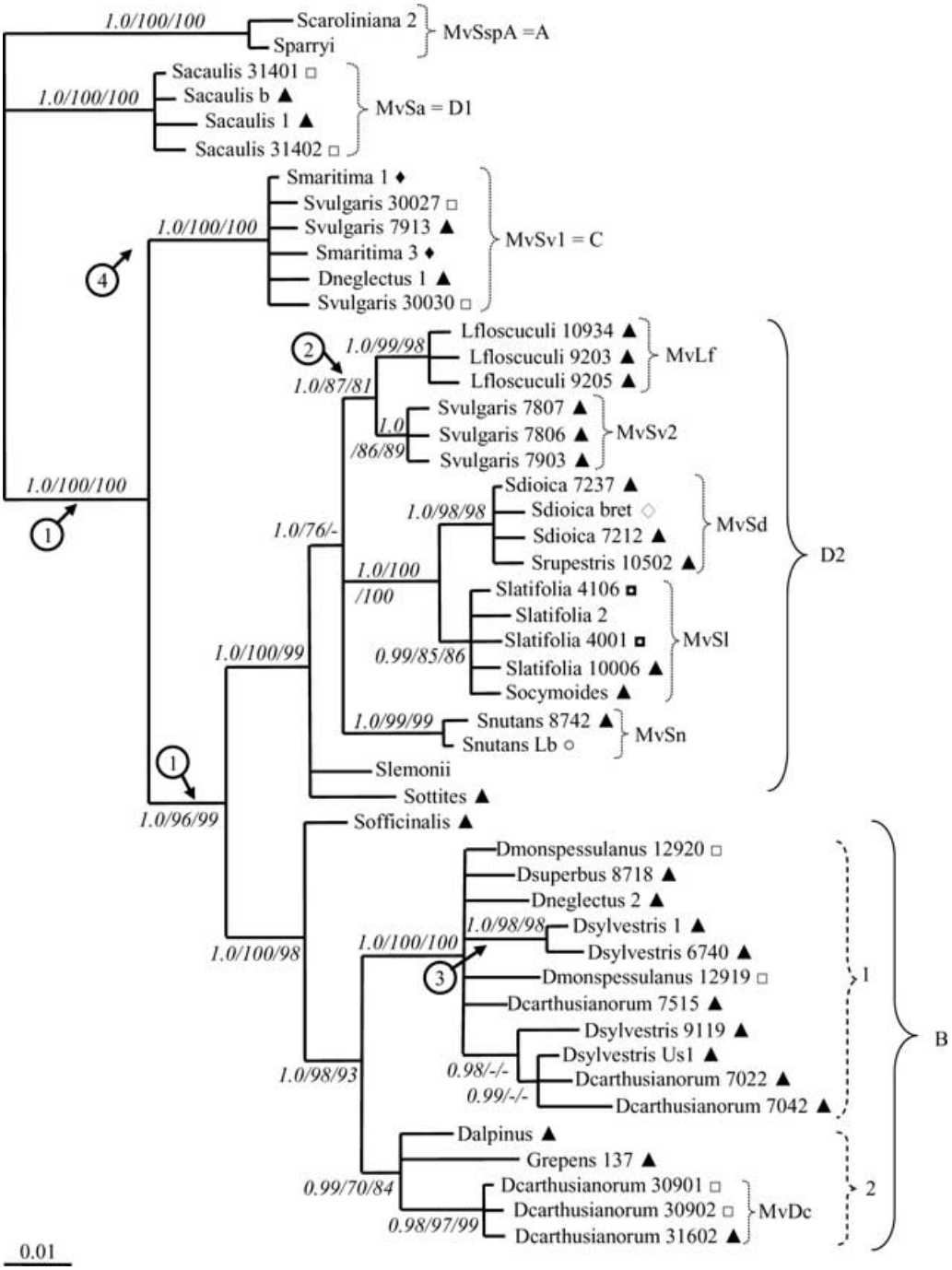

Figure 3. Bayesian 50% majority-rule consensus tree based on the γ-tub gene (mean lnL = −2238). See Figure 1 for the legend of the bootstraps, arrows, and taxa labels.

those of the *C. caespitosa* sample that did not appear actually to be an outgroup.

