This is the submitted version of a paper published in *Evolution*.

Citation for the original published paper (version of record):

http://dx.doi.org/10.1111/evo.12374

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:nrm:diva-382
ABSTRACT

Fungal mycoparasitism – fungi parasitizing other fungi – is a common lifestyle in some basal lineages of the basidiomycetes, particularly within the Tremellales. Relatively non-aggressive mycoparasitic fungi of this group are in general highly host specific, suggesting cospeciation as a plausible speciation mode in these associations. Species delimitation in the Tremellales is often challenging since morphological characters are scant. Host-specificity is therefore a great aid to discriminate between species but appropriate species delimitation methods that account for actual diversity are needed to identify both specialist and generalist taxa and avoid inflating or underestimating diversity. We use the Biatoropsis-Usnea system to study factors inducing parasite diversification. We employ morphological, ecological and molecular data – methods including Genealogical Concordance Phylogenetic Species Recognition (GCPSR) and the general mixed Yule-coalescent (GMYC) model – to assess the diversity of fungi currently assigned to Biatoropsis usnearum. The degree of cospeciation in this association is assessed with two cophylogeny analysis tools (ParaFit and Jane 4.0). Biatoropsis constitutes a species complex formed by at least 7 different independent lineages and host switching is a prominent force driving speciation, particularly in host-specialists. Combining ITS and nLSU is recommended as barcode system in tremellalean fungi.

Keywords: cospeciation, coevolution, GMYC, Tremellales, species complex, integrative taxonomy
**INTRODUCTION**

Cophylogeny is the study of the evolutionary histories, captured as phylogenies, of two ecologically linked groups (Charleston 2002). Coevolution, particularly between parasites and their hosts, is among the most common targets of cophylogenetic studies. The terms coevolution and cospeciation are often used as synonyms, which can lead to misleading interpretations in cospeciation studies (Smith et al. 2008). Coevolution can be seen as a microevolutionary process of reciprocal adaptations, which does not necessarily imply cospeciation (Page 2003). It leads to reciprocal selection for improvements in host-parasite recognition mechanisms (de Vienne et al. 2013).

Cospeciation is a process whereby two symbionts speciate at the same time and results in congruent cophylogenies when it occurs. Cospeciation is often considered as one of the strongest evidences of coevolution (Cutchill and Charleston 2012). Most frequently, however, host and parasite phylogenies are not fully congruent (Huelsenbeck et al. 2003) and associations can then be explained as a result of processes such as host switching (Brooks and McLennan 1991), independent speciation of parasites (“duplication”) (Page 1990; 1996), lineage sorting (“loss”) (Paterson and Gray 1997; Paterson et al. 1999) or failure of the parasites to speciate (Page 2003). Coevolution can promote rapid diversification (Nunn et al. 2004; Thrall et al. 2007), and cospeciation as a result of coevolution was considered to be prominent in systems with high host specificity (Page 2003; Hughes et al. 2007; Light & Hafner 2008). However, parasite diversification can also be driven by host specialization, in the absence of cospeciation. For instance, Desdevises et al. (2002) found that despite the lack of cospeciation between fish and their monogean parasites, the high host-specificity in this system indicates a strong ecological association that could influence the diversification of parasites.
The study of these patterns in fungal mycoparasites is particularly appealing, since this type of association has never been studied in an evolutionary context before. Fungal mycoparasitism – fungi parasitizing other fungi – is a common and well-represented lifestyle in fungi. Within the basidiomycetes, mycoparasitism is mainly represented in basal lineages and has probably played an important role as evolutionary motor within this group (Weiss et al. 2004). Relatively non-aggressive mycoparasitic fungi are in general highly host-specific, which could suggest cospeciation with their hosts as one possible force driving speciation. This may well be the case in the Tremellales (Tremellomycetes, Basidiomycota, Fungi), one of the earliest diverging groups of the basidiomycetes where host-specific mycoparasites are particularly well represented (Boekhout et al. 2011; Millanes et al. 2011).

As a suitable system, we will investigate the association between the tremellalean Biatoropsis usnearum and its lichenized hosts – the last all belonging to Usnea and Protousnea (Parmeliaceae, Ascomycota, Fungi). Biatoropsis is currently monospecific, and distributed worldwide, but has been suggested to represent several species (Diederich and Christiansen 1994). It is morphologically very variable (Fig. 1), and can infect a large number of Usnea species. Previous work confirmed its placement within the Tremellales (Millanes et al. 2011), although the Biatoropsis clade showed high infraspecific genetic variation, and no specimens growing on Protousnea were studied. Biatoropsis usnearum may therefore be a species complex, providing an excellent system for studying early stages of speciation in conjunction with coevolution or cospeciation with its hosts.

Integrative taxonomy promotes the combination of several lines of evidence as a base for delineating species boundaries (DeSalle et al. 2005; Damm et al. 2010; Pires and Marinoni 2010). Adding information from coevolutionary studies to the
combination of different methods of DNA analysis and morphological studies could
greatly improve our understanding of the diversity of parasitic fungi, and its
evolutionary origin.

To investigate the diversity within *Biatoropsis* we will use two analytical
methods of species delimitation: 1) the general mixed Yule-coalescent (GMYC) model
(Pons et al. 2006; Fontaneto et al. 2007; Fujisawa and Barraclough 2013, Monaghan et
al. 2009), and 2) a variant of the Phylogenetic Species Recognition method—
Genealogical Concordance Phylogenetic Species Recognition (GCPSR) proposed by
Taylor et al. (2000). The GMYC model has proven to be particularly appropriate to
delineate species related to ecological variables and niche-specificity, including relation
to host species (Jousselin et al. 2009; Powell et al. 2011). This method has been
successfully used to study mycorrhizal diversity in environmental fungal samples
(Powell et al. 2011) and to assess diversity in lichens (Parnmen et al. 2012; Pérez-
Ortega et al. 2012), but it has otherwise not been much utilized in fungi, and it has never
been applied for species delimitation studies in parasitic fungi. It could potentially be
appropriate for obligate parasitic fungi, which are difficult to identify, and represent a
large component of the undescribed diversity (Hawksworth and Rossman 1997). On the
contrary, Genealogical Concordance Phylogenetic Species Recognition (GCPSR) has
largely proven to be a good tool for species delimitation in fungi (Grube and Kroken
2013).

