ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

Molecular Phylogenetics and Evolution xxx (2006) xxx–xxx

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.elsevier.com/locate/ympev

Phylogenetic relationships, host affinity, and geographic structure of boreal and arctic endophytes from three major plant lineages

K. Lindsay Higgins [1], A. Elizabeth Arnold *, Jolanta Miadlikowska, Snehal D. Sarvate [2], François Lutzoni

Department of Biology, Duke University, Durham, NC 27708, USA

Received 18 August 2005; revised 3 July 2006; accepted 19 July 2006

Abstract

Although associated with all plants, fungal endophytes (microfungi that live within healthy plant tissues) represent an unknown proportion of fungal diversity. While there is a growing appreciation of their ecological importance and human uses, little is known about their host specificity, geographic structure, or phylogenetic relationships. We surveyed endophytic Ascomycota from healthy photosynthetic tissues of three plant species (*Huperzia selago*, *Picea mariana*, and *Dryas integrifolia*, representing lycophytes, conifers, and angiosperms, respectively) in northern and southern boreal forest (Québec, Canada) and arctic tundra (Nunavut, Canada). Endophytes were recovered from all plant species surveyed, and were present in <1–41% of 2 mm^2 tissue segments examined per host species. Sequence data from the nuclear ribosomal internal transcribed spacer region (ITS) were obtained for 280 of 558 isolates. Species-accumulation curves based on ITS genotypes remained non-asymptotic, and bootstrap analyses indicated that a large number of genotypes remain to be found. The majority of genotypes were recovered from only a single host species, and only 6% of genotypes were shared between boreal and arctic communities. Two independent Bayesian analyses and a neighbor-joining bootstrapping analysis of combined data from the nuclear large and small ribosomal subunits (LSUrDNA, SSUrDNA; 2.4 kb) showed that boreal and arctic endophytes represent Dothideomycetes, Sordariomycetes, Chaetothyriomycetidae, Leotiomycetes, and Pezizomycetes. Many well-supported phylotypes contained only endophytes despite exhaustive sampling of available sequences of Ascomycota. Together, these data demonstrate greater than expected diversity of endophytes at high-latitude sites and provide a framework for assessing the evolution of these poorly known but ubiquitous symbionts of living plants.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Ascomycota; Coevolution; Endophytic fungi; Geographic structure; Host specificity; ITS; LSUrDNA; SSUrDNA; Symbiosis

1. Introduction

Fungal endophytes are microfungi that live inside plants without causing apparent symptoms of infection. The vertically transmitted (seed-borne) endophytes of some grasses (Poaceae) can benefit their hosts via increased drought tolerance, protection against herbivory, and resistance to pathogens (reviewed by Clay and Schardl, 2002), but represent just one fungal family (Clavicipitaceae, Ascomycota) in symbiosis with a single plant lineage.

In contrast, all plants studied to date, including liverworts, mosses, ferns, conifers, and angiosperms, are infected by multiple lineages of horizontally transmitted fungal endophytes (Stone, 1987; Johnson and Whitney, 1992; Saikkonen et al., 1998; Davis et al., 2003). The majority of these endophytes are Ascomycota, although some Basidiomycota and a few Zygomycota are known (e.g., Petrini,

* Corresponding author. Present address: Division of Plant Pathology and Microbiology, Department of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA. Fax: +520 621 9290.
E-mail address: Arnold@ag.arizona.edu (A.E. Arnold).
[1] Present address: Department of Biology, University of Utah, Salt Lake City, UT 84112, USA.
[2] Present address: University of Virginia Health System, Charlottesville, VA 22908, USA.

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.07.012

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

1986; Zheng and Jiang, 1995). In general, horizontally transmitted endophytes associated with photosynthetic tissues such as leaves are not thought to benefit their hosts (see Saikkonen et al., 1998), but in most cases their ecological roles have not been assessed. Several recent studies have provided evidence for important, if previously overlooked, roles of endophytes in shaping plant defenses against herbivores, pathogens, and abiotic stressors (see Carroll and Carroll, 1978; Redman et al., 2002; Arnold et al., 2003; Arnold and Lewis, 2005). Together, these cryptic associates of foliage are thought to represent a significant, yet unknown proportion of the 1.5 million species of fungi thought to exist (Fröhlich and Hyde, 1999; Arnold et al., 2000; Hawksworth, 2001).

To date, foliar endophytes have been surveyed in few plant taxa and in relatively few sites, and their phylogenetic relationships have not been elucidated. Most research on the diversity of endophytes has focused on angiosperms and conifers in the temperate zone (reviewed in Saikkonen et al., 1998; Stone et al., 2000), although there is growing evidence that tropical endophyte diversity far exceeds that in the temperate zone (see Fröhlich and Hyde, 1999; Arnold and Lutzoni, 2006). Little is known regarding the endophytes associated with plants in high-latitude sites or in relatively extreme environments (e.g., arctic tundra).

Previous studies of plant–symbiotic fungi at higher latitudes have focused primarily on root-inhabiting fungi, including dark-septate endophytes (a diverse, polyphyletic group of ascomycetous endophytes characterized by melanized septa; Jumpponen, 2001) and arbuscular mycorrhizal fungi (AMF). Hambleton and Currah (1997) examined AMF and endophytic fungi from the roots of arctic Ericaceae and Olsson et al. (2004) determined that root-endophytic fungi are more prevalent in arctic sites than are AMF. These studies confirmed that arctic plants frequently engage in symbioses with fungi. However, the diversity and abundance of foliage-inhabiting endophytes have received little attention. Fisher et al. (1995) showed that leaves of *Dryas octopetala* (Rosaceae) harbored more endophytes than did twigs or roots, and that diversity in the arctic was lower than in subarctic sites. Johnson and Whitney (1992) examined endophyte colonization of healthy foliage of *Picea mariana* (Pinaceae) in boreal forest as a function of leaf age, revealing a higher abundance of endophyte isolates in older needles. However, to our knowledge, the molecular diversity, host preference, geographic structure, or phylogenetic relationships of high-latitude endophytes of foliage have not yet been assessed.

