Evolution of nutritional modes of Ceratobasidiaceae (Cantharellales, Basidiomycota) as revealed from publicly available ITS sequences

Vilmar VELDRE, Kessy ABARENKOV, Mohammad BAHRAM, Florent MARTOS, Marc-Andre SELOSSE, Heidi TAMM, Urmas KÖJALG, Leho TEDERSOO

Article info

Article history:
Received 1 September 2012
Revision received 5 March 2013
Accepted 6 March 2013
Available online 24 April 2013
Corresponding editor:
Havard Kauserud

Keywords:
Ceratobasidium
Ectomycorrhiza
Evolutionary ecology
Orchid mycorrhiza
Phylogeny
Rhizoctonia
Saprotroph—pathogen continuum
Thanatephorus

Abstract

Fungi from the Ceratobasidiaceae family have important ecological roles as pathogens, saprotrophs, non-mycorrhizal endophytes, orchid mycorrhizal and ectomycorrhizal symbionts, but little is known about the distribution and evolution of these nutritional modes. All public ITS sequences of Ceratobasidiaceae were downloaded from databases, annotated with ecological and taxonomic metadata, and tested for the non-random phylogenetic distribution of nutritional modes. Phylogenetic analysis revealed six main clades within Ceratobasidiaceae and a poor correlation between molecular phylogeny and morphological–cytological characters traditionally used for taxonomy. Sequences derived from soil (representing putative saprotrophs) and orchid mycorrhiza clustered together, but remained distinct from pathogens. All nutritional modes were phylogenetically conserved in the Ceratobasidiaceae based on at least one index. Our analyses suggest that in general, autotrophic orchids form root symbiosis with available Ceratobasidiaceae isolates in soil. Ectomycorrhiza-forming capability has evolved twice within the Ceratobasidiaceae and it had a strong influence on the evolution of mycoheterotrophy and host specificity in certain orchid taxa.

© 2013 Elsevier Ltd and The British Mycological Society. All rights reserved.

Introduction

Fungi play a fundamental role in nutrient and carbon cycling in terrestrial ecosystems and sediments. They act as decomposers of dead organic matter, provide mineral nutrition to plants as mycorrhizal symbionts, or devastate plant populations as phytopathogens. Despite their major ecological and economic importance, the taxonomic and
The family Ceratobasidiaceae (Cantharellales, Basidiomycota) consists of the two closely related sexual genera, *Ceratobasidium* and *Thanatephorus* along with their *Rhizoctonia* asexual forms. They form one of such ‘cryptic’ fungal groups that play important ecological roles as crop pathogens, orchid mycorrhizal symbionts, saprotrophs and endophytes (Parmeter 1970, Sneh et al. 1996; Roberts 1999). These fungi spend most of their life cycle in the morphologically simple asexual (i.e. ‘anamorphic’) stage during which they can only be observed macroscopically as irregular sclerotia or, for phytopathogens, as necrotic lesions in the tissues of a tremendous range of host plants. Even when the sexual (i.e. ‘teleomorphic’) stage occurs, it represents an inconspicuous, fragile, web-like layer of generative hyphae covering living leaves and stems of hosts, or plant debris (Roberts 1999). Difficulties with the induction of fruiting in culture and the lack of morphological characters to distinguish among biological species have considerably complicated taxonomic and ecological research on these fungi.

To provide a practical method for identification, plant pathologists implemented a tractable test based on number of nuclei per cell and anastomosis compatibility in co-culture instead of morphology-based binomial taxonomy (Carling 1996; Sharon et al. 2008). Traditionally, strains of binucleate *Rhizoctonia* (BNR; binucleate strains are occasionally placed in the anamorphic genus *Ceratobasidium*) and multinucleate *Rhizoctonia* (MNR) are considered anamorphs of *Ceratobasidium* and *Thanatephorus*, respectively, although a few species of *Thanatephorus* are binucleate (Roberts 1999). Both the MNR and BNR are separated into anastomosis groups (AG), including 13 MNR AGs (AG-1 to AG-13) and 16 BNR AGs (AG-A to AG-S, excluding AG-J), AG-M and AG-N; Sharon et al. (2008), some of which are further divided into anastomosis subgroups based on more detailed analysis of anastomosis compatibility, morphology, pathogenicity and other criteria. All multinucleate AGs correspond to the teleomorphic species *T. cucumeris* (anamorph *Rhizoctonia solani*). The binucleate AG-A, AG-B(o), AG-D, AG-P and AG-Q correspond to *Ceratobasidium cornigerum*, whereas AG-Ba and AG-Bb correspond to *Ceratobasidium setariae*. However, telemorphs of the remaining binucleate AGs remain unknown (Roberts 1999).