In this study we test the monophyly of *Biatoropsis usnearum*, including
specimens growing both on *Usnea* and *Protousnea*, and in parallel assess the potential
species diversity within *Biatoropsis*, employing different analytical methods. We further
test for significant fit between the phylogenetic reconstructions of *Biatoropsis* and their
Usnea and Protousnea hosts, and investigate the extent of cospeciation occurring, in order to elucidate whether this is the main force driving diversity within Biatoropsis, or if other evolutionary processes explain the observed host-parasite associations. This is the first time the GMYC model is used to study the amount of diversity in a group of lichenicolous fungi, and the first time that a study on the joint evolution between lichenicolous fungi and their hosts is undertaken.
MATERIAL AND METHODS

Sampling

To confirm the phylogenetic position of Biatoropsis samples growing on Usnea and Protousnea within the Tremellales, we used a combined data set of nSSU, ITS and nLSU rDNA, which included all tremellalean groups distinguished by Boekhout et al. (2011), and Millanes et al. (2011). One taxon of the Holtermanniales (Wuczkowski et al. 2011) was used as out-group.

To focus specifically on the diversity within Biatoropsis, we used a combined data set of mtSSU, ITS and nLSU rDNA sequences, covering all macroscopic morphological variation observed in the genus, and including representatives from Asia, Australia, Europe, North America, and South America. To test hypotheses of coevolution between Biatoropsis and Usnea/Protousnea we used a dataset including the ITS and nLSU, for the parasites and hosts. We did so in order to compare homologous gene regions from the hosts and parasites, since mtSSU sequences could not be obtained for most of the Usnea samples. In 5 cases, the DNA from the exact host specimen could not be amplified. In these cases we used sequences from a different specimen of the same host, or sequences from GenBank. Host identification was conducted using morphological and chemical data. Secondary substances present in the Protousnea and Usnea samples were identified using thin layer chromatography (TLC; Culberson and Amman 1979) or high performance thin layer chromatography (HPTLC; Arup et al. 1993).

Supplementary Table S1 includes newly sequenced representatives, in which the species name, voucher information, and GenBank accession numbers are given, as well as the accession numbers of specimens retrieved from GenBank.
DNA extraction, amplification, and sequencing protocols are included as supplementary material (Material & Methods).

Species delimitation, GCPSR and the GMYC model

For the dataset including only samples within *Biatoropsis* s. str., we used the GCPSR (Taylor et al. 2000) and the generalized mixed Yule coalescent (GMYC) model (Pons et al. 2006; Fontaneto et al. 2007) to delimit independent evolutionary lineages from the topology of the trees.

The GCPSR method requires concordance of single gene phylogenies to identify the limits of independently evolving lineages (i.e., species), and is based on the simple visual comparison of more than one gene genealogy (Taylor et al. 2000). We therefore analyzed the independent datasets individually by maximum parsimony using PAUP* 4.0b10 (Swofford 2002), maximum likelihood using the graphical user interface of RAxML: RAxMLGUI (Stamatakis 2006; Silvestro and Michalak 2012) and Bayesian analyses using MrBayes 3.2.1 (Ronquist et al. 2012), with the corresponding settings used for each partition of the combined dataset as described in the supplementary material.

The GMYC method combines the neutral coalescent theory (Hudson 1991; Wakeley 2006) with Yule speciation models (Yule 1924) and, under a maximum likelihood approach, assesses the branching rates along an ultrametric tree and detects the point where the transition from rates corresponding to speciation events (speciation or extinction) to rates indicating coalescence events (i.e., population-level evolutionary processes) occurs most likely. This transition point (threshold) determines species boundaries in the way that nodes before the threshold are identified as species diversification events, whereas nodes after the threshold are considered clusters.
following coalescent processes, i.e., populations (Pons et al. 2006; Fontaneto et al. 2007). Under this approach, the null model assumes that the sample of individuals derives from a single entity, and follows a coalescent process, whereas the alternative model assumes that the sample contains several independently evolving entities. A modification of this method – the multiple threshold GMYC model - was developed by Monaghan et al. (2009) to allow the transition point to vary among lineages, i.e., multiple threshold model. Fujisawa and Barraclough (2013) introduced a correction in the implementation of the GMYC method, treating the threshold times as model constraints instead of as parameters. As a consequence the single and multiple threshold methods cannot be compared by likelihood ratio test, since they have the same number of parameters. The ultrametric trees used in the analyses were generated under a Bayesian approach using BEAST version 1.7.5 (Drummond and Rambaut 2007).

Procedures to generate the ultrametric trees and to optimize the GMYC model are available as supplementary material (Material & Methods).

**Statistic analyses**

We searched for statistical differences among lineages in the size of the galls, where the sampling was large enough to achieve reliable statistical analyses. Since the data did not adjust to a normal distribution, we used the non-parametric Kruskall-Wallis test (Zar 1984; PAST for windows), analogue to the one factor ANOVA, to examine whether there were significant differences among the groups. Additionally, the correction factor was used to compute the $H_c$ Kruskal-Wallis statistics (Zar 1984:179). If the Kruskal-Wallis test was significant, Mann-Whitney pairwise comparisons were then used to test for differences among unique pairs of clusters. The value of alpha was adjusted dividing by the number of comparisons (i.e., Bonferroni correction) as a conservative correction.
Cospeciation analyses

Cospeciation between hosts and parasites was assessed by distance/topology based methods or event-cost based methods. Distance-based methods determine if hosts and their parasites are associated randomly by comparing genetic distances from homologous gene regions for the associated taxa. Event-cost based methods compare only the branching structure of host and parasite trees to determine if tree topologies are more similar than would be expected by chance (Light and Hafner 2008). Such methods use heuristics to find solutions that minimize the overall cost of a historical reconstruction given a cost regime for different events including cospeciation, host switching, lineage sorting, etc.