We used a culture-based approach to examine endophytic fungi in photosynthetic tissues of *Picea mariana* (Pinaceae), *Dryas integrifolia* (Rosaceae), and *Huperzia selago* (Lycopodiaceae) from three arctic and boreal sites. These hosts represent three major lineages of terrestrial plants (conifers, angiosperms, and lycophytes, respectively) that co-occur in high-latitude sites in North America. Fungal isolates from asymptomatic tissues were grouped into morphotypes, screened for unique genotypes using the fast-evolving nuclear internal transcribed spacer region (ITS), and their phylogenetic relationships inferred using Bayesian analyses of the nuclear ribosomal large and small subunits (LSUrDNA and SSUrDNA). This study is the first to examine fungal symbiont diversity from arctic and boreal plants in a global phylogeny of Ascomycota, and the first to investigate patterns of infection, molecular diversity, host preference, and geographic structure of endophytes associated with three co-occurring plant lineages. Studies of endophytic fungi in relatively extreme environments and from phylogenetically distinct lineages of plants are likely to recover novel species, which will be important in turn for understanding fungal diversity, the cryptic ecology of microfungal symbionts, and the evolution of plant–fungal symbioses.

2. Methods

2.1. Study sites and species

Tissue samples were collected from three host species in July 2003 at two boreal forest sites in Québec, Canada, and one arctic tundra site in Nunavut, Canada. *Picea mariana* (black spruce; Pinaceae) is a characteristic evergreen tree of boreal forests in Canada (Lauriault, 1989). *Dryas integrifolia* (mountain avens; Rosaceae), a wide-ranging pioneer shrub, forms low mats in rocky and gravelly sites at boreal and arctic latitudes (Porsild and Cody, 1980). *Huperzia selago* (fir clubmoss; Lycopodiaceae) is herbaceous with evergreen microphylls, and is distributed widely throughout northeastern Canada (Kartesz and Meacham, 1994). We sampled three individuals per focal host species, as available, in each site.

The study site at Schefferville, Québec (54.8°N, 66.8°W) consists of open, low-canopy *P. mariana* forest at the northern edge of its range (northern boreal forest/subarctic; Grondin and Melançon, 1980). The understory comprises dense mats of lichens (primarily *Cladonia rangiferina*) and mosses (*Pleurozium schreberi*), with small angiosperms including *Vaccinium angustifolium*, *V. vitis-idaea*, *Kalmia angustifolia*, *Empetrum nigrum*, and *Ledum groenlandicum*. Île Niapiskau, Mingan Archipelago, Québec (50.3°N, 63.6°W) is a coastal, southern boreal forest with mixed stands of *Abies balsamea* and *P. mariana*. *Picea glauca* is also present in the forest, and the understory is dominated by *C. rangiferina, Pl. schreberi*, *E. nigrum*, *K. angustifolia*, and *V. angustifolium*. The arctic site is located near Iqaluit, Nunavut (63.8°N, 68.5°W) in tundra dominated by *Cassiope tetragona*, *Dryas integrifolia*, and diverse lichens.

From each individual, photosynthetic tissues were collected from three branches (*Picea* at the Mingan Archipelago and Schefferville) or three stems (*Huperzia* at Schefferville; *Dryas* at Mingan Archipelago and Iqaluit). In sites where host species co-occurred, we sampled individuals of each species in close proximity to one another (1–2 m). Tissue samples were kept at ambient temperature

in sealed plastic bags with clean paper towels for up to 96 h during transit from collection sites to the lab for endophyte isolation.

2.2. Endophyte collection and isolation

Whole leaves (*Picea, Dryas*) or photosynthetic stems with microphylls (*Huperzia*) were washed in tap water and then surface-sterilized by rinsing briefly in 95% ethanol, followed by immersion in 10% Clorox (2 min) and 70% ethanol (2 min). This method effectively eradicates viable fungal spores and hyphae from leaf surfaces (Arnold et al., 2000). Using sterile technique, tissues were cut into ca. 2 mm^2 pieces and plated onto 2% malt extract agar (MEA), which supports growth by a wide diversity of fungal species (Fröhlich et al., 2000). Eighty tissue segments were plated per branch or stem sample ($N = 240$ segments/individual, 720 segments/species/site, 3600 segments overall). Plates were sealed and incubated at room temperature for up to one year, or until fungal growth was observed. Emergent hyphae were transferred and purified on new MEA plates, and whole-colony morphology was documented and photographed. Living vouchers have been deposited at the Robert L. Gilbertson Mycological Herbarium at the University of Arizona.