INTRAGENE RECOMBINATION

Maximum χ^2 and Maximum mismatch χ^2 methods each indicated multiple recombination events in Ef1α, but only one recombination event was detected by both methods and was therefore considered as significant. The Ef1α alleles of the strains involved belonged to clade C and resulted from recombination between a haplotype similar to the *Silene lemonii* strain and a haplotype similar to the *S. acaulis* strains. The entire clade C may therefore have experienced recombination, causing the position of this clade in the Ef1α phylogeny to remain uncertain (incongruency 4, Figs. 1, 2, and 3). Estimated beginning breakpoints for the insertion of the *S. acaulis*-like haplotype in the *S. lemonii*-like haplotype varied between the positions 357 and 507 of the aligned

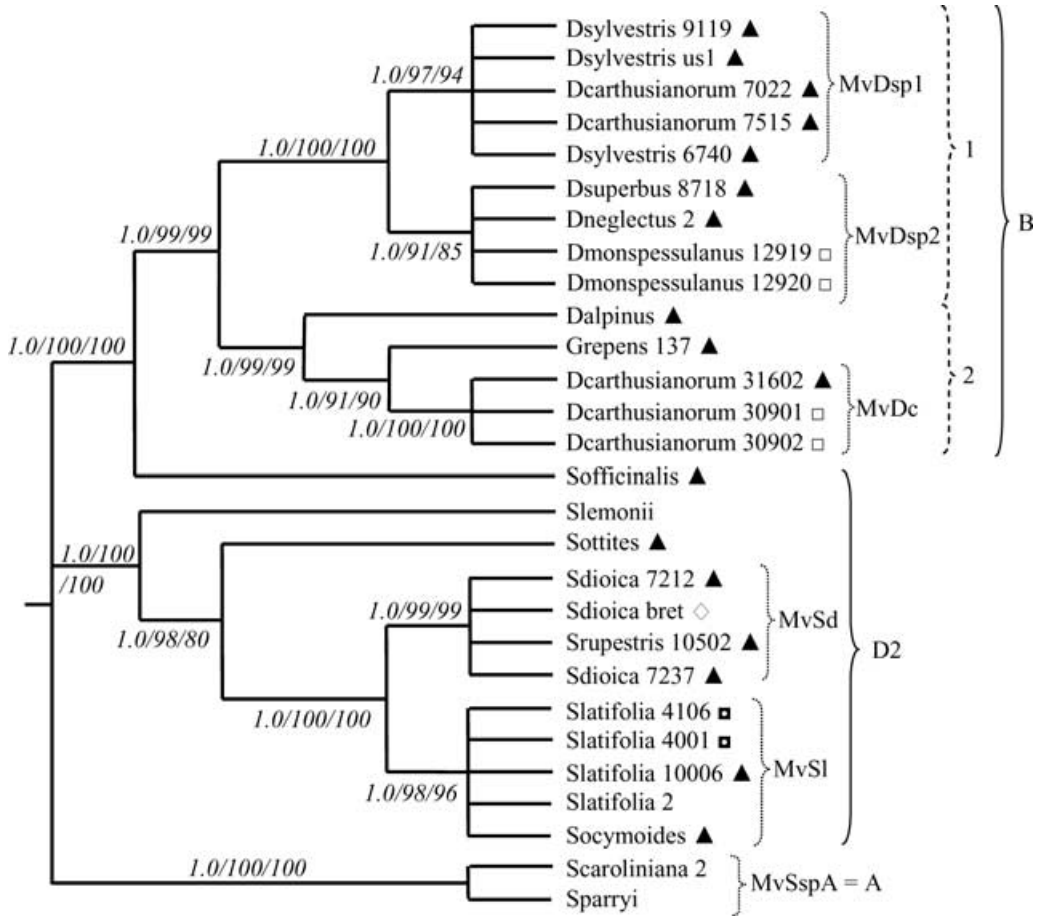

Figure 4. **Topology of the tree inferred from the concatenated analyses using the restricted combinable dataset. See Figure 1 for the legend of the bootstraps, arrows, and taxa labels.**

datafile, depending on the method and on the focal sequence. Estimated ending breakpoint for this insertion was the position 695 of the aligned datafile. None of the four tests of recombination (RDP, Bootscanning, Maximum χ^2, and Maximum mismatch χ^2) detected evidence of recombination within β-tub nor γ-tub.