We used both a distance based and an event-cost based test of cospeciation. We used a distance method –ParaFit Global test (Legendre et al. 2002) – to test the null hypothesis \( H_0 \) that each parasite species is associated with hosts selected randomly along the host phylogenetic tree, against the alternative hypothesis \( H_1 \) that the positions of the individual host-parasite associations are not random but associated to phylogenetic distances between hosts and parasites. Analyses were conducted using the ParaFit method as implemented in the R package “ape” (Paradis et al. 2004). Three matrices were constructed, i.e., \( A \): a matrix describing the associations between hosts and parasites, \( B \): the parasite phylogeny transformed into a distance matrix, and \( C \): the host phylogeny transformed into a distance matrix. Since neither the parasite, nor the host phylogenies were fully resolved, we used matrices of phylogenetic distances – instead of patristic distances – computed directly from the row data (Legrende et al. 2002). The method, as implemented in R (Paradis et al. 2004), calculates principal
coordinates matrices from matrices B and C, to calculate a fourth matrix (the four
corner matrix D), from which it calculates a statistic that tests for significance of the
associations between hosts and parasites. Statistical significance of the ParaFitGlobal
statistic was assessed by comparison with a null distribution created by 999
permutations of the host-parasite association matrix while keeping matrices B and C
fixed. If the observed ParaFitGlobal statistic is larger than the randomized statistic in ≥
95% of cases, the null hypothesis of independent evolution can be rejected, and
cospeciation can be assumed. Testing the significance of individual host-parasite
phylogenetic associations is also possible by testing their contribution to the
ParaFitGlobal. For both the distance-based and the event-cost based analyses we used a
pruned ITS and nLSU alignment including only one representative per putative species
of hosts and parasites. The pruned topologies were obtained using Bayesian analyses
with the same substitution models and settings described in the phylogenetic analyses
(see supplementary material). Since topologies were congruent with the analyses
performed using the complete sampling, we used the 50% majority rule consensus trees
from the pruned Bayesian analyses, for the tree-based coevolutionary analyses. We used
Jane 4.0 (Conow et al. 2010) to test for significant congruence between the *Biatoreopsis*
and *Usnea* topologies. Cophylogeny mapping in Jane uses heuristics to reconstruct
histories that explain the similarities and differences between associated phylogenies. It
prioritizes minimizing the overall cost, given a cost regime for evolutionary events
including ‘cospeciation’ (a host and a parasite speciate simultaneously), ‘duplication’ (a
parasite speciates and both of the new species remain on the same host), ‘duplication
and host switch’ (a parasite speciates and one of the new species switches onto a
different host), ‘loss’ or ‘lineage sorting’ (a host speciates and the parasite remains only
on one of the new host species) and ‘failure to diverge’ (a host speciates and the parasite
remains on both new host species) (Charleston 1998; Conow et al. 2010). It is often
difficult to accurately estimate the relative costs of events. We therefore tested several
cost regimes, with settings detailed in Table 4. The different cost regimes were selected
in order to test the overall costs of reconstructions, when minimizing the costs of
different events, penalizing cospeciations or host switches, or giving all events the same
weight. In all analyses the number of generations was set to 100, and the population size
to 300, with a maximum of 99999 stored solutions in each run. Statistical analyses were
then performed in order to test whether the cost of the reconstructions obtained was
significantly lower than expected by chance. Jane 4.0 achieves this by generating a
pseudo-random sample of minimal costs from a null distribution of problem instances
with the same model phylogeny. The null distribution is generated by repeatedly
randomising either the host-parasite associations (Random Tip Mapping), or the branch
pattern of the host tree (Random Parasite Tree). We assessed the significant matching of
host and parasite phylogenies by computing the costs of 1000 replicates – with the same
settings described above – to compare the resulting costs to the cost of the original
cophylogeny. Polytomies in the tree topologies were included in the analyses, since
Jane 4.0 is able to resolve soft polytomies in both trees to minimize the cost of the
cophylogenetic reconstruction.

We selected ParaFit as distance-based method and Jane 4.0 as event-cost based
method, because both approaches accept: 1) unbalanced numbers of hosts and parasites,
2) multihost parasitism, and 3) phylogenies that are not completely resolved.
RESULTS

Results on DNA data generated and on phylogenetic analyses are available as supplementary material (Results).

Species delimitation by GCPSR and GMYC analyses

Seven species were inferred by GCPSR (Taylor et al. 2000), consisting of lineages that fulfilled the assumption of concordance among independent gene trees (Table 2; Fig. 2A; Supplementary Fig. S2). Lineages A1 and A2 were separated by the nLSU topology, and this separation was not contradicted by the other two markers.

Independently of the dataset used, both the single-threshold and multiple-threshold GMYC models resulted in a significant better fit to the ultrametric tree than the null model, except when using the nLSU individually (Table 1). In all cases the single threshold model delimited 6 potential species and the multiple-threshold model delimited 7 potential species (Fig 3, Table 1). When using a single-threshold model, the narrowest confidence interval for the number of entities (i.e., potential species) corresponded to the combined analysis and to the single locus analyses using mtSSU and ITS. When using a multiple-threshold model, the narrowest confidence interval for the number of entities corresponded to the combined analysis and to the single locus analysis using mtSSU (Table 1). Using this approach lineage F splits in two independent lineages, namely F1 and F2 (Table 1; Fig. 2A; Fig. 3A).