Telemorphs of Ceratobasidiaceae are characterized by aseptate basidia (a derived, apomorphic feature), but also self-replicating spores and large indeterminate sterigmata (ancestral features; Roberts 1999). Numerous minor genera have been described within the Ceratobasidiaceae, but later synonymized with either *Ceratobasidium* (characterized by a single layer of basidia arising from horizontally branching hyphae of <10 μm in diameter) or *Thanatephorus* (characterized by multiple layers of basidia arising from vertically branching hyphae of >10 μm in diameter; Roberts 1999). Roberts (1999) included the genera *Waita* and *Scotomyces* in the Ceratobasidiaceae, but molecular phylogenetic analyses place both taxa outside Ceratobasidiaceae (Larsson 2007; K-H. Larsson, pers. comm. 02.08.2012). We, therefore, consider the Ceratobasidiaceae family to consist of the two genera *Ceratobasidium* and *Thanatephorus* only. The Ceratobasidiaceae is currently considered a peripheral member of the order Cantharellales, although its phylogenetic relations to other members of this group are not well resolved (Moncalvo et al. 2006; Hibbett et al. 2007).

Most attention to the members of the Ceratobasidiaceae has been sparked by their role as widespread soil-borne crop pathogens. Their necrotrophic capability is remarkably non-specific and affects a multitude of plant taxa. There is a continuum between the necrotrophic parasitism and saprotrophy in the Ceratobasidiaceae. Most strains possess some saprotrophic capability, but aggressive pathogens are poor saprotrophic competitors and depend on nutrients acquired from living plant tissue (Papavizas 1970). By contrast, strains that have lost their ability to infect and cause serious damage to living organs, may function as commensal or even mutualistic endophytes (i.e. fungi that grow diffusely in tissues, without causing any visible symptoms; Sen et al. 1999) and increase their hosts’ resistance to pathogenic strains (Sneh 1998). However, little is known about the frequency and role of Ceratobasidiaceae endophytes (Sen et al. 1999).

As a further major interaction, the Ceratobasidiaceae includes a large number of taxa that form orchid mycorrhiza (OrM). Orchids have an unusual relationship with their mycorrhizal fungi compared to other mycorrhizal plants: they receive all nutrients, including carbon, from their fungi during their heterotrophic germination (Smith & Read 2008; Dearnaley et al. 2012). At the adult stage, most orchids develop photosynthetic capability, but clearly continue to benefit from their mycobiota by receiving mineral nutrients. Some species associated with *Ceratobasidium* allow a net carbon flow from adult orchids to the fungus (Cameron et al. 2006, 2008), while in some other species, adult orchids still obtain carbon from fungi (Selosse & Roy 2009; Yagame et al. 2012). Thus, the balance between costs and benefits for fungi remains poorly understood in OrM (Dearnaley et al. 2012), and we treat here the association between autotrophic orchids and their mycorrhizal fungi as symbiotic along the mutualism–parasitism continuum (Egger & Hibbett 2004). As a side finding in the research on OrM, it was discovered that some Ceratobasidiaceae isolates also form ectomycorrhiza (EcM) (Warcup 1991; Yagame et al. 2008, 2012; Bougoure et al. 2009), and indeed some Ceratobasidiaceae have been sporadically reported in community analyses of EcM fungi (e.g. Rosling et al. 2003). However, the phylogenetic and biogeographic distribution of these relatively uncommon EcM groups remains poorly understood (Tedersoo et al. 2010).

Development of molecular methods has led to rapid unravelling of the systematics and ecology of microbes in the past few decades. Barcoding with the Internal Transcribed Spacer (ITS) region of the ribosomal DNA locus has become a standard means of identification in fungi (Schoch et al. 2012). ITS sequences offer suitable resolution for identification of the Ceratobasidiaceae strains and provide an invaluable tool for differentiating AGs and their subtypes both from cultured strains, soil and plant tissue (Johnson et al. 1998; Sharon et al. 2006). Phylogenetic analyses suggest that most AGs are monophyletic, but still correspond to complexes of molecular and ecological species (Gonzales et al. 2001, 2006; Sharon et al. 2006, 2008). AG subtypes usually have different ecological niches and their within-group ITS sequence similarity is over
Materials and methods

Data assembly

All available full-length ITS sequences belonging to the Ceratobasidiaceae were retrieved from INSD in Jun. 2010. A direct search by taxonomic assignment to the family Ceratobasidiaceae or the genera Ceratobasidium, Thanatephorus, Rhizoctonia or Ceratoderma in GenBank (3 167 sequences), search for matching unidentified sequences (299 additional sequences) using the web tool emerencia (Nilsson et al. 2005) and the 1 000 best BLAST matches for highly deviating sequences AJ633124 (49 matching sequences) and AB000014 (one matching sequence) retrieved a total of 3 516 INSD entries. In addition, 84 sequences were included from the UNITE database. These were supplemented with a reference dataset of 946 partial (at least 100 bp missing) or low-quality sequences representing other taxonomic groups or genes, 26 highly divergent sequences that potentially belong to Ceratobasidiaceae and 946 partial (at least 100 bp missing) or low-quality sequences were excluded. Sequences with low read quality or of chimeric nature were tagged accordingly in the INSD copy of the UNITE database as implemented in the PlutoF workbench (Abarenkov et al. 2010b). The remaining 2 257 nearly full-length ITS sequences formed the ‘full conservative’ dataset that is central in the subsequent analyses. Because AG-D-II and AG-H were only represented by notoriously low-quality or incomplete sequences, a ‘liberal’ dataset including an additional 285 ITS sequences was also analyzed. In addition, all 44 nuclear rDNA Large Subunit (LSU) sequences (at least 900 bp) of the Ceratobasidiaceae were retrieved from INSD and UNITE. These were supplemented with a reference dataset of 95 sequences covering the Cantharellales and most other orders of the Agaricomycotina to determine a suitable outgroup for the ITS-based analysis.