IDENTIFICATION OF INDEPENDENT EVOLUTIONARY LINEAGES

Using the criterion of congruence between individual gene phylogenies, 11 independent evolutionary lineages were identified, labeled on the figures by the initials of their main host species (e.g., MvSa for *M. violaceum* evolutionary unit infecting *S. acaulis*, or MvSspA for *M. violaceum* evolutionary unit infecting *Silene* species native from America). Figure 5 summarizes the supported phylogenetic relationships among these 11 independent evolutionary lineages, as inferred from the different trees constructed. The evolutionary lineages each included mainly strains infecting the same host species, and all the evolutionary lineages geographically overlapped with several other lineages. Five individual strains (from *S. lemonii*, *Silene ottites*, *Silene officinalis*, *Dianthus alpinus,* and *Gypsophila repens*) were each strongly supported as evolutionarily independent of the 11 lineages identified above, and were therefore not included in any of these lineages. We, however, did not name specific lineages for them because these strains each branched alone in the trees.

Discussion

DETECTION OF SEVERAL HOST-SPECIFIC INDEPENDENT EVOLUTIONARY UNITS WITH OVERLAPPING GEOGRAPHIC RANGES

The multiple gene phylogenies approach in *M. violaceum* provided evidence of several independent evolutionary lineages evolving without gene flow for enough time for the ancestral polymorphism to be either lost or fixed in the descendant populations. Several groups were strongly supported as monophyletic by the three independent gene phylogenies. Moreover, the phylogenetic relationships supported by these three nuclear genes are also fully congruent with the results obtained with ITS sequence data from

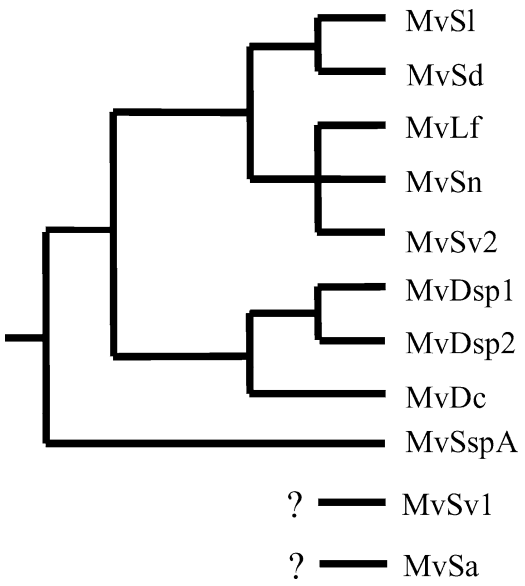

Figure 5. Summary of the phylogenetic relationships established between the 11 independent evolutionary lineages identified. The two lineages displaying incongruencies for which the position in the tree could not be resolved are not connected to it.

an independent set of strains from the same range of hosts (M. E. Hood, M. Le Gac, and T. Giraud, unpubl. ms.). We detected 11 independent evolutionary lineages, MvSspA, MvSa, MvSv1, MvSl, MvSn, MvSd, MvSv2, MvLf, MvDsp1, MvDsp2, and MvDc, that had largely overlapping geographic ranges. However, it remains to be determined whether regions or host species where sampling was limited may contain some unresolved cryptic species, and whether lineages that appear to infect more than one host species may actually be very host-specific but too recently divergent for detection using our markers.

Our results show that there is not frequent introgressions among the different lineages despite the possibility of cross-infections observed in experimental inoculations between some host plants and even in nature (Biere and Honders 1996; Shykoff et al. 1999; Antonovics et al. 2002; Hood et al. 2003; Van Putten et al. 2005) and the possibility of hybridization between strains from different hosts (Goldschmidt 1928; Biere and Honders 1996; Shykoff et al. 1999; Van Putten et al. 2003). The few incongruencies detected between individual gene phylogenies and the intragene recombination event generally correspond to ancient events. This is in agreement with the idea that in vitro interfertility, and even the presence of hybrids in nature, do not suggest that frequent introgression is likely (Coyne and Orr 2004, p. 42).