Table 2 summarizes the lines of evidence supporting each potential species in the complex. Four lineages (B-E) are supported by all molecular analytical methods, and at least by one additional line of evidence. The separation between lineages A1-A2 is recovered by GCPSR, but not by the GMYC approach. Two other lines of evidence support the split A1-A2, i.e., gall size and host specificity of A2 after a host switch,
including five samples. The split F1-F2 is more controversial since it is only supported
by the GMYC (multiple-threshold) model, the colour of the galls, and host specificity of
F2 after a host switch, all based on a single sample. We therefore considered four
possible species delimitation hypotheses: *i*) six potential species recovered by the
GMYC single-threshold model (i.e.: A, B, C, D, E and F; Fig. 3A); *ii*) seven potential
species recovered by the GMYC multiple-threshold model (i.e.: A, B, C, D, E, F1 and
F2; Fig. 3B); *iii*) seven lineages inferred by GCPSR (i.e.: A1, A2, B, C, D, E and F; Fig.
2A; Supplementary Fig. S2); and finally, *iv*) eight lineages including the entities
recovered by the GMYC multiple-threshold model, but splitting lineage A in two
independent ones as inferred by GCPSR, (i.e.: A1, A2, B, C, D, E, F1 and F2; Fig. 2A;
Supplementary Fig. S2). These species-delineation scenarios were considered for
morphological statistical analyses and for coevolutionary analyses.

Fig. 2B depicts the host phylogeny. In total we include 34 terminals in the
phylogeny of *Usnea* and *Protousnea*, representing 16 species. As a conservative
approach regarding coevolutionary studies, we considered samples that had been
morphologically determined as the same species, but where genetic distance or tree
topology suggested the possibility of two different lineages, as two different species.
This was the case for the two samples of *U. flavocardia* and one species in the *U.
cornuta* group (AM296), which appeared out of the clade including other
representatives of *U. cornuta*. These two species have been shown to be
morphologically highly variable and could potentially include several undescribed
species (Truong et al. 2011; 2013). Contrarily, we considered the clade including three
unidentified samples of *Usnea*, collected in a small area in Chile, as the same lineage.
Samples of *U. florida* and *U. subfloridana* were considered one lineage, following
Articus et al. (2002) (Fig. 2B).
Statistical analyses of morphological traits

We studied the size of the galls to compare two alternative species delimitation hypotheses: i) six lineages recovered by the GMYC single-threshold model (Fig. 3A) – which would correspond to the lowest diversity inferred – and iv) eight lineages identified by combining the species recovered by GCPSR and the GMYC multiple-threshold model (Fig. 2A and 3B; Supplementary Fig. S2) – which corresponds to the hypothesis accounting for the higher diversity within Biatoropsis. In both cases we found significant differences in the morphology among lineages (p < 5·10^{-20}; Table 3). The level of significance was notably higher when comparing the lineages considered by ‘hypothesis iv’ of species delimitation (p = 9,808·10^{-23}; Table 3B) than when considering the GMYC entities using a single-threshold model, i.e., ‘hypothesis iii’ (p = 1,168·10^{-20}; Table 3A). Lineages A1, A2 and B were the maximum contributors to the significance of the global test.

Cospeciation and Cophylogeny tests

The global ParaFit test did not allow rejecting the null hypothesis of random association, independently of the species delimitation hypothesis considered for parasites: i) (GMYC single threshold) p-value=0.377; ii) (GMYC multiple threshold) p-value=0.236; iii) (GCPSR) p-value=0.197; (GCPSR and GMYC multiple threshold) p-value=0.285. However, event-cost based tests using Jane 4.0 detected significant congruence (i.e., > 95% of random solutions were worse than the solution reconstructed by Jane) between Biatoropsis and Usnea/Protousnea phylogenies, in most cases (Table 5). The only exception appeared when using cost regime 7, in which losses were not penalized. In this case, all random solutions were equally good compared to the solution reconstructed by Jane 4.0. Using the species delimitation hypotheses considering...
lineages A1 and A2 as independent (i.e., hypotheses iii and iv) always resulted in a lower overall cost (Table 5). Assigning a lower cost to host switches than to cospeciations (cost regimes 4-6; Tables 4 and 5) obviously increased the number of host switches inferred, but also decreased the overall cost in all cases, compared to reconstructions when the cost of cospeciations was lower than that of host-switches, or when both events were assigned the same cost (cost regimes 1-3; Tables 4 and 5).

Penalizing cospeciations always resulted in lower overall costs, independently of the species delimitation hypothesis considered and of the cost regime utilized (Tables 4 and 5). Moreover, penalizing cospeciations over host switches yielded always the same particular four or five host switches, also independently of the species delimitation hypothesis considered, and of the cost regime utilized (cost regimes 4-6; Tables 4 and 5). The same host switches were recovered when the cost of failure to diverge was set to 0 (cost regime 6; Tables 4 and 5). All host switches reconstructed were related to the origin of host-specialized lineages (i.e., A2, B, D, E and F2; Fig. 4). Particularly the host switch explaining the origin of lineage A2 was supported in all reconstructions including A2 as a potential species. Assigning cost 0 to losses and failure to diverge resulted indeed in less costly reconstructions, but these were never significantly better than the reconstructions generated in the randomized tests (Cost regime 7; Tables 4 and 5). These reconstructions were therefore not considered. However, all significant reconstructions included a similar number of failures to diverge – ranging from 9 to 11 – independently of the species delimitation hypotheses and cost regimes utilized (Table 5). Among the significant reconstructions (Table 5), the lower overall costs were recovered using the species delimitation hypotheses iii and iv, and utilizing the cost regime ‘5’ – which penalized cospeciations and assigned cost=0 to failures to diverge.
These reconstructions consisted of 1 cospeciation, 1 duplication, 4 to 5 host switches, 8 losses, and finally 9 to 10 failures to diverge (Table 5; Fig. 4).

DISCUSSION

_Biatoropsis usnearum_ is a species complex formed by at least 7 potential independent lineages. Cospeciation does not seem to have played a major role in the diversification of the group. Contrarily, host switching appears as a prominent force driving speciation, particularly in host-specialized parasites. Losses and failure to diverge are also frequent in the history of the _Biatoropsis-Usnea/Protousnea_ association.