2.3. DNA extraction, amplification, and data assembly

Total genomic DNA was extracted directly from mycelia in culture following Arnold et al. (2006) and Arnold and Lutzoni (2006). Three regions of the nuclear ribosomal RNA tandem repeats were amplified: the internal transcribed spacer region (ITS; ca. 0.6 kb), large subunit (LSUrDNA; ca. 1.4 kb), and small subunit (SSUrDNA; ca. 1.0 kb). Primers ITS5 or ITS1F, and ITS4 or LR3, were used to amplify ITS (White et al., 1990; Vilgalys and Hester, 1990; Gardes and Bruns, 1993). Various combinations of the forward primers nssu131 and nssu97a (Kauff and Lutzoni, 2002) and reverse primers NS17, NS22 (Gargas and Taylor, 1992) and nssu1088 (Kauff and Lutzoni, 2002) were used to amplify SSUrDNA. LSUrDNA was amplified using primers LR0R and LR7 (Vilgalys and Hester, 1990). All primer sequences are listed at www.biology.duke.edu/fungi/mycolab and www.lutzonilab.net/pages/primer.shtml. PCR protocols followed Arnold and Lutzoni (2006): 25 μl reactions included 2.5 μl dNTPs (10 μM), 2.5 μl bovine serum albumin (BSA), 2.5 μl PCR buffer, 1.25 μl of each primer (10 μM), 13.875 μl water, 0.125 μl *Taq* polymerase, and 1.0 μl of diluted template. Cycling reactions were run on PTC-200 Thermal Cyclers (MJ Research) using the following parameters: 4 min at 95 °C; 35 cycles of 30 s at 95 °C, 30 s at 50°, 52° or 54 °C, and 90 s at 72 °C; 10 min at 72 °C. Gel electrophoresis using SYBR Green demonstrated single bands for all products. PCR products were cleaned using Qiagen purification columns and sequenced for both forward and reverse reads on an ABI 3700 using PCR primers, with some SSUrDNA sequences also requiring internal primers SR11R and SR7, and all LSUrDNA sequences requiring primers LR3 and LR3R or LR5 (Vilgalys and Hester, 1990; Spatafora et al., 1995).

The majority of endophytes recovered in this study remained sterile in culture. Therefore, isolates were grouped to morphotypes following Arnold et al. (2000), and up to eight isolates of each morphologically distinct colony type were sequenced for ITS. ITS data were obtained for 280 representative isolates (50.2% of isolates recovered). Consensus sequences were assembled into genotype groups based on 90% ITS sequence similarity using Sequencher v. 4.2.2 (Gene Codes Corporation, Ann Arbor, Michigan, USA). Forty-seven unique ITS genotype groups were identified, of which five had highest BLAST affinity for members of the Basidiomycota, and 42 had highest BLAST affinity for members of the Ascomycota. We chose up to five isolates from each ascomycetous genotype group for multilocus sequencing, for a total of 67 isolates. Forty-six of those isolates were successfully sequenced for both LSUrDNA and SSUrDNA, representing 30 unique ITS genotypes.

For each locus, base calls of complementary reads were assigned quality scores using phred (Ewing et al., 1998) and assembled into contigs using phrap (Ewing and Green, 1998). A BioPython script written by F. Kauff (Duke University; available on request from FL) facilitated the use of phred and phrap. Automatically assembled contigs were verified by visual inspection in Sequencher v. 4.2.2, and ambiguous base calls were determined manually. All sequences obtained in this study (280 ITS, 46 LSUrDNA, and 46 SSUrDNA sequences) have been submitted to GenBank under Accession Nos. DQ979418–DQ979789.

2.4. Ecological analyses

Because the majority of endophytes remained sterile in culture, we used ITS data to designate operational taxonomic units (OTU) for ecological analyses. Sequencher was used to delimit groups corresponding to 90%, 95%, 97%, and 99% ITS similarity, without considering differences in sequence length (after Arnold et al., 2006). Previous studies have used 90–97% ITS sequence similarity as a proxy for species boundaries in fungi (e.g., 97%, O'Brien et al., 2005; 95%, Arnold and Lutzoni, 2006; 90%, Hoffman and Arnold, in review). However, the degree of ITS divergence within and among species of fungi, and whether such percent divergences are consistent among different clades of fungi, remains to be determined. Therefore, we used ITS genotypes defined by 90%, 95%, 97%, and 99% ITS similarity as OTU to assess species richness, estimated richness, and diversity. To conservatively estimate geographic and host specificity, OTU were defined as ITS genotypes based on 90% sequence similarity.

Species accumulation curves, bootstrap estimates of total richness, and diversity were inferred using EstimateS 6.0 (R. K. Colwell, http://purl.oclc.org/estimates). Diversi-

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

ty was measured using the Shannon index, Simpson index, and Fisher's alpha (α), which is robust for comparisons among samples of different sizes (Leigh, 1999). Fisher's α is defined implicitly by the formula

$$S = \alpha^* ln(1 + n/\alpha)$$

where S is the number of taxa, n is the number of individuals (defined by numbers of isolates), and α is Fisher's α (Leigh, 1999). To examine similarity in communities of fungi within and among sites and hosts, we used similarity indices based on presence/absence data only (Jaccard's index) and isolation frequency (Morisita-Horn index) (Arnold et al., 2003). All indices were calculated using EstimateS.