INDEPENDENT EVOLUTIONARY LINEAGES INFECTING THE SAME HOST

Interestingly, the host species *S. vulgaris* appears to be parasitized by two independent divergent lineages of *M. violaceum*, as previously suggested based on microsatellites (Bucheli et al. 2000). Here, we showed that the species *Dianthus carthusianorum* also appears to be parasitized by two distinct *M. violaceum* lineages. In both cases, the two lineages infecting the same host plant occur in the same mountains and even in the same host population for *S. vulgaris* strains (Svulgaris_7903 and Svulgaris_7913), raising questions of the conditions that allow their coexistence. Further investigations should be conducted to address why different lineages on the same host do not fuse or why one lineage is not outcompeted by the other.

COMPLEX PHYLOGEOGRAPHY

The lineage MvSspA includes most of the strains from native U.S. hosts and therefore corresponds to the native North American hosts strains clade described by Freeman et al. (2002). These authors reported a clear distinction between strains infecting native North American hosts and European hosts, suggesting a single and ancient divergence event. Their study was, however, based on a single gene, ITS, included collections from few European host species, and was restricted to east-coast populations in North America. In the present study, a strain from an additional native host in the western regions of North America, *S. lemonii,* was included and did not branch within the established North American clade A, but at the base of a clade containing European strains. This suggests a more complex geographic colonization and diversification history of *M. violaceum* than was previously known (Freeman et al. 2002).

INTROGRESSION AND HYBRIDIZATION

Four incongruencies were detected among the individual gene phylogenies. Because the third incongruence (incongruency 3) is located within a lineage, MvDsp1, and even within a host race, it probably only reflects usual intraspecific recombination. The three remaining incongruencies, strongly supported by comparing topologies and supports of the three phylogenies and AU tests (incongruencies 1 and 2) or by testing for intragene recombination (incongruency 4), may be due to genetic exchange between lineages, to the maintenance of ancestral polymorphism, or to convergence. Convergence appears highly improbable because apomorphies yielding incongruencies were numerous and were due to synonymous or intronic substitutions (data not shown). Ancestral polymorphism and ancestral recombination are not likely either given the great divergence between the different DNA fragments detected as incongruent and the fact that long stretches of DNA have the same evolutionary history. For instance γ-tub and Ef1α from *S. acaulis* strains branch in two very distant positions in the *Microbotryum* tree, but within each gene no site contradicts their respective position. The incongruencies 1, 2, and 4 therefore suggest the existence of genetic exchange events between divergent lineages of *M. violaceum*, that is, ancient introgressions

or speciation by hybridization. These incongruencies indeed involved all the strains from the given lineages. This indicates an ancient recombination event that may have occurred during speciation, and may thus reflect a speciation by hybridization, or genetic exchange just after speciation, when introgression was still possible. Signatures of ancient introgressions have been reported in other fungal species complexes, for instance, in *Epichloë* spp. (Schardl et al. 1997) and *Botrytis* spp. (Staats et al. 2005).

THE ROLE OF ADAPTATION AND GEOGRAPHY IN THE SPECIATION WITHIN *M. VIOLACEUM*

Samples of *M. violaceum* collected from the same host plants but in different geographic regions were genetically much closer than smuts collected on different host species in the same fields. This suggests a major role of specialization in the origin and maintenance of the *M. violaceum* lineages. However, this also makes the role of geographic isolation during speciation much more difficult to assess. Speciation events may have occurred in allopatry, and the current geographic distribution of species is the result of postspeciation dispersal events. It is indeed theoretically difficult to explain how two lineages can emerge in sympatry without divergent adaptations. It is therefore likely that two lineages infecting the same host species, as it is the case on *S. vulgaris* or *D. carthusianorum,* have first diverged in allopatry.