Host switching, rather than cospeciation, originates host-specialists within _Biatoropsis_

Cospeciation is a process in which speciation in one lineage is accompanied by speciation in an associated but unrelated lineage (Clayton et al. 2003) and implies time congruence in speciation events of both associated lineages. Cospeciation was traditionally the prevalent hypothesis to explain diversification in parasites, especially when they were host-specialized or had a narrow host-range (Paterson et al. 2000; Perez-Losada et al. 2006; Light and Hafner 2008). However, the cospeciation paradigm in host-parasite systems gradually lost strength, as studies on coevolutionary mechanisms showed that other processes were more frequent in these associations (Peterson et al. 2010; Badets et al. 2011; Jansen et al. 2011). In a recent critical review, de Vienne et al. (2013) claim that only 7% of such studies constitute convincing cases of cospeciation. Our ParaFit tests were not significant independently of the species delimitation hypothesis considered, which rules out cospeciation as a main diversification mode in _Biatoropsis_. Our results add to the multiple examples where
cospeciation does not explain the evolutionary story of a host-parasite association. On the contrary, in the *Biatoropsis-Usnea* system, host-switching events are more frequent than cospeciation events in the best cophylogeny reconstructions. Host switching has been recognised as a main mode of speciation in many coevolutionary studies (Roy 2001; Charleston and Robertson 2002; de Viene et al. 2013; Perlman et al. 2003; Weckstein 2004; Huyse et al. 2005; Shafer et al. 2009; Lee and Stock 2010; Lewis-Rogers and Crandall 2010; Gottschling et al. 2011; Longdon et al. 2011; Cui et al. 2012; Irwin et al. 2012; Mcleish and Noort 2012), including systems of plant pathogenic fungi (Jackson 2004; Refrégier et al. 2008). Host switches have been related to bursts of species diversification (Fordyce 2010). In particular mycoparasitism has shown to increase speciation rates in parasitic fungi after a host switch from plant-hosts to fungal-hosts (Chaverri and Samuels 2013). Whether speciation rates are higher in mycoparasitic lineages of the Tremellales, and if this habit was the result of one or several host switches remains to be tested, but the abundance of mycoparasitism within the group suggests its link with episodes of accelerated diversification.

It has been argued that coevolution processes can promote speciation through host specialization (Summers et al. 2003), but in those cases speciation occurs more frequently through host switches than through cospeciation (de Vienne et al. 2013). There are alternative explanations for host restriction including limited dispersal or limited adaptation to hosts (Timms and Read 1999). Some parasites may have evolved a restricted host range due to geographical barriers, for instance when hosts occur in allopatry, or because limited dispersal abilities, without any evolutionary processes of reciprocal adaptation being involved (Refrégier et al. 2008). Werth et al. (2013) compared the demographic stories of the host-specific *Tremella lobariacearum* and its *Lobaria* hosts and showed that host specificity, and not geographical barriers among
islands, was the most important factor structuring the populations of the lichenicolous species. This is also the case in the *Biatoropsis/Usnea* system where host-specificity, appears to influence speciation in the lichenicolous partner independently of geography. Alternatively to geographical constraints, host specificity may arise because of adaptive specialization (van Tienderen 1991), i.e., when being a generalist reduces fitness, and host-specificity may be an advantage even in sympatry (Refrégier et al. 2008). Our results reveal that different host-specific species of *Biatoropsis* often occur in sympatry suggesting speciation by adaptive specialization. A better knowledge on the reproduction and dispersal methods of parasites can help to interpret speciation through host specialization – understood as reproductive isolation accompanying host-specificity. Little if any is known on the dispersal mechanisms of *Biatoropsis* or on its reproductive mode. *Biatoropsis* can potentially reproduce sexually (see Bandoni 1995, for details on the life-cycle of the Tremellales). However, the low genetic divergence of samples collected in geographically distant areas (Supplementary Fig. S3) suggests long distance dispersal and vertical transmission of the parasite together with the host, by means of asexual reproduction. It has been demonstrated that asexual organisms can diversify into independently evolving and distinct entities equivalent to species, and that adaptation to different ecological niches is among the possible causes of divergence (Fontaneto et al. 2007).

In a general context, host specificity has been claimed to be the reason for ‘non-host resistance’ – i.e., resistance of species that are not considered to be hosts of a certain parasite –, opposite to ‘evolved resistance’ or resistance acquired by evolution of the host towards the parasite (implying coevolution) (Antonovics et al. 2012). The range of *Usnea* and *Protousnea* hosts that can be infected by *Biatoropsis* is currently difficult to assess, but of the ca. 350 species included in *Usnea* and *Protousnea* together, only 54
have been reported to harbour basidiomata of the parasite (Diederich and Christiansen 1994; own unpublished observations). Field studies have shown that in mixed populations of *Usnea* species, not all species are attacked (Diederich and Christiansen 1994). The reasons for host specificity and host resistance in this system, as well as in other systems of mycoparasitic fungi, have never been investigated, and deserve further attention in future evolutionary studies.

**Losses, failure to diverge, and generalist species**

Generalist species are also present in the *Biatoropsis-Usnea* system, if we consider species that grow on different species of the same genus, ‘generalists’. The importance of recognizing generalist species in parasite diversity surveys has been already pointed out (Refrégier et al. 2008; de Vienne et al. 2012), particularly in order to make sounder assumptions in cospeciation studies. The existence of systems where both generalists and specialists occur are not rare (Little et al. 2006; Poullain et al. 2008) and might be influenced by a combination of selection pressures and compensations of e. g. being able to grow on a wider range of hosts, or a higher rate of adaptation of specialists (de Vienne et al. 2013). Loss and failure to diverge are in general more abundant in cophylogeny reconstructions than might initially be thought. They are prevalent in the *Biatoropsis-Usnea* association, and the latter seem to be related to the origin of generalist species in the complex. Extinctions have been reported to be frequent in parasites, either due to the gain of resistance by the host (Ricklefs et al. 2004) or to a decline in host population size (de Castro and Bolker 2005). When a host speciates, the population size of the emerging new species can often be too small for the parasite to survive, and this can easily become extinct on the incipient host (Altizer et al. 2007; Gibson et al. 2010). On the contrary, failure to diverge is one of the evolutionary
processes generating multi-host parasites (Johnson et al. 2003). It occurs when parasite populations survive and maintain genetic flow despite the hosts diverging, generally between parasite populations that grow on originally closely related hosts (Banks and Paterson 2005). The considerable number of failures to diverge reconstructed in deeper nodes of our parasite phylogeny suggests that generalist species are ancient lineages within *Biatoropsis*.