Metadata on isolation source, interacting taxon, locality and AG were compiled for all full-length ITS sequences from INSD and associated publications (Table S1). Isolation source and substrate formed a basis for the statistical metadata analysis. The main sources of isolation included diseased tissue of crop plants (implying pathogenic interaction), roots of orchids (OrM interaction) and soil (see below), followed by ectomycorrhizal root tips (EcM interaction) and healthy non-mycorrhizal root or symptomless above-ground tissue of wild plants (endophytic interaction; Sen et al. 1999). Some non-mycorrhizal isolates from crop plants were found to be non-pathogenic or even beneficial in experimental studies, and were therefore treated as endophytes. Soil-derived sequences were separated into entries obtained from soils of: (1) crop fields by plant pathologists in studies targeting pathogens; or (2) natural or semi-natural ecosystems addressing soil fungal communities. The first category was considered suggestive of saprotrophy, although we anticipate that pathogenic and EcM strains may also be present in the soil of natural habitats. All annotated metadata are publicly available in the UNITE database and through a link-out function in the European Nucleotide Archive (ENA).

Phylogenetic analyses

The full-length ITS sequences ranged from 580 to 680 nucleotides. The final alignments subjected to phylogeny reconstruction were created with PRANKSTER (Löytynoja & Goldman 2005) using default parameters, and corrected manually. Two segments of ITS1 (ca. 15–30 and ca. 20–60 bp, varying among sequences) and one segment of ITS2 (ca. 10–20 bp) were omitted from phylogenetic analyses, because these were highly variable and could not be reliably aligned across the whole dataset. The length of individual sequences was thereby reduced to 540–580 nucleotides. The final full alignment spanned 1 178 sites, of which 571 were variable and 444 were informative. For Maximum Likelihood (ML) and
Bayesian phylogeny reconstruction, RAxML 7.2.8 (Stamatakis 2006) and MrBayes 3.1 (Husonbeck & Ronquist 2001) were respectively used. Only RAxML was able to handle the full dataset with reasonable speed. In all RAxML analyses, GTR + I + I evolutionary model and 1 000 fast bootstrap replicates were used. To use both methods for comparison to confirm validity of tree topology and to reduce tree size for convenient display, the dataset was collapsed into 288 entries by clustering the sequences with BLASTclust (Biegert et al. 2006) based on 99 % sequence similarity and 90 % coverage criteria (99 % threshold was used instead of 97 %, to keep all anastomosis subgroups in separate clusters). The longest sequence of each cluster served as a representative sequence. Phylograms were constructed from this ‘collapsed conservative’ dataset with both RAxML and MrBayes (burn-in = 2 000; evolutionary model GTR + I + I as revealed from Modeltest; Posada & Crandall 1998). All Bayesian analyses in this study were run for 10 million generations sampled every 1 000 generations. Burn-in value was determined according to at which generation the log likelihood scores reached stationary level. To include all ecological metadata in a single tree, a phylogram was constructed directly from the ‘full conservative’ dataset with RAxML. Patristic distances (pairwise total path length between two terminals) were exported from this full dataset tree using PDAP package of Mesquite 2.75 (Midford et al. 2003; Maddison & Maddison 2009), and subjected to statistical analyses regarding the nutritional mode of sequenced isolates (see below). Trees were visualized together with metadata using the online tool iTOL (Letunic & Bork 2007; http://itol.embl.de/), and are publicly available as shared projects of the user ‘CeratobasidiuinThanatephorus’ (case sensitive). For improved readability, clade names of the Ceratobasidiaceae are labelled based on the phylogenetic distribution of known AGs preceded by a slash (Moncalvo et al. 2002).

The LSU sequences were aligned with PRANKSTER using default parameters and minimal manual correction. The LSU alignment spanned 1 622 sites, of which 706 were variable and 517 were informative. LSU phylograms were generated both with RAxML and MrBayes (GTR + I evolutionary model as revealed from Modeltest) to assess the monophyly of the Ceratobasidiaceae and to obtain an outgroup for rooting ITS phylograms. All alignments are available in TreeBase (accession TB2:S13952).

To remove pseudoreplicates from the analyses of distribution of nutritional modes, we excluded redundant sequences from each study by keeping the longest representative sequence from each MOTU (barcoding threshold of 97 % sequence similarity and 90 % of coverage) per study, resulting in 508 representative sequences. The Net Relatedness Index (NRI, standardized mean of distances between all pairs of sequences from compared categories) and the Nearest Taxon Index (NTI, standardized mean of distances between pairs of the nearest neighbours) were calculated based on patristic phylogenetic differences of an ultrametric tree to test clustering and overdispersion of nutritional modes on larger and smaller phylogenetic scales, respectively (Webb 2000). Sampling species as random draws from the phylogeny without replacement served as a null model and statistical significance was calculated based on 999 permutations as implemented in Phylocom (Webb et al. 2008). To address phylogenetic distinctness of nutritional modes, we calculated the pairwise Nearest Taxon distances (NTD) for all these ecological guilds. Differences between observed and randomized NTD that exceeded 2 SD of the randomized dataset were considered statistically significant. In addition, pairwise UniFrac distances were calculated in the Fast UniFrac server (http://bmf2.colorado.edu/unifrac/) by use of 999 permutations. UniFrac distance metric measures the phylogenetic distance among communities by calculating the proportional length of the tree branches that lead to descendants from each single community but not from both communities, and tests whether this distance is different from a random draw from the communities (Hamady et al. 2010). All these distance metrics were calculated with and without the pathogen-rich/mainly-MNR clade to assess the robustness of results. Significance values of the UniFrac pairwise distances were subjected to reduction of false discovery rate by use of Benjamini–Hochberg FDR correction.