2.5. Phylogenetic analyses

To reconstruct the phylogenetic relationships of endophytes in the context of other Ascomycota, endophyte sequences generated for this study were incorporated into a core alignment of 241 sequences of Ascomycota generated for the Assembling the Fungal Tree of Life project (AFTOL; Lutzoni et al., 2004). GenBank accession and identification numbers for these taxa are given in Lutzoni et al. (2004). For additional comparisons of endophyte taxon composition with regard to latitude and host species, we included LSUrDNA and SSUrDNA sequences for 24 endophytes isolated from tropical and temperate hosts (Arnold and Lutzoni, 2006; Arnold et al., 2006). These endophytes were isolated as described above. Temperate endophytes included cultures from healthy, >1-year-old foliage of *Pinus taeda* (Pinaceae) and *Magnolia grandiflora* (Magnoliaceae) in Duke Forest (Orange County, NC, USA; 35°58′N, 79°05′W) (Arnold et al., 2006). Tropical endophytes included cultures from six angiosperm taxa: *Swartzia simplex* (Caesalpiniaceae), *Theobroma cacao* (Malvaceae), *Faramea occidentalis* (Rubiaceae), *Laetia thamnia* (Flacourtiaceae), *Gustavia superba* (Lecythidaceae), and *Trichilia tuberculata* (Meliaceae), which were isolated from healthy, mature foliage in primary and late secondary forest at Barro Colorado Island, Panama (9°9′N, 79°51′W) (Arnold and Lutzoni, 2006). In each case, focal endophytes were chosen to represent the most common morphotypes and ITS genotypes characterizing these tropical and temperate sites.

Alignments of LSUrDNA and SSUrDNA data, each containing 311 sequences, were generated in MacClade v. 4.06 (Maddison and Maddison, 2003) according to the secondary structure model (Kjer, 1995) of *Saccharomyces cerevisiae* (Saccharomycetes, Accession No. U53879) published by Cannone et al. (2002). Ambiguously aligned regions (sensu Lutzoni et al., 2000) were excluded from analyses.

A reciprocal 70% bootstrap support value was used as a criterion to detect conflicts between data partitions (LSUrDNA vs. SSUrDNA) following Reeb et al. (2004). A hierarchical likelihood ratio test was implemented using Modeltest 3.6 (Posada and Crandall, 1998), yielding a GTR (General Time Reversible) + I (Invariant) + G (Gamma) model for each dataset. Using this model, we performed a neighbor-joining bootstrap analysis using PAUP* 4.0b8a (Swofford, 2004) on both the LSUrDNA and SSUrDNA data sets using maximum likelihood distances, replicated 1000 times. Conflicts between the topologies generated by each data set were considered significant if a given taxon was a part of two different relationships that both had bootstrap support $\geqslant$70% (Mason-Gamer and Kellogg, 1996). We found no significant conflict between the LSUrDNA and SSUrDNA data sets, and the data were combined for subsequent analyses.

Modeltest was then used to determine the model that best fit the combined data set. The following parameters were calculated: GTR + I + G model, with base frequencies of A, C, G, and T equal to 0.2350, 0.2424, 0.3303, and 0.1923, respectively; substitution rate matrix ([A–C] = 0.9294, [A–G] = 2.5896, [A–T] = 1.4933, [C–G] = 0.5343, [C–T] = 6.1776, and [G–T] = 1.000); gamma distribution parameter ($\alpha = 0.5714$); and proportion of invariable sites (I = 0.3494). This model and these parameters were used for subsequent neighbor-joining bootstrapping analysis of the combined data set.

Metropolis coupled/Markov chain/Monte Carlo analysis was conducted on the combined LSUrDNA and SSUrDNA data set using MrBayes v.3.0b4 (Huelsenbeck and Ronquist, 2001). For each of the two data partitions we applied the GTR + I + G model. Ten independent analyses, each with 5,000,000 generations, were initiated with random trees and sampled every 100th tree. The analysis was run on four separate, simultaneous chains. The resulting average likelihood scores differed substantially among the 10 analyses.

Ten random trees from the 20,000 last trees generated by the highest-likelihood run were used as starting trees to initiate 10 new runs with the same settings as the first round except for the quality of the starting tree. At the same time, we initiated 10 runs of 10,000,000 generations each, starting from a random tree, which were sampled every 500th tree. We compared the average likelihood scores of the last 10,000 trees from each run and found that the best likelihood generated by any of the longer runs was inferior to the likelihoods generated by the runs of 5,000,000 generations that used a non-random starting tree. Therefore, we used trees generated by the run of 5,000,000 generations that started from a tree generated by the previous run to begin all future analyses.

After checking the distribution of the likelihood scores of the last 10,000 trees generated by the 10 runs of 5,000,000 generations, we chose the last 10,000 trees from eight of the 10 runs that provided the highest likelihood scores (−42746.691 to −42622.63). The resulting pool of 80,000 trees was used to compute a majority-rule consensus tree in PAUP*. This consensus tree was rooted with *Taphrina communis*, and was used as the final tree.

Support for nodes was provided both by bootstrap values generated by the neighbor-joining bootstrap analysis with maximum likelihood distances, and by posterior probabilities generated by both MrBayes and BayesPhylogenies (Pagel and Meade, 2004a). Using BayesPhylogenies, we applied up to six different models of evolution to the combined data set, including gamma distributed rate heterogeneity. We first estimated the optimal number of partitions across the SSUrDNA and LSUrDNA datasets. We initiated nine runs with various settings (2 partitions with GTR, 2 partitions with GTR + G, 3 partitions with GTR, 3 partitions with GTR + G, etc., up to 6 partitions with GTR). After 1,000,000 generations, we computed likelihood scores by separately averaging the likelihood values of the last 2000 trees for each run. The likelihood score recovered using 5 partitions with GTR + G (−41414.998) was significantly better (significance = 27.6 units, Pagel and Meade, 2004b) than the next best run (6 partitions with GTR; average likelihood score of −41508.079). Using the optimal settings, we initiated 20 final runs of 1,000,000 generations, each starting with a random tree selected from the 2000 trees from the previous best run. Average likelihoods of the last 1000 trees (sampled every 100th generation) of each of the 20 runs ranged from −41147.196 to −41236.49. We selected the 10 runs with the highest likelihood scores and combined the last 1000 trees from each to compute a majority-rule consensus tree from 10,000 trees. The posterior probabilities generated by this analysis were used for comparison with other measures of support.