Conclusion

The degree of specialization and differentiation within *M. violaceum* has been debated for many decades. Genetic data suggested that this fungus was divided in several host-specific and genetically differentiated groups (Perlin 1996; Perlin et al. 1997; Shykoff et al. 1999; Bucheli et al. 2000, 2001; Giraud et al. 2002; Van Putten et al. 2005), although it was not clear whether these represented host races or true species. In contrast, cross-inoculation studies suggested that there was no strong specialization in the anther smut (Biere and Honders 1996; Shykoff et al. 1999; Van Putten et al. 2005), and the viability of hybrids both in vitro and in natura suggested that gene flow could be frequent among strains of different host species (Van Putten et al. 2005). The present study settles this issue, clearly showing the existence of independent evolutionary lineages, not exchanging gene for a long time, even when in sympatry, and establishing clear genetic boundaries for these lineages. Furthermore, this study suggests for the first time the existence of ancient introgression or hybridization events and large-scale species dispersal.

ACKNOWLEDGMENTS

The authors thank Y. Brygoo, B. Faivre, and A. Gautier for help with sequencing; A. Thomas, M. Bartoli, I. Till-Bottraud, M. Arroyo, A. Widmer, J. Antonovics, S. Triki-Teurtroy, S. Ribstein, and T. Darras for help in spore collection; J. A. Shykoff and R. Yockteng for critically reading the manuscript; H. Hollocher and three anonymous referees for constructive comments, and the French Pyrenees National Park for collection permit. TG acknowledges a grant "ACI Jeunes Chercheurs" from the French Ministry of Research and MEH. acknowledges award DEB-0346832 from the National Science Foundation. An ECOS grant (FONDECYT Cooperative Grant 7020956) allowed the spore collection in Chile. MLG was supported by a 2004 Systematics Research Fund award from the Linnean Society and the Systematics Association.

LITERATURE CITED

Antonovics, J., M. E. Hood, and J. Partain. 2002. The ecology and genetics of a host shift: *Microbotryum* as a model system. Am. Nat. 160:S40–S53.

Avise, J. C., and K. Wollenberg. 1997. Phylogenetics and the origin of species. Proc. Natl. Acad. Sci. USA 94:7748–7755.

Biere, A., and S. Honders. 1996. Host adaptation in the anther smut fungus *Ustilago violacea* (*Microbotryum violaceum*): infection succes, spore production and alteration of floral traits on two host species and their F1-hybrid. Oecologia 107:307–320.

Bucheli, E., B. Gautschi, and J. A. Shykoff. 2000. Host-specific differentiation in the anther smut fungus *Microbotryum violaceum* as revealed by microsatellites. J. Evol. Biol. 13:188–198.

———. 2001. Differences in population structure of the anther smut fungus *Microbotryum violaceum* on two closely related host species, *Silene latifolia* and *S. dioica*. Mol. Ecol. 10:285–294.

Burbidge, M. L., R. M. Colbourne, H. A. Robertson, and A. J. Baker. 2003. Molecular and other biological evidence supports the recognition of at least three species of brown kiwi. Conserv. Genet. 4:167–177.

Burnett, J. H. 2003. Fungal populations and species. Oxford University Press, Oxford.

Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Darlu, P., and G. Lecointre. 2002. When does the incongruence length difference test fail? Mol. Biol. Evol. 19:432–437.

Delmotte, F., E. Bucheli, and J. A. Shykoff. 1999. Host and parasite population structure in a natural plant-pathogen system. Heredity 82:300–308.

Dettman, J. R., D. J. Jacobson, and J. W. Taylor. 2003a. A multilocus genealogical approach to phylogenetic species recognition in the model eukaryotes *Neurospora*. Evolution 57:2703–2720.

Dettman, J. R., D. J. Jacobson, E. Turner, A. Pringle, and J. W. Taylor. 2003b. Reproductive isolation and phylogenetic divergence in *Neurospora*: comparing methods of recognition in a model eukaryote. Evolution 57:2721–2741.

Dres, M., and J. Mallet. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Philos. T. Roy. Soc. B. 357:471–492.