Based on the ML phylogram of the ‘full conservative’ dataset, we generated ancestral state reconstructions for major clades and well-supported subgroups as implemented in the Ape package of R (Paradis et al. 2004). We particularly focused on the evolution of ECM habit within the Ceratobasidiaceae, because ECM sequences were mainly concentrated in two subclades. All nutritional modes were allowed to change to any other state at different probabilities, because no a priori model exists.

Biogeographic patterns and host specificity of the most common MOTUs and major clades were assessed based on both the full and non-redundant datasets. Because of highly biased sequence availability in different countries and hosts, we sought the patterns of endemism on a continental scale and host specificity at the plant tribe (OrM) and family (pathogens) levels.

**Results**

**Taxonomic and phylogenetic patterns**

The Bayesian (ASDSF = 0.011) and ML analyses revealed similar tree topology in the LSU-based phylogenetic reconstruction. The LSU-based phylogeny confirmed monophyly of the Ceratobasidiaceae within Cantharellales (Fig 1). The Tulasnellaceae clade was inferred as a sister group to the Ceratobasidiaceae, but with no statistical support. Within the Ceratobasidiaceae, Thanatephorus (syn. Uhatobasidiun) fuisporus formed a basal branch that was supported in both the Bayesian (PP = 1.0) and ML (BS = 95) analyses. Only the /BD clade was not represented by LSU sequences, but it was nested among four other clades that were only partly supported (PP = 0.76; BP = 75) as a sister group to the /fuisporus clade based on the ITS analysis (Fig 2). Therefore, we rooted the ITS phylogram at the /fuisporus clade. The topology of the LSU phylogram was congruent with the ITS phylogram. The /mainly-MNR clade was monophyletic and nested within the BNR clades in both the LSU- and ITS-based analyses.

In the ‘collapsed conservative’ ITS dataset, phylograms revealed in both the ML and Bayesian analyses (ASDSF = 0.019)
Fig 1 – Large subunit maximum likelihood phylogram demonstrating the phylogenetic placement of Ceratobasidiaceae among Agaricomycotina. Tremella mesenterica serves as outgroup. Values above branches indicate bootstrap support ≥70%, while thick branches indicate high posterior probabilities (PP ≥ 0.95) as revealed from a parallel Bayesian analysis. Bar indicates 0.1 substitutions per site.
Fig 2 — ITS maximum likelihood phylogram of the reduced conservative dataset, ancestral states of selected nodes and frequency of occurrence in different nutritional modes. The phylogram was rooted at the Fusisporus clade. Values above branches indicate bootstrap support ≥ 70%, while thick branches indicate high posterior probabilities (PP ≥ 0.95) as revealed from a parallel Bayesian analysis. Bar indicates 0.05 substitutions per site. Circular diagrams below branches indicate the probability of nutritional modes being ancestral: red, pathogenic (pathog); black, saprotrophic (sapr); blue, orchid mycorrhizal (OrM); yellow, ectomycorrhizal (EcM); purple, endophytic (endoph); white, unknown (unkn). The table of counts of nutritional modes is based on the full conservative dataset — the first sub-column excluding and second sub-column including redundant sequences (99% similarity within a study).
produced comparable results that differed only in statistical support (Fig 2). Phylograms of the ‘full conservative’ dataset displayed slight differences in tree topology and weaker branch support, perhaps due to a greater amount of noise owing to accumulation of erroneous base calls in larger datasets.

All ITS-based phylogenetic analyses resulted in similar tree topology and revealed six major clades (Fig 2). All MNR AGs were clustered in one well-supported clade (‘mainly-MNR’) together with some BNR AGs (AG-E, AG-F, AG-P, AG-R, AG-S). The remaining BNR sequences were divided between other clades, four of which included known AGs (‘AK, ‘BD, ‘CHI and ‘GLO). The /fusisporus clade was not associated with any described AGs, but contained a sequence of the binucleate teleomorph T. fusisporus (DQ398957). The ‘mainly-MNR, ‘AH and ‘CHI clades were supported in the Bayesian analysis (PP ≥ 95 %), whereas the ‘mainly-MNR, ‘AK, ‘CHI and /fusisporus clades received bootstrap support over 70 % in the ML analysis. Thus, the monophyly of the ‘BD and ‘GLO clades is poorly supported despite their consistent formation in all different analyses and datasets.