3. Results

3.1. Diversity of boreal and arctic endophytes

Endophytic fungi were isolated from all surveyed individuals of *Picea mariana*, *Dryas integrifolia,* and *Huperzia selago*. In total, 558 isolates of endophytic fungi were obtained from *Picea* (460 isolates), *Dryas* (55 isolates), and *Huperzia* (43 isolates). Infection frequencies ranged from 0.97% (*Dryas*, Iqaluit) to 41.30% (*Picea*, Schefferville) of tissue segments (Table 1). ITS data were obtained for 280 representative isolates, including 40–100% of endophytes isolated from each host species in each site (Table 1). Over the entire dataset, species accumulation curves remained non-asymptotic regardless of the degree of ITS sequence similarity used to delimit operational taxonomic units (Fig. 1). Bootstrap estimates of species richness significantly exceeded the observed species richness under all OTU definitions (Table 2). The majority of ITS genotypes occurred only once (Table 2). Diversity ranged from 16.29–25.93 (Fisher's α), 2.69–2.99 (Shannon index), and 6.73–7.77 (Simpson index) depending on the stringency of ITS genotype groupings (Table 2).

Diversity of endophytes (Fisher's α) ranged from 3.43 (*Dryas* at Mingan Archipelago) to 10.92 (*Picea*, Mingan Archipelago) (Fig. 2). Thirty-one unique ITS genotypes (90% similarity) were observed among 190 sequenced isolates from *Picea* (Fisher's $\alpha = 10.64$). Endophytes of *Dryas* represented 15 unique genotypes among 54 samples (Fisher's $\alpha = 6.87$). *Huperzia* endophytes displayed the lowest richness and diversity, with nine unique genotypes among 36 samples (Fisher's $\alpha = 3.86$).

3.2. Host and geographic specificity of endophytes

From 280 representative isolates sequenced for ITS, we recovered 47 distinct genotype groups based on 90% ITS sequence similarity. Of these, 26 genotypes (55.3%) were represented by only one sequence (hereafter, singletons), while 21 genotypes were represented by multiple isolates (non-singletons) (Table 2). Of the 21 non-singleton genotypes, 13 were isolated from only one of the three host species (Appendix 1). Representatives of eight genotypes were found in two host taxa. No genotype was recovered from all three plant species.

Among all ITS genotypes obtained from *Picea*, 77.4% were found only in that host, despite sampling of sympatric *Huperzia* at Schefferville and *Dryas* at the Mingan Archipelago. Six of nine genotypes (66.7%) found in *Huperzia* were recovered only from that host species, and nine of 15 genotypes (60%) from *Dryas* were unique to that host.

Among endophytes of *Picea,* genotype AE (highest BLAST affinity for Cudoniaceae) comprised 20.6% of isolates (Appendix 1 and 2). Only one other genotype comprised $\geqslant 5.0\%$ of isolates from that host (genotype AI; Rhytismatales; Appendix 1 and 2). Genotype AD (Mycosphaerellaceae) was the most commonly isolated endophyte from *Dryas* (34.5% of isolates) and *Huperzia* (37.2% of isolates), but the next most common endophytes were distinctive for each host (*Dryas,* AF, Botryosphaeria-

Table 1
Infection frequencies, total number of isolates, sequencing intensity for ITS, and representation of ITS genotypes among isolates sequenced for LSUrDNA and SSUrDNA for endophytes isolated from boreal forest at Mingan Archipelago, Québec (MA), boreal forest at Schefferville, Québec (SF), and arctic tundra at Iqaluit, Nunavut (IQ)

Host species	Site	Segments infected (%)	Isolates recovered	Isolates sequenced, ITS (%) (N)	ITS genotypes (90%)	ITS genotypes (and isolates) in phylogenetic analysis
Picea mariana	MA	22.60	163	43.5 (71)	23	12 (14)
	SF	41.30	297	40.1 (119)	15	8 (9)
Dryas integrifolia	MA	6.67	48	97.9 (47)	9	5 (13)
	IQ	0.97	7	100 (7)	6	4 (4)
Huperzia selago	SF	5.97	43	83.7 (36)	9	3 (6)

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

ARTICLE IN PRESS

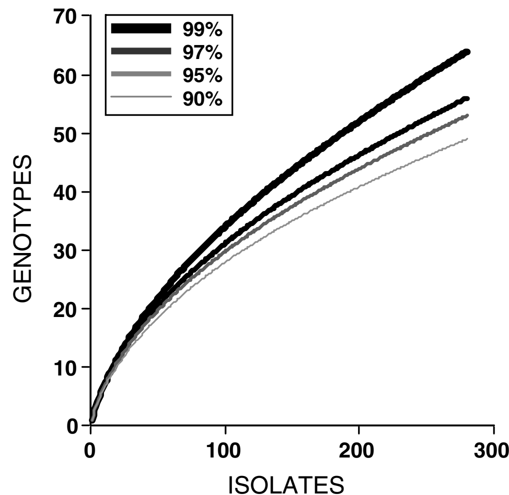

Fig. 1. Non-asymptotic species accumulation curves for fungal endophytes from healthy photosynthetic tissues of three host species (*Picea mariana, Dryas integrifolia,* and *Huperzia selago*) in boreal and arctic sites. Distinct curves indicate operational taxonomic units defined by 90–99% ITS sequence similarity.