One important question remains: which are the factors favouring the origin of specialist and generalist species in this system? It has been hypothesized that lichen secondary metabolites could play important roles in the specialization of lichenicolous fungi (Lawrey and Diederich 2003). Werth et al. (2013) suggested that hosts can create a selective environment for the lichenicolous *Tremella lobariacearum* since, despite their wide geographic distribution, the same parasite haplotypes consistently associate with a given host species. Host specific lineages within *Biatoropsis* all grow on hosts that have rather distinct morphological and chemical characteristics compared to the original hosts (Calvelo et al. 2005; Clerc 2006), supporting speciation in sympatry. In particular, A2 evolved host specificity on a different genus (*Protousnea*) after a host switch. Whether secondary lichen compounds or morphological characteristics can be influencing hosts specialization is, however, not known. Other authors have suggested the evolution of degradative specific enzymes by specialized parasites as alternative adaptive mechanism (Lawrey, 2002). Some plant pathogens suppress costly metabolic pathways, when they get particular nutrients directly from the host, which conditions their survival to obligate biotrophy on selected plant species (Kemen et al. 2011). In the case of generalist parasitic lineages, diverging host species should present biochemical similarities that allow the same parasite to adapt to the hosts. A thorough study of the
biochemical mechanisms of pathogenicity and resistance constitutes another highly interesting future line of research in mycoparasitic fungi.

**Overlooked diversity within Biatoropsis**

Fungal species complexes are abundant (Le Gac et al. 2007; Crespo and Pérez-Ortega 2009; Crespo and Lumbsch 2010; Diogo et al. 2010; Hyde et al. 2010; Rodríguez-Estrada et al. 2010; Cai et al. 2011), which make fungi suitable organisms in which to investigate early stages of speciation and to understand mechanisms generating diversity (Guiraud et al. 2008). The drawback is that species delimitation is often challenging in some groups (including Tremellales) due to scarcity of morphological characters. In host-specific parasitic fungi, species delimitation relies frequently on hosts, but it is important to utilize adequate species delimitation methods that account for actual diversity, identifying both specialist and generalist species (Refrégier et al. 2008). One of the approximations that integrative taxonomy offers to delineate species boundaries is the so-called “taxonomic circle”, which requires the coincidence of at least two independent lines of evidence to allow supporting species-recognition hypotheses (DeSalle et al. 2005). Lines of evidence can be based on DNA data, morphology, geography, reproduction, and ecology (Ferri et al. 2009; Damm et al. 2010; Cruz-Barraza et al. 2012). Following this approach we combined several evidences, including a) DNA data (using different analysis methods for species recognition), b) morphology, and c) ecology in order to outline species-delimitation hypotheses within *Biatoropsis usnearum*. Under this approach, the genus harbours at least 7 species, and lineage F2 deserves further investigations as a possible independent evolving entity (Table 2). Diederich and Christiansen (1994) had earlier suggested that *Biatoropsis* might include several species, but they refrained from separating
morphologically distinct taxa (i.e., based on gall colour and morphology) due to the abundance of intermediate forms. Indeed, among all lineages within Biatoropsis only B is clearly distinct morphologically, regarding size and colour of the galls. For the rest, even when gall sizes generate statistical differences, overlapping does not allow to use this character to identify species. Also the colour of the galls is too variable, except for lineage B, to base circumscriptions on this trait. Therefore, an integrative strategy is especially useful to 1) delineate species boundaries in difficult diagnosable groups 2) to identify the methods and lines of evidence that can contribute more importantly in particular species-delimitation problems, and 3) to identify well supported species and those where additional investigations would be necessary.

Differences between species-delimitation methods and the problem of rare species

GCPSR is currently the most widely used species recognition criterion in fungi and it has proven to perform more accurately than other discriminating methods (e.g., morphological species recognition –MSR), or to be more convenient for particular groups where other criteria as biological species recognition (BSR) are not applicable, (e.g., asexual fungi) (Guiraud et al. 2008). The method detects the separation between lineages A1 and A2, which is also sustained by the distinct ecology of species A2 – restricted to Protousnea dusenii – as a consequence of a host-switch. Phylogenetic studies within the lichen-forming family Parmeliaceae indicated that Protousnea forms a distant lineage from Usnea (Crespo et al. 2007, Truong et al. 2013), but several Usnea species are sympatric with Protousnea in Southern South America, which might have favoured the host-switch. However, the GMYC approach does not support the independence of lineages A1 and A2. It has been demonstrated with simulated data that recent speciation events are often not detected by this method (Fujisawa and
Barraclough 2013). On the contrary, the GMYC method performs relatively well to identify rare species where sampling is limited to one or a few representatives. For instance, the separation of lineages F1 and F2 is only supported by a GMYC multiple-threshold approach. Lineage F2 constitutes a singleton (i.e., species represented by a single specimen) and differs from specimens included in lineage F1 in gall morphology (black galls instead of orangish galls) and host (*U. ceratina*).