We found no clear distinction between the teleomorph genera Thanatephorus and Ceratobasidium (Fig S1). While most Thanatephorus fruit-bodies (predominantly T. cucumeris) were placed in the ‘mainly-MNR clade, fruit-bodies with Thanatephorus morphology also appeared in the binucleate ‘GLO.
Distribution of sequences and MOTUs (97 % similarity threshold; in parentheses) representing different C. cornigerum specimens) belonged to the BNR clades, but four AG-B and AG-D that in turn contained monophyletic subgroups /mainly-MNR clade (among AG-E, AG-P, AG-R and AG-4-HGI).

Most of the AGs were monophyletic, including AG-1, AG-4, AG-B and AG-D that in turn contained monophyletic subgroups (BS > 70; Fig 2, Fig S2). By contrast, the AG-Fa formed a distinct monophyletic group (BS = 82) near the base of the /mainly-MNR clade, whereas AG-Fb constituted a subgroup of the multifluclate AG-6. AG-2 was polyphylectic as a whole, but each of its subgroups was monophyletic. Although AG-C, AG-I and AG-7 have not been separated into subgroups, members of these AGs were distributed in seemingly unrelated subclades within the /CHI (AG-C, AG-I) and /mainly-MNR (AG-7) clades. AG-S and AG-Q were represented by only one and two sequences, respectively, and therefore the monophyly of these AGs cannot be assessed. A large proportion of the /mainly-MNR clade was covered with known AGs (Fig 2B), while the majority of terminal taxa of the BNR clades lacked anastomosis grouping information (Fig 2A).

Ecological patterns

Based on metadata, putative pathogens, OrM symbionts, saprotrophs and Ecm symbionts were represented by 1 129 (50.0 %), 401 (17.8 %), 69 (3.1 %) and 73 (3.2 %) sequences, respectively (Table 1). The putative ecology of 597 (26.5 %) sequences remained unknown due to the lack of metadata about the isolation source. In total, 82.9 % of sequences of putative pathogens were affiliated to the /mainly-MNR clade. Only 14.3 % of sequences representing other ecological strategies belonged to this clade. Sequences of putative Ecm fungi clustered either in the /CHI or in the /GLO clade. A group of Australian Ecm isolates sampled by Bougoure et al. (2009) was exceptional as each of these strains comprised two ITS copies, one clustering with other Ecm sequences in the /GLO clade and another clustering with non-Ecm sequences in the /BD clade. Phylogenetic analyses and ancestral state reconstructions suggested that Ecm habit has evolved twice in the Ceratobasidiaceae (P < 0.01), but revealed no ancestral mode of nutrition with confidence, except that Ecm habit is unlikely to be ancestral (Fig 2A).

Based on sequence clustering at 97 % similarity cut-off, the ‘full conservative’ dataset was divided into 157 MOTUs of which 53 (33.8 %) were represented by a single sequence and 23 (14.6 %) were represented by only two sequences. Damaged plant tissue, orchid roots, soil, Ecm root tips, and healthy shoot tissue or non-mycorrhizal root tissue were the dominant isolation source in 52 (33.1 %), 47 (29.9 %), 14 (8.9 %), 14 (8.9 %) and 7 (4.5 %) MOTUs, respectively (Table 1). The main isolation source of 13 (8.2 %) MOTUs remains unknown, whereas sequences of ten MOTUs (6.4 %) were obtained from multiple sources in equal abundance (usually a combination of pathogenic, OrM and/or soil-derived sequences). The ‘liberal’ dataset was divided into 252 MOTUs, including an additional 84 singletons, eight doubletons, and three taxa with more than two sequences.

The six major clades of the Ceratobasidiaceae differed in the available taxonomic information and relative sampling depth. The /mainly-MNR clade comprised 68.4 % of sequences, but only 38.9 % of MOTUs, indicating relative oversampling of harmful pathogens by plant pathologists. Anastomosis grouping information was available for 64.0 % of MOTUs in the /mainly-MNR clade, but only for 16.4 % of species in the BNR clades taken together. No AGs were assigned to members of the /fusisporus clade.

All nutritional modes except endophytes (0.1 < P < 0.5) were significantly phylogenetically clustered compared to the null model based on the NRI index and inclusion or exclusion of the /mainly-MNR clade (P < 0.05 in all cases; Table 2). Based on NTD index, all nutritional modes except pathogens (P > 0.5) were significantly phylogenetically clustered. In all analyses, OrM fungi displayed the strongest phylogenetic clustering among all modes of nutrition. The NTD pairwise distance metric revealed several phylogenetic dissimilarities among the nutritional modes (Table 2). OrM sequences differed from pathogens (NTD = −20.63; P < 0.001) and Ecm fungi (NTD = −3.36; 0.01 < P < 0.05), but not from endophytes and saprotrophs (0.1 < P < 0.5). Ecm fungi were significantly different from pathogens (NTD = −19.15; P < 0.001) and saprotrophs (NTD = −3.36; 0.01 < P < 0.05), whereas endophytes differed significantly only from pathogens (NTD = −5.59; 0.001 < P < 0.01). Saprotrophs were phylogenetically clearly distinct from pathogens (NTD = −8.18; 0.001 < P < 0.01). The Unifrac pairwise distance metric corroborated these findings of phylogenetic differentiation (i.e. separation of communities; Table 2).