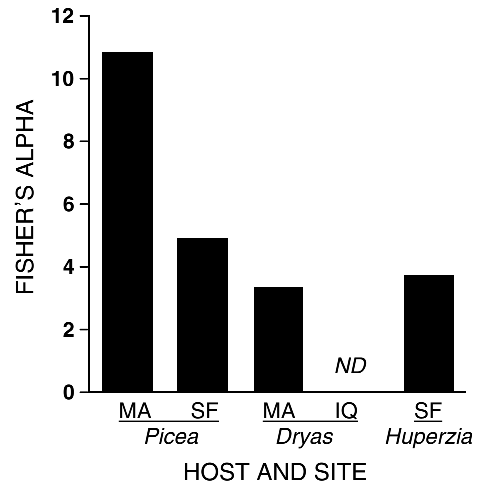

Fig. 2. Diversity (Fisher's α) as a function of host species and study site, based on conservatively delimited ITS genotype groups (90% sequence similarity).

ceae: 20% of isolates; *Huperzia*, AC, Pleosporales: 18.6% of isolates).

Communities of endophytes in boreal forest sites were more similar to each other than to the endophyte community at the arctic tundra site (Table 3). Based on non-singletons only, similarity in endophyte communities among sites ranged from 0.07 (Schefferville vs. Iqaluit) to 0.43 (Schefferville vs. Mingan Archipelago) based on presence–absence data (Jaccard's index), and from 0.02 (Schefferville vs. Iqaluit) to 0.51 (Schefferville vs. Mingan Archipelago) based on isolation frequencies (Morisita-Horn index; Table 3).

Endophyte communities also appear to be structured by host species. Based on both presence/absence and isolation frequency data, endophyte communities in *Picea* at the Mingan Archipelago are significantly more similar to communities from *Picea* at Schefferville than to any others (Table 4). Similarity was low for comparisons among species in different sites, with the exception of *Dryas* at the Mingan Archipelago and *Huperzia* at Schefferville: based on isolation frequency, endophyte communities of these hosts appear to be highly similar (Table 4). However, this pattern is driven by a single genotype (AD; Appendix 1 and 2) that dominated the endophyte community in both hosts. As demonstrated by the much lower similarity based on presence/absence data (Jaccard's index = 0.10), *Dryas* and *Huperzia* share relatively few endophyte genotypes overall (Table 4, Appendix 1).

Among 15 non-singleton genotypes recovered from *Picea,* eight genotypes (53.3%) were recovered from that host in only one site (Schefferville or the Mingan Archipelago; Appendix 1). All unique, non-singleton ITS genotypes from *Dryas* were recovered from that host in only one site (Iqaluit or the Mingan Archipelago). However, several genotypes were shared among different host species at geo-

Table 3
Similarity of endophyte communities as a function of study site (MA, boreal forest at Mingan Archipelago; SF, boreal forest at Schefferville; IQ, arctic tundra at Iqaluit)

	MA	SF	IQ
MA	—	**0.43**	0.17
SF	**0.51**	—	0.07
IQ	0.22	0.02	—

Data reflect similarity based on presence/absence data only (Jaccard's index, above diagonal) and isolation frequency (Morisita-Horn, below diagonal). Bold font indicates highest observed similarity for each index, showing the relatively high similarity between the two boreal forest sites.

Table 2
Richness, percent singletons, and diversity among 280 representative isolates of endophytic fungi from boreal and arctic sites, as a function of operational taxonomic units based on 90%, 95%, 97%, and 99% ITS sequence similarity

ITS similarity	S_{obs}	S_{obs} SD	Bootstrap	Singletons N (%)	Fisher's α	Shannon	Simpson
90%	47	5.90	59.68*	26 (55.32)	16.29	2.69	6.73
95%	53	5.99	64.89*	30 (56.60)	19.35	2.81	7.24
97%	56	6.01	68.73*	32 (57.14)	21.04	2.87	7.36
99%	64	6.75	79.65*	40 (62.50)	25.93	2.99	7.77

S_{obs} indicates observed richness. S_{obs} SD indicates standard deviation inferred from 50 randomizations of sample order. Bootstrap values indicate the inferred estimate of total species richness; asterisks indicate bootstrap estimates that significantly exceed $S_{obs} + 1$ SD ($\alpha = 0.05$). Singletons reflect the number and percent of genotypes represented only once. Diversity data are given as Fisher's α, Shannon index, and Simpson index values.