Species delimitation of rare species, often based on single specimens, is indeed problematic, but such species are otherwise common in diversity surveys (Hajibabaei et al. 2006; Pons et al. 2006; Coddington et al. 2009; Monaghan et al. 2009). Pons et al. (2006) suggested that, since the GMYC model is a branching-based approach, independent of the recognition of population limits, ‘the approach permits the inclusion of rare species represented by only a single individual, which is problematic in population based species concepts’. Although some authors have been skeptical about the sensitiveness of the method in cases of undersampling of taxa (Lohse 2009), the modification of the GMYC method achieved by Monaghan (2009) – allowing the transition point from coalescent to cospeciation to vary along branches – reveals as specially appropriate to identify species boundaries even under conditions of scant sampling for rare species (Papadopoulou et al., 2009). The simulations achieved by Reid and Casterns (2012) showed that the GMYC model can be particularly useful when taxon sampling is incomplete. Caution is recommendable however when using the multiple threshold approach, since its tendency to over-split has been also reported (Fujisawa and Barraclough 2013). Of the two singletons recovered in our study, lineage F2 is therefore considered with some hesitation, but there are signs (different host and different colour of the galls) indicating that additional sampling would be worthwhile to confirm or reject F2 as a species. The multiple-threshold GMYC method should not be
neglected as a tool to bring attention to rare or possibly undersampled lineages, on which further investigations should be conducted.

The combination of ITS and nLSU as a barcode system for the Tremellales

The GMYC method is widely used for delimiting species from single locus data, but can also be used to delimit lineages from any ultrametric trees, including those obtained by combining multilocus data (Fujisawa and Barraclough 2013). Our comparisons of species-delimitations by GMYC employing single locus data, or the concatenated dataset, suggest that the method performs better on the ultrametric tree derived from the combined dataset (Table 2; Fig. 3) regarding the level of significance supporting the GMYC model over the null model of coalescence, and the restricted confidence interval. The mitochondrial small subunit rDNA (mtSSU) retrieves even higher values of significance and narrower confidence intervals, but this could be an effect of the lower number of haplotypes sampled in the mtSSU analyses. The use of the internal transcribed spacer and the large subunit of the rDNA (ITS + nLSU) as a single marker gives a total coincidence on lineage-delineation with the combined dataset, independently of using the single or the multiple-threshold variant of the GMYC model. Some conflicts were observed with ITS alone and mtSSU and the nLSU used alone did not yield significant results, although species recognized by this marker coincide also with the recovered using the combined dataset (Fig. 3). ITS has recently been proposed as the official primary barcoding marker for fungi (Eberhardt 2010; Schoch et al. 2012). It has been mentioned, however, that the nuclear ribosomal large subunit (nLSU), had superior species resolution in some taxonomic groups such as early diverging lineages and basidiomycete yeasts (Schoch et al. 2012). The possibility of a two-marker barcoding system for fungi (i.e., combining ITS and nLSU) has sometimes been
preferred among mycologists, particularly by researchers working on tremellalean yeasts (Fell et al. 2000; Scorzetti et al. 2002), Glomeromycota (Stockinger et al. 2010) or environmental samples (Klaubauf et al. 2010). Schoch et al. (2012) showed that in reality ITS and LSU performed very similarly as barcodes, and that differences in these sequences correlated well with current species concepts in fungi. ITS is preferred in general, however – for the majority of fungal groups – due to its higher variability (especially in recent lineages of fungi) and higher level of success in PCR amplifications. Our GMYC results suggest that combining ITS and nLSU constitutes a more accurate barcode system for tremellalean species, coinciding with Fell et al. (2000) and Scorzetti et al. (2002). Moreover, only the nLSU gene tree supported the split between lineages A1 and A2 recovered by GCPSR. We therefore recommend the use of this marker, in addition to ITS, in future species delimitation studies in the Tremellales.

In summary, our results show that combining GCPSR with the GMYC model is a good strategy for identifying and delimiting independent fungal lineages, even in the case of limited sampling. Despite uncertainty in the selection of competing delimitation hypotheses, the combination of the two methods detects potential independent evolving lineages that might otherwise remain unnoticed. We recommend adopting the combination of GCPSR and the GMYC model in future studies of speciation in parasitic fungi, to provide comparable data from similar systems. Moreover, it will be interesting to test whether the combination of these two approaches can be applicable to other organisms and ecological situations too. Our study has also interesting implications for the understanding of the evolution modes in parasitic fungi and their contribution to the origin of diversity. We have shown here that, in addition to other lines of evidence
widely used in integrative taxonomy (such as DNA data, morphology, reproductive ability, geography, or ecology), understanding the evolutionary processes driving speciation (here host switching) is a valuable addition to support the distinction of species in fungal parasites.

ACKNOWLEDGEMENTS

We thank the curators of the herbaria sending material on loan, and Guillermo Amo de Paz, Gregorio Aragón, Brian Coppins, Pradeep K. Divakar, Jana Kocourková, Sonia Merinero, Sergio Pérez-Ortega, and Mercedes Vivas who very kindly provided specimens for the present study. The Laboratory of Molecular Systematics at the Swedish Museum of Natural History in general and Bodil Cronholm in particular are warmly thanked for excellent technical support. Raul García Camacho, Marcos Méndez, and David Sanchez Pescador provided key guidelines for the use of R, and Diego Fontaneto and Tomochika Fujisawa are thanked for discussions on the GMYC model. We are grateful to two anonymous referees for very helpful comments. The research was supported by grants from the Swedish Research Council (VR 621-2009-5372 and VR 621-2012-3990) and The Swedish Taxonomy Initiative to M.Wedin, the Spanish Ministry of Economy and Competitiveness (CGL2012-40123) to A.M. and M.Wedin, and the SYNTHESYS project SE-TAF 354, and Universidad Rey Juan Carlos (‘Estancias Breves’) Madrid, to A.M.


gilchristii, sp. nov. within the human pathogenic fungus Blastomyces dermatitidis.

PLoS ONE 8: e59237.


Clerc. P. 2006. Synopsis of Usnea (lichenized Ascomycetes) from the Azores with
additional information on the species in Macaronesia. Lichenologist 38:191–212.