Biogeographic patterns and host specificity could be addressed in sufficient detail only in the seven most common MOTUs (found in more than five studies; representing AG-1-A, AG-2-1, AG-2-2, AG-4-HGI, AG-4-HGI, AG-5, AG-A) due to

| Table 1 – Distribution of sequences and MOTUs (97 % similarity threshold; in parentheses) representing different nutritional modes among six major clades of Ceratobasidiaceae. For MOTUs, only the most common nutritional mode is scored |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| /mainly-MNR     | 0 (0) | 3 (1) | 47 (11)| 936 (34)| 26 (3) | 531 (9) | - (3) | 1 543 (61) |
| /AK             | 0 (0) | 4 (0) | 1 (0)  | 86 (4)  | 4 (0)  | 18 (0)  | - (0) | 113 (4)   |
| /BD             | 13 (1)| 2 (2) | 151 (13)| 34 (6)  | 5 (3)  | 24 (1)  | - (0) | 229 (26)  |
| /CHI            | 15 (7)| 7 (3) | 90 (9) | 21 (4)  | 25 (3) | 13 (3)  | - (5) | 171 (34)  |
| /GLO            | 45 (6)| 14 (1)| 46 (6) | 52 (9)  | 8 (4)  | 9 (0)   | - (0) | 174 (25)  |
| /fusisporus     | 0 (0) | 0 (0) | 24 (6) | 0 (0)   | 2 (0)  | - (0)   | 27 (7) |
| Total           | 73 (14)| 30 (7)| 359 (47)| 1 129 (52)| 69 (14)| 597 (13) | - (10)| 2 257 (157) |
limited sample size. All these MOTUs were distributed in several continents and most of these displayed no evidence for host specificity, infecting a wide range of plant families (Fig S1). Only AG-3 showed some host specificity, infecting predominantly species of the Solanaceae family, particularly potato (Solanum tuberosum). Anastomosis groups that were relatively widely distributed in the phylogenetic tree (AG-C, AG-I, AG-2, AG-7) all had wide geographic distribution in the Northern hemisphere, with sympatric occurrence of MOTUs. In OrM, there was no evidence for specificity between orchid tribes and MOTUs or major clades of the Ceratobasidiaceae (Fig S1). Because most orchid species were each represented by only a single study and geographic origin, we cannot address OrM specificity at orchid species level. At higher taxonomic level, most of the six major clades had cosmopolitan distribution and none of them were restricted to a single continent or biome (Fig S1).

Discussion

Phylogenetic implications

Phylogenetic analyses revealed that the Ceratobasidiaceae is monophyletic within the Cantharellales and the /mainly-MNR clade is nested within BNR groups. Within Ceratobasidiaceae, the /fusisporus clade was inferred as a basal branch given the position of T. fusisporus in the LSU phylogram. Therefore, we rooted the ITS phylogram at the /fusisporus clade, but conservatively anticipate the unlikely possibility that the true rooting point lies within this clade. Use of protein-encoding genes will be necessary for further in-depth phylogenetic reconstruction of the Ceratobasidiaceae (Gonzales et al. 2006).

We found strong phylogenetic evidence for non-monophyly in both the teleomorph genera and the nuclear count-based system of anamorphs. Species of the two teleomorph genera Ceratobasidium and Thanatephorus are subtle and intermediate forms exist (Roberts 1999), so that the characteristic features may reverse in Ceratobasidiaceae evolution. Our data support that the simplest basidioma shape characteristic for Ceratobasidium (Roberts 1999) is likely ancestral and at least evolved several times into the more complex Thanatephorus basidioma shape. Although all MNR strains were representative saprotrophs, it seems that orchids preferentially change is evident based on the sister relationship between AG-1 and AG-P. Although most AGs were monophyletic, large AGs usually consisted of multiple MOTUs that were not closely related (as low as 81 % ITS sequence similarity between different subgroups of the same anastomosis group; Gonzales et al. 2001; Sharon et al. 2006, 2008). The topology found suggests that the multinucleate organization is a derived feature that arose multiple times in the Ceratobasidiaceae evolution, and/or with many reversions to the binucleate state. While there seems to have been one main shift from binucleate to multinucleate organization, at least one more change is evident based on the sister relationship between AG-1 and AG-P. Although most AGs were monophyletic, large AGs usually consisted of multiple MOTUs that were not closely related (as low as 81 % ITS sequence similarity between different subgroups of the same anastomosis group; Gonzales et al. 2001; Sharon et al. 2006, 2008). The polyphyly of teleomorph genera and some AGs as well as multiple shifts in the number of nuclei imply that both the morphology-based and anastomosis behaviour-based taxonomies lack species-level resolution in the Ceratobasidiaceae. The limited number of known morphological characters available for fungi, especially for taxa with no or resupinate fruit-bodies, often results in lower taxonomic resolution when using morphology, as compared to molecular features (Taylor et al. 2006).