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

Table 4
Similarity of endophyte communities as a function of host species and study site, reflecting similarity based on presence/absence data only (Jaccard's index, above diagonal) and isolation frequency (Morisita-Horn, below diagonal)

	Dryas—MA	*Picea*—MA	*Huperzia*—SF	*Picea*—SF	*Dryas*—IQ
Dryas—MA	—	0.11	0.10	0.14	0.00
Picea—MA	0.03	—	0.06	**0.47**	0.23
Huperzia—SF	**0.68**	0.01	—	0.08	0.00
Picea—SF	0.04	**0.58**	0.00	—	0.09
Dryas—IQ	0	0.34	0.00	0.02	—

Bold font indicates values that significantly exceed mean similarity for all comparisons (Wilcoxon sign-rank test; $\alpha = 0.05$), revealing high similarity of endophyte communities of (1) *Picea* at SF and MA (Jaccard and Morisita-Horn), and (2) *Huperzia* at SF and *Dryas* at MA (Morisita-Horn only).

graphically distant sites. For example, *Dryas* at the Mingan Archipelago shared one genotype with *Huperzia* from Schefferville (AD, with highest BLAST affinity for Mycosphaerellaceae; Appendix 1 and 2). *Picea* at Schefferville shared two genotypes with *Dryas* at the Mingan Archipelago (AE and AK, both unidentified species based on BLAST results). Overall, three of the genotypes recovered from *Dryas* at Iqaluit were also found in boreal sites (6% of all genotypes recovered). However, two of these genotypes contained isolates that differed in their top BLAST matches, with the *Dryas* endophytes from Iqaluit distinct from the samples collected at boreal sites (Appendix 1 and 2).

3.3. Phylogenetic analyses

Each alignment contained 311 sequences, including 46 sequences of arctic and boreal endophytes (see Appendix 1 for culture, origin, and BLAST information), and 24 sequences of temperate and tropical endophytes (Arnold and Lutzoni, 2006; Arnold et al., 2006). The LSUrDNA alignment consisted of 4378 characters: 991 included characters, and 3387 characters excluded as introns or because of ambiguous alignment. The SSUrDNA alignment included 7818 characters, of which 956 characters were included.

Analysis of the combined data sets resulted in a highly resolved phylogeny of Ascomycota that largely agrees with previously published trees (Fig. 3; see Lutzoni et al., 2004; Reeb et al., 2004; Lumbsch et al., 2005). The Dothideomycetes + Arthoniomycetes was recovered as a monophyletic lineage composed of five subclades, delineated and named where relevant according to Lutzoni et al. (2004): Dothideomycetes 3 (clade A on Fig. 3), Dothideomycetes 2 (clade B), a clade of endophytes and *Botryosphaeria ribis* (clade C), the Arthoniomycetes (clade D), and Dothideomycetes 1 (clade E). For the first time, the monophyletic Arthoniomycetes was nested within the Dothideomycetes, with significant support from both Bayesian analyses (see Section 4.). The overall monophyly of the Dothideomycetes + Arthoniomycetes was highly supported by both estimates of posterior probabilities.

The Sordariomycetes (clade F) was recovered as monophyletic with high bootstrap and Bayesian support values. The monophyly of the Sordariomycetes + Leotiomycetes 1 (clade G) received high support from Bayesian analyses. The Chaetothyriomycetidae (clade H) was recovered as sister to the Eurotiomycetidae (clade I) to form the Eurotiomycetes, which is monophyletic but without strong support. Relationships among the Eurotiomycetes, Acarosporomycetidae (clade J), Lecanoromycetidae (contained within clade K) and Ostropomycetidae (clade L) were unresolved, but together with *Strangospora* and members of the Umbilicariaceae they form a monophyletic group with strong support from both Bayesian analyses. The Pezizomycetes were recovered as a monophyletic group (clade M), but without significant support. The Saccharomycotina (represented by *Candida albicans* and *Saccharomyces cerevisiae*) are well supported by all three measures.

Posterior probabilities generated by MrBayes (MB) and BayesPhylogenies (BP) were generally consistent with one another (Fig. 3). Both methods demonstrated significant support ($\geqslant$95%) for 145 internodes, compared to 89 internodes supported by bootstrap values $\geqslant$70%. There were 19 internodes for which MrBayes demonstrated significant support while BP did not, and six internodes for which BP indicated significant support while MB did not. Seventy-seven internodes supported by bootstrap values $\geqslant$70% were significantly supported by both MB and BP (86.5%). Two nodes were supported by bootstrap ($\geqslant$70%) and MB ($\geqslant$95%) only; one node was supported by bootstrap ($\geqslant$70%) and BP ($\geqslant$95%) only; and nine nodes were supported only by bootstrap values ($\geqslant$70%).

3.4. Phylogenetic affinities of endophytes

Endophytes from arctic, boreal, temperate and tropical sites were associated with several major fungal lineages (Fig. 3): Dothideomycetes 1, Dothideomycetes 2, Dothideomycetes 3, Sordariomycetes, Leotiomycetes 1, Chaetothyriomycetidae, and Pezizomycetes. One boreal endophyte (5607) showed high affinity to *Phaeotrichum benjaminii*, but the placement of these two taxa was unresolved.

Three major lineages (Dothideomycetes 2, Sordariomycetes, and Chaetothyriomycetidae) contain endophytes from both northern (boreal and arctic) and southern (temperate and tropical) sites. However, representative endophytes from northern and southern sites frequently demonstrated different phylogenetic affinities at the class

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

level. Twenty-five of 32 representative endophytes within the Dothideomycetes were isolated from northern sites. The Dothideomycetes 1 and 3 (clades A and E) contain only northern endophytes. Within the Dothideomycetes, eight endophytes representing both southern and northern fungi form a clade sister to *Botryosphaeria ribis* (clade C). In contrast, the Sordariomycetes contain mostly temperate and tropical endophytes: 16 of 19 representative endophytes within this group were collected from southern sites. The Leotiomycetes 1 contained only northern endophytes, including 11 from *Picea* (Schefferville and Mingan Archipelago) and three from *Dryas* (Mingan Archipelago). The Pezizomycetes contained two northern endophytes. One northern and one southern endophyte displayed affinity to the Chaetothyriomycetidae. Endophyte 4466, representing an ITS genotype group of three isolates from two different *Picea* individuals (Appendix 2 and 3), is nested within a paraphyletic group of lichen-forming fungi including Verrucariales (*Verrucaria* and *Dermatocarpon*) and Pyrenulales (*Pyrenula*). No endophytes were recovered among the primarily or exclusively lichen-forming lineages Acarosporomycetidae, Lecanoromycetidae, Ostropomycetidae, and Lichinomycetes, nor among the Arthoniomycetes (here represented by lichen-forming taxa). Endophytes were also absent among the non-lichenized Leotiomycetes 2 and Eurotiomycetidae.