Farris, J. S., M. Kallersjo, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence. Cladistics 10:315–319.

Felsenstein, J. 1985. Confidence-limits on phylogenies—an approach using the bootstrap. Evolution 39:783–791.

Fournier, E., T. Giraud, C. Albertini, and Y. Brygoo. 2005. Partition of the *Botrytis cinerea* complex in France using multiple gene genealogies. Mycologia 97:1251–1267.

Freeman, A. B., K. K. Duong, T.-L. Shi, C. F. Hughes, and M. H. Perlin. 2002. Isolates of *Microbotryum violaceum* from North American host species are phylogenetically distinct from their European host-derivated counterparts. Mol. Phyl. Evol. 23:158–170.

Garber, E. D., M. L. Baird, and L. M. Weiss. 1978. Genetics of *Ustilago violacea*. II. Polymorphism of color and nutritional requirements of sporidia from natural populations. Bot. Gaz. 139:261–265.

Garber, E. D., and M. Ruddat. 2002. Transmission genetics of *Microbotryum violaceum* (*Ustilago violacea*): a case history. Adv. Appl. Microbiol. 51:107–127.

Geiser, D. M., J. I. Pitt, and J. W. Taylor. 1998. Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. Proc. Natl. Acad. Sci. USA 95:388–393.

Giraud, T. 2004. Patterns of within population dispersal and mating of the fungus *Microbotryum violaceum* parasitising the plant *Silene latifolia*. Heredity 93:559–565.

Giraud, T., E. Fournier, D. Vautrin, M. Solignac, and J. A. Shykoff. 2002. Isolation of 44 polymorphic microsatellite loci in three host races of the phytopathogenic fungus *Microbotryum violaceum*. Mol. Ecol. Notes 2:142–146.

Giraud, T., O. Jonot, and J. A. Shykoff. 2005. Outcrossing and selfing in the fungus *Microbotryum violaceum* inferred from artificial inoculations on the plant *Silene latifolia*. Int. J. Plant Sci. 166: 649–657.

Giraud, T., L. Villaréal, F. Austerlitz, M. Le Gac, and C. Lavigne. 2006. Importance of the life cycle in host race formation and sympatric speciation in parasites. Phytopathology 96:280–287.

Goldschmidt, V. 1928. Vererbungsversuche mit den biologischen arten den Antherenbrandes (*Ustilago violacea* Pers.). Zeitschrift fur Botanik 21:1–90.

Hood, M. E. 2002. Dimorphic mating-type chromosomes in the fungus *Microbotryum violaceum*. Genetics 160:457–461.

Hood, M. E., and J. Antonovics. 2004. Mating within the meiotic tetrad and the maintenance of genomic heterozygosity. Genetics 166:1751–1759.

Hood, M. E., J. Antonovics, and H. Heishman. 2003. Karyotypic similarity identifies multiple host-shifts of a pathogenic fungus in natural populations. Infect. Genet. Evol. 2:167–172.

Kasuga, T., T. J. White, G. Koenig et al. 2003. Phylogeography of the fungal pathogen *Histoplasma capsatulatum*. Mol. Ecol. 12:3383–3401.

Koufopanou, V., A. Burt, T. Szaro, and J. W. Taylor. 2001. Gene genealogies reveal cryptic species, and molecular evolution in the human pathogen *Coccidioides immitis* and Relatives (Ascomycota, Onygenales). Mol. Biol. Evol. 18:1246–1258.

Koufopanou, V., A. Burt, and J. W. Taylor. 1997. Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. Proc. Natl. Acad. Sci. USA 94:5478–5482.

Martin, D., and E. Rybicki. 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16:562–563.

Martin, D. P., C. Williamson, and D. Posada. 2005. RDP2: recombination detection and analysis from sequence alignments. Bioinformatics 21: 260–262.

Martinez-Espinoza, A. D., M. D. Garcia-Pedrajas, and S. E. Gold. 2002. The Ustilaginales as plant pests and model systems. Fung. Genet. Biol. 35:1–20.

Mason-Gamer, R. J., and E. A. Kellogg. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst. Biol. 45:524–545.