Evolution of nutritional modes

Although represented in several major clades, both the pathogenic, OrM, EcM, endophytic and putatively saprotrophic isolates displayed phylogenetically clustered distribution within the Ceratobasidiaceae. This indicates that these ecological roles tend to be evolutionarily conserved. Sequences derived from OrM and soil formed a phylogenetically coherent group, and the fact that they originated from studies with different aims and sampling protocols emphasizes this result. Assuming that soil-derived isolates mostly represent saprotrophs, it seems that orchids preferentially

Table 2 - Phylogenetic clustering and pairwise phylogenetic distances of nutritional modes. Unifrac pairwise distances indicate phylogenetic dissimilarity and these are given to the left from the diagonal; nearest taxon (NT) pairwise distances are given to the right from the diagonal. For net Relatedness index (NRI), nearest taxon index (NTI) and NT, positive values denote phylogenetic clustering and negative values indicate overdispersion (or avoidance, NT). Significant values are highlighted in bold.

<table>
<thead>
<tr>
<th>n³</th>
<th>NRI</th>
<th>NTI</th>
<th>Pairwise UnifracNTD (comdistNT) phylogenetic distances</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcM</td>
<td>18</td>
<td>3.46</td>
<td>5.34</td>
</tr>
<tr>
<td>Endophyte</td>
<td>19</td>
<td>1.94</td>
<td>0.83</td>
</tr>
<tr>
<td>OrM</td>
<td>82</td>
<td>7.44</td>
<td>6.87</td>
</tr>
<tr>
<td>Pathogen</td>
<td>216</td>
<td>-1.58</td>
<td>2.64</td>
</tr>
<tr>
<td>Saprotroph</td>
<td>35</td>
<td>3.57</td>
<td>2.67</td>
</tr>
<tr>
<td>Unknown</td>
<td>138</td>
<td>-1.58</td>
<td>3.17</td>
</tr>
</tbody>
</table>

a The number of non-redundant sequences (one sequence per MOTUs per study).

Note: The number of non-redundant sequences (one sequence per MOTUs per study).
establish symbiosis with the less aggressive side of the pathogen-saprotroph continuum in the Ceratobasidiaceae. One could conclude that association with pathogenic strains could result in death of the orchids before they germinate or reach maturity and become available for sampling (Taylor et al. 2003). However, some orchids germinate as successfully with isolates from strongly pathogenic AGs as with those from non-pathogenic or weakly pathogenic AGs in vitro (Masuhara et al. 1993, Masuhara & Katsuya 1994; Pope & Carter 2001). Alternatively, pathogenic isolates could be simply less available for orchids in soil, because pathogens tend to be more limited to tissues and rhizosphere of infected plants. We tested the latter hypothesis by separately considering only the MOTUs found from soil. There was no significant difference in the association of OrM isolates with non-pathogenic or pathogenic MOTUs (chi-square test: n = 29; P = 0.775), which lends no support for an avoidance strategy. Therefore, our analyses suggest that in general, orchids associate with available Ceratobasidiaceae strains in soil irrespective of their pathogenicity.

Sequences of plant pathogens were phylogenetically distinct from sequences of all other nutritional modes. Phylogenetic differences between soil-derived sequences and pathogens may represent a trade-off between abilities to compete for nutrients in debris dispersed in soil and to attack living plant tissue. To our knowledge, phylogenetic distinctness of saprotrophic and parasitic guilds has not been explicitly tested at the genus or family level in other taxonomic complexes of fungi, but this could be a common phenomenon in basidiomycetes as suggested by published phylograms (e.g. Cryptococcus: Findley et al. 2009). Within Ceratobasidiaceae, pathogenic groups were potentially derived from putative soil saprotrophs in several instances (Sharon et al. 2006). One could argue that soil-derived isolates are non-pathogenic to most plants in natural conditions where they evolved, but may become aggressive against fertilized crop plants or when introduced to exotic habitats (Wingfield et al. 2001; Desprez-Loustau et al. 2007). This is certainly an uncommon phenomenon in the Ceratobasidiaceae family, because pathogenic isolates are relatively well-sampled.

Bougoure et al. (2009) reported that several cultured strains contained two alleles (cf. ‘cistrons’) that shared only 80% ITS sequence similarity. While one of the alleles belonged to the /ceratobasidium1 EcM lineage in the /GLO clade, the other allele was nested in the distantly related /BD clade, within other OrM and soil-derived isolates. This may represent a case of hybridization or gene transfer between an EcM fungus and a distantly related non-mycorrhizal fungus (Bougoure et al. 2009). Some anastomosis groups (AG-1, AG-2, AG-B) encompass lineages with only 81–83% similarity (Sharon et al. 2006, 2008), which may allow the coexistence of distantly related nuclei and thus create a possibility for recombination events (Xie et al. 2008). Such hybridization within sympatric, anastomosis-compatible groups may result in shifts in host range and give rise to novel harmful pathogens (Brasier 2000; Desprez-Loustau et al. 2007).