Coevolution of Mimetic Heliconius Butterflies. PLoS ONE 7: e36464


Fell, J. W., T. Boekhout, A. Fonseca, G. Scorzetti, and A. Statzell-Tallman. 2000. Biodiversity and systematics of basidiomycetous yeasts as determined by large-


Phylogenetic species delimitation in ectomycorrhizal fungi and implications for
barcoding: the case of the *Tricholoma scalpturatum* complex (Basidiomycota).


between *Brachycaudus* and *Buchnera aphidicola*: bacterial genome helps define

Kemen, E., A. Gardiner, T. Schultz-Larsen, A. C. Kemen, A. L. Balmuth, A. Robert-
Seilanianzt, K. Bailey, E. Holub, D. J. Studholme, D. MacLean, J. D. G. Jones.
2011. Gene gain and loss during evolution of obligate parasitism in the white rust

Kirk, R. E. 1982. Experimental design: procedures for the behavioural sciences, 2nd

Strauss, and M. Gorfer. 2010. Molecular diversity of fungal communities in
agricultural soils from Lower Austria. Fungal Divers. 44:65–75.

Heidelberg, Germany.

Lawrey, J. D., and P. Diederich. 2003. Lichenicolous Fungi: interactions, evolution, and
biodiversity. Bryologist 106:80–120.

relationships between *Steinernema* spp. (Nematoda: Steinernematidae) and their
bacterial symbionts *Xenorhabdus* spp. (c-Proteobacteria: Enterobacteriaceae).

Syst. Parasitol. 77:1–12.


Pons, J., T.G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. P. Duran, S. Hazell, S.


FIGURE CAPTIONS

Fig. 1. Morphological variation of galls induced by Biatoropsis usnearum on Usnea and Protousnea thalli. Scale bars = 1 mm. Specimens are representatives of lineages shown in Figs. 2A and 3 (see also Supplementary Table S1 for details). A1: specimen AM190 (lineage ‘A1’); A2: specimen AM139 (lineage ‘A2’); B: specimen AA9 (lineage ‘B’); C: specimen AM203 (lineage ‘C’); D: specimen AA10 (lineage ‘D’); E: specimen AM213 (lineage ‘E’), F1: specimen AM10 (lineage ‘F1’); and F2: specimen AM166 (lineage ‘F2’).

Fig. 2. Fifty percent majority rule Bayesian consensus trees with average branch lengths from the combined analyses of ITS, nLSU and mtSSU datasets for Biatoropsis (A) and from the combined analyses of ITS and nLSU for their Usnea and Protousnea hosts (B). Black thick branches indicate Bayesian posterior probabilities (BPP) values ≥ 0.95 and ML and MP bootstrap values ≥ 70%. Grey thick branches indicate branches supported only by one or two methods. In these cases Bayesian posterior probabilities (BPP) are indicated over the branches and ML (left) and MP (right) bootstrap support values are indicated below the branches. Branch lengths are scaled to the expected number of nucleotide substitutions per site. 2A: For each Biatoropsis specimen, the geographical origin of the sample is indicated (see Supplementary Table S1). Coloured boxes in the Biatoropsis topology represent lineages A, B, C, D, E, and F, as potential species recovered using the GMYC single threshold model. Subdivisions of these lineages obtained using the GMYC multiple threshold model or the genealogy concordance phylogenetic species recognition (GCPSR) are also indicated along the tree (A1, A2, B1, B2, C, D, F1 and F2; see Fig. 3 and Supplementary Fig. S2). Speciation hypotheses
where each particular lineage is present are indicated in italics and within parentheses.

Asterisks indicate parasite lineages that are host-specific. **2B**: Colours in the *Usnea/Protousnea* phylogeny represent the mapping of the parasites on the host tree.

The names of the parasite lineages corresponding to each host are also indicated for each terminal of the host tree.

**Fig. 3.** Evolutionary reconstruction of the single-threshold (A) and multiple threshold (B) general mixed Yule-coalescent (GMYC) models, using the combined dataset described in Fig. 2A, i.e., ITS, nLSU and mtSSU, or each gene independently. The scale indicates relative time. The thresholds (i.e., the points where the transition from rates corresponding to speciation events to rates indicating coalescence events occurred most likely) are indicated with red vertical lines. The GMYC entities and clusters recovered are named as clades in Fig. 2A, or as subdivisions of these clades. Four columns to the right represent the coincidence or not of the lineages recovered when using single locus datasets. Black-lined white circles indicate lineages supported also using individual genes. Red-lined white circles indicate lineages recovered—but without support—using individual genes. Red crosses indicate conflicts in lineages obtained using the complete dataset or some of the independent genes. In this last case, different lineages recovered by single genes are indicated. A question mark in the mtSSU column indicates that DNA data is missing for this marker and lineage F2.

**Fig. 4.** Least costly coevolutionary reconstructions between *Biotaropsis* and *Usnea/Protousnea* phylogenies, achieved using Jane 4.0 and using pruned phylogenies including one representative per potential species of parasite and host. **A)** *Biotaropsis* species delimitation hypothesis **iii** (lineages recovered by GCPSR). **B)** *Biotaropsis*
species delimitation hypothesis (lineages recovered combining GCPSR and the GMYC multiple-threshold model) and their respective hosts (see Figs. 2 & 3, Table 5, and material and methods). Black branches represent the host phylogeny and blue branches the parasite phylogeny. Violet lines represent original polytomies resolved by Jane in order to minimize the overall cost of the solution. The cost regime utilized for the reconstruction was as follows: (cospeciation = 2, rest of events = 1, failures to diverge = 0) corresponding to cost regime 5 (Table 4). White filled circles represent cospeciations, yellow solid circles represent duplications, red solid circles represent host switches, dashed lines represent losses, and uneven (not straight) lines represent failures to diverge. Support values over 95% are indicated next to each event. These values give the percentage of solutions in which a specific event of the parasite tree is mapped to a given location on the host tree. The red arrow indicates a host switch recovered in 100% of the reconstructions that included lineage A2 – independently of the cost regime utilized (Table 5).
Fig. 1
A) Biatoropsis s. lat.

B) Usnea s. lat.