Maynard Smith, J. 1992. Analysing the mosaic structure of genes. J. Mol. Evol. 34:126–129.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends Ecol. Evol. 16:358–364.

O'Donnell, K., H. C. Kistler, B. K. Tacke, and H. H. Casper. 2000. Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum,* the fungus causing wheat scab. Proc. Natl. Acad. Sci. USA 97:1905–7910.

Perlin, M. H. 1996. Pathovars or *formae speciales* of *Microbotryum violaceum* differ in electrophoretic karyotype. Int. J. Plant Sci. 157:447–452.

Perlin, M. H., C. Hughes, J. Welch, S. Akkaraju, D. Steinecker, A. Kumar, B. Smith, S. S. Garr, S. A. Brown, and T. Andom. 1997. Molecular approaches to differentiate subpopulations or *formae speciales* of the fungal phytopathogen *Microbotryum violaceum*. Int. J. Plant Sci. 158:568–574.

Posada, D. 2002. Evaluation of methods for detecting recombination from DNA sequences: empirical data. Mol. Biol. Evol. 19:708–717.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.

Posada, D., and K. A. Crandall. 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. USA 98:13757–13762.

Pringle, A., D. Baker, J. Platt, J. Wares, J. Latgé, and J. Taylor. 2005. Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. Evolution 59:1886–1899.

Ravaoarimanana, I. B., B. Tiedemann, D. Montagnon, and Y. Rumpler. 2004. Molecular and cytogenetic evidence for cryptic speciation within a rare endemic Malagasy lemur, the Northern Sportive Lemur (*Lepilemur septentrionalis*). Mol. Phyl. Evol. 31:440–448.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Salminen, M. O., J. K. Carr, D. S. Burke, and F. E. McCutchan. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by Bootscanning. AIDS Res. Hum. Retroviruses 11:1423–1425.

Schardl, C. L., A. Leuchtmann, K.-R. Chung, D. Penny, and M. R. Siegel. 1997. Coevolution by common descent of fungal symbiont (*Epichloë* spp.) and grass hosts. Mol. Biol. Evol. 14:133–143.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492–508.

Shimodaira, H., and M. Hasegawa. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17: 1572–1574.

Shykoff, J. A., A. Meyhöfer, and E. Bucheli. 1999. Genetic isolation among host races of the anther smut fungus *Microbotryum violceum* on three host plant species. Int. J. Plant Sci. 160:907–916.

Sites, J. W. J., and J. C. Marshall. 2004. Operational criteria for delimiting species. Annu. Rev. Ecol. Evol. Syst. 35:199–227.

Staats, M., P. Van Baarlen, and J. A. L. Van Kan. 2005. Molecular phylogeny of the plant pathogenic genus *Botrytis* and the evolution of host specificity. Mol. Biol. Evol. 22:333–346.

Swofford, D. L. 2003. PAUP*: Phylogenetic analysis using parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, MA.

Taylor, J. W., D. J. Jacobson, S. Kroken, T. Kasuga, D. M. Geiser, D. S. Hibbett, and M. C. Fisher. 2000. Phylogenetic species recognition and species concept in fungi. Fung. Genet. Biol. 31:21–32.

Thomas, A., J. Shykoff, O. Jonot, and T. Giraud. 2003. Sex-ratio bias in populations of the phytopathogenic fungus *Microbotryum violaceum* from several host species. Int. J. Plant Sci. 164:641–647.

Thrall, P. H., A. Biere, and J. Antonovics. 1993. Plant life-history and disease susceptibility the occurence of *Ustilago violacea* on different species within the Caryophyllaceae. J. Ecol. 81:489–498.

Van Putten, W. F., A. Biere, and J. M. M. Van Damme. 2003. Intraspecific competition and mating between fungal strains of the anther smut *Microbotryum violaceum* from the host plants *Silene latifolia* and *S. dioica*. Evolution 57:766–776.

———. 2005. Host related differentiation in the anther smut fungus *Microbotryum violaceum* in sympatric, parapatric and allopatric populations of two host species *Silene latifolia* and *S. dioica*. J. Evol. Biol. 18:203–212.

Associate Editor: H. Hollocher