Sequences isolated from EcM root tips fell consistently into two well-supported lineages in the /BNR /GLO and /CHI clades. Ancestral state reconstructions support these findings indicating that EcM lifestyle has secondarily arisen twice in the Ceratobasidiaceae. Tedersoo et al. (2010) referred to these EcM fungal lineages as /ceratobasidium1 and /ceratobasidium2, respectively, but provided no phylogenetic support for this hypothesis. Present data indicates that the /ceratobasidium1 lineage includes both EcM and OrM strains from SW Japan, Thailand, Australia, Malaysia, Zambia and Madagascar, suggesting a subtropical and tropical distribution. The Japanese (Yagame et al. 2008) and Australian (Bougoure et al. 2009) OrM strains readily form EcM in experimental conditions and participate in tripartite interactions with mycoheterotrophic (non-photosynthetic) orchids in nature. The evolution of EcM habitat in this lineage may have facilitated the development of mycoheterotrophy in several orchid taxa (Chamaeagastrodia sikokiana and Rhizanthella gardneri), because mycoheterotrophic orchids usually associate with EcM fungi in temperate habitats. It is believed that EcM fungi provide a more stable and reliable carbon source compared to saprotrophic fungi (Dearnaley et al. 2012). The /ceratobasidium2 lineage represents a collection of sequences from EcM root tips in several community analyses in the boreal and temperate forests of Europe, North America and Japan, suggesting a circumboreal distribution. Recently, Yagame et al. (2012) showed that Platanthera minor, a partially mycoheterotrophic (= mixotrophic) green orchid, associates with both the /ceratobasidium1 (cf. types 11 and 12) and /ceratobasidium2 EcM lineages (I4). Both EcM lineages lack a sister group with well-defined nutritional mode, and the ancestral state reconstruction was unable to resolve the ancestral mode of nutrition for them.

Biodiversity and biogeography

Sequence clustering reveals that both the /mainly-MNR and especially BNR clades of the Ceratobasidiaceae comprise high cryptic diversity of taxa that lack sequenced representative cultures and fruit bodies. Culturing and anastomosis tests of pathogenic isolates have been performed for decades by phytopathologists, but soil-derived and orchid root-associated strains have usually escaped these trials, although they are often isolated. Moreover, recent in situ molecular identification studies from roots and soil have provided an unprecedented wealth of information about the diversity of BNR clades, suggesting that many Ceratobasidiaceae taxa have never been obtained into pure culture. Both the ‘full conservative’ and ‘liberal’ datasets revealed a large number of MOTUs represented by only one or two sequences, indicating that part of the uncultured richness is yet to be captured by molecular techniques. Because sequence analysis of soil and root material is performed mostly in boreal and temperate regions, tropical soils probably contribute to a large proportion of the undetected richness and understanding of biogeographic patterns. For example, the ‘full conservative’ dataset included 11 Ceratobasidiaceae MOTUs from 15 orchid species in Réunion Island (Martos et al. 2012). The present evaluation of biogeography and host specificity was restricted to the most common MOTUs of pathogens and OrM fungi and to the six major clades. While the major clades have cosmopolitan distribution, the common species lack both host specificity and endemism that may be at least partly ascribed to anthropogenic dispersal for crop pathogens. We cannot rule
out the possibility that host specificity occurs at the strain level as shown for the Fusarium–Gibberella complex (Ma et al. 2010). In OrM fungi, the lack of specificity for plant groups is consistent with the facultative nature of this symbiosis for fungi, as revealed by the lack of phylogenetic fidelity on the fungal side (Martos et al. 2012).

The use of only ITS sequences for inferring biogeographic and ecological questions has, however, a few limitations. First, tracing the origin and specificity of pathogens is best approached by use of population genetics techniques, because the ITS region has insufficient resolution at the population level. Second, ITS-based phylogenetic trees often exhibit low phylogenetic resolution because of abundant insertions and deletions that are difficult to handle by alignment and phylogenetic programs. Another source of error is the paucity of metadata. Both phylogenetic uncertainty and missing metadata render the results of evolutionary ecology studies less statistically supported due to greater noise to signal ratio and reduced sample size, respectively. However, such noise is unlikely to bias the qualitative patterns when it is distributed randomly along the phylogram.

Conclusions

Our global analysis of public ITS sequences of the Ceratobasidiaceae family sheds light onto phylogenetic relations and distribution of ecological strategies within this large, ecologically and economically important fungal family. All major nutritional modes such as saprotrophy, pathogenic, OrM, EcM and endophytic interactions were phylogenetically conserved. Although pathogens have arisen multiple times independently (Gonzales et al. 2001, 2006), they are phylogenetically distinct from most other functional guilds. Orchid root symbionts are phylogenetically overlapping with putative saprotrophs from soil samples, suggesting that saprotrophic strains from natural soils are more easily accessible for orchids. EcM lifestyle has evolved separately in two major clades of Ceratobasidiaceae. Probably through improved carbon nutrition, at least one of these events may have triggered loss of photosynthesis in certain orchid taxa associating with these clades of Ceratobasidiaceae. Our study emphasizes that the Ceratobasidiaceae is both functionally and taxonomically highly diverse, and that the classical morphological, cytological and anastomosis investigations have limited power of resolution at both the species and genus levels. Further sampling in natural habitats, especially in tropical sites, will probably enable us to recover many novel species and to address biogeographic patterns within the Ceratobasidiaceae.

Acknowledgements

This work was initially led by V. Veldre, but was finalized by co-authors following the accidental death of the first author. We thank A. Stamatakis and A. Aberer for advice regarding phylogenetic analyses. The bulk of this project was funded from Estonian Science Foundation grants PUT171, 8235 and 9286, and FIBIR. M.-A. Selosse is funded by the Agence Nationale de la Recherche (ANR program SYSTRUF), and F. Martos by the Région Réunion.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.funeco.2013.03.004.

References