Genotype groups defined by 90% ITS sequence similarity consistently clustered together in our LSUrDNA + SSUrDNA tree (see Appendix 2 and 3). However, phylogenetic structure was occasionally recovered within groups of isolates that represented the same ITS type. For example, within the Dothideomycetes 2 (clade B), endophytes 2712 and 2722 (from *Huperzia*, Schefferville) and 3358A, 4140, and 4221 (from *Dryas*, Mingan Archipelago) represent the same ITS genotype group, but were reconstructed with significant support as two distinct clades: one containing *Huperzia* endophytes, and the other containing *Dryas* endophytes.

Although broad-scale patterns of coevolution with hosts are not evident from this analysis, our results provide evidence for host-structuring of endophyte lineages. Phylotypes (defined by significant support values) of boreal and arctic endophytes typically represented only a single host taxon. For example, multiple ITS genotypes of *Picea* endophytes were reconstructed as close relatives of one another within the Leotiomycetes 1 (Fig. 3).

4. Discussion

Fungal diversity is typically thought to increase with decreasing latitude, concomitant with many other terrestrial organisms (see Hawksworth, 1991, 2001). However, this study reveals a high diversity of endophytic fungi in aboveground tissues of arctic and boreal plants. Although diversity was lower among endophytes of *Huperzia* and *Dryas,* the diversity of endophytes recovered from healthy foliage of *Picea* in a single site (Mingan Archipelago; Fisher's $\alpha = 10.64$) is consistent with that observed with several hosts in temperate and tropical sites (Arnold and Lutzoni, 2006). Non-asymptotic species accumulation curves (Fig. 1), the disparity between observed and estimated richness (Table 2), and the occurrence of most genotypes only once (>55%; Table 2) suggest that many endophytic fungi have yet to be recovered in these putatively low-diversity environments.

The relatively high similarity between boreal forest sites at Schefferville and the Mingan Archipelago contrasts markedly with the very low similarity between the boreal and arctic sites (Table 3). These patterns likely reflect the similarity of plant communities between the Mingan and Schefferville sites, but also may be driven by abiotic factors that could shape the fungal community in each site. Identifying the mechanisms of apparent geographic structure represents an important direction for future work.

The high diversity of endophytes at the species level parallels the phylogenetic diversity of these fungi, highlighting the occurrence of foliage-symbiotic fungi in many of the major lineages of Ascomycota. This study did not provide consistent evidence for clade-level specificity to hosts or sites, but instead demonstrated that clades of endophytes may contain representatives from many hosts or geographic regions. However, host affinity was manifested by the most commonly isolated endophytes. Genotypes AE, AQ, and AP were found almost exclusively in leaves of *Picea*; AI was unique to *Picea*; AF, AG, and AL were unique to *Dryas*; AC was unique to *Huperzia*; and AD was found only in *Dryas* and *Huperzia* (Appendix 1). These observations do not exclude the importance of abiotic or geographic factors in shaping the endophytic community for a given host. Both host and bioclimatic zones seem to be determinant factors in explaining the observed variation in endophytic communities.

→

Fig. 3. Phylogenetic affinities of 46 endophytes from *Huperzia selago, Dryas integrifolia*, and *Picea mariana* sampled from arctic and boreal sites (this study; white circles), and 24 endophytes from conifers and angiosperms in temperate and tropical sites (Arnold et al., 2006; Arnold and Lutzoni, 2006; Arnold, 2006; black circles). The phylogenetic tree includes 311 terminal branches and is based on LSUrDNA + SSUrDNA data. Support values are posterior probabilities generated by MrBayes (before slash) or BayesPhylogenies (after slash), and bootstrap proportions (beneath line) calculated using a neighbor-joining bootstrapping analysis. Sites are abbreviated as follows: QU-MA, Québec-Mingan Archipelago (southern boreal forest); QU-SF, Québec-Schefferville (northern boreal forest); IQ, Iqaluit, Nunavut (arctic tundra); NC, North Carolina (temperate semi-deciduous forest); PAN, Panama (lowland, moist tropical forest). Lichen-forming taxa are shaded in grey. Major fungal lineages are identified by bars along the right margin. Focal clades referred to in the text are marked with letters; clade names, where relevant, follow Lutzoni et al. (2004). The asterisk following *Leotia lubrica* denotes an unexpected placement in the Sordariomycetes, and the asterisks following *Sarea resinosa* denote an unexpected affiliation for the Leotiomycetes 1. The identity of the sources of these sequences should be examined. The ‡ following *Balansia* denotes the endophytic symbioses formed by members of this genus with grass hosts.

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

Phylotypes consisting only of endophytic fungi were common in this study, and frequently contained both boreal/arctic and temperate/tropical endophytes (e.g., the well-supported monophyletic group in clade C containing endophytes 469 and 435b from *Laetia* in Panama, 1182 from *Magnolia* in North Carolina, 6730 and 6731 from *Swartzia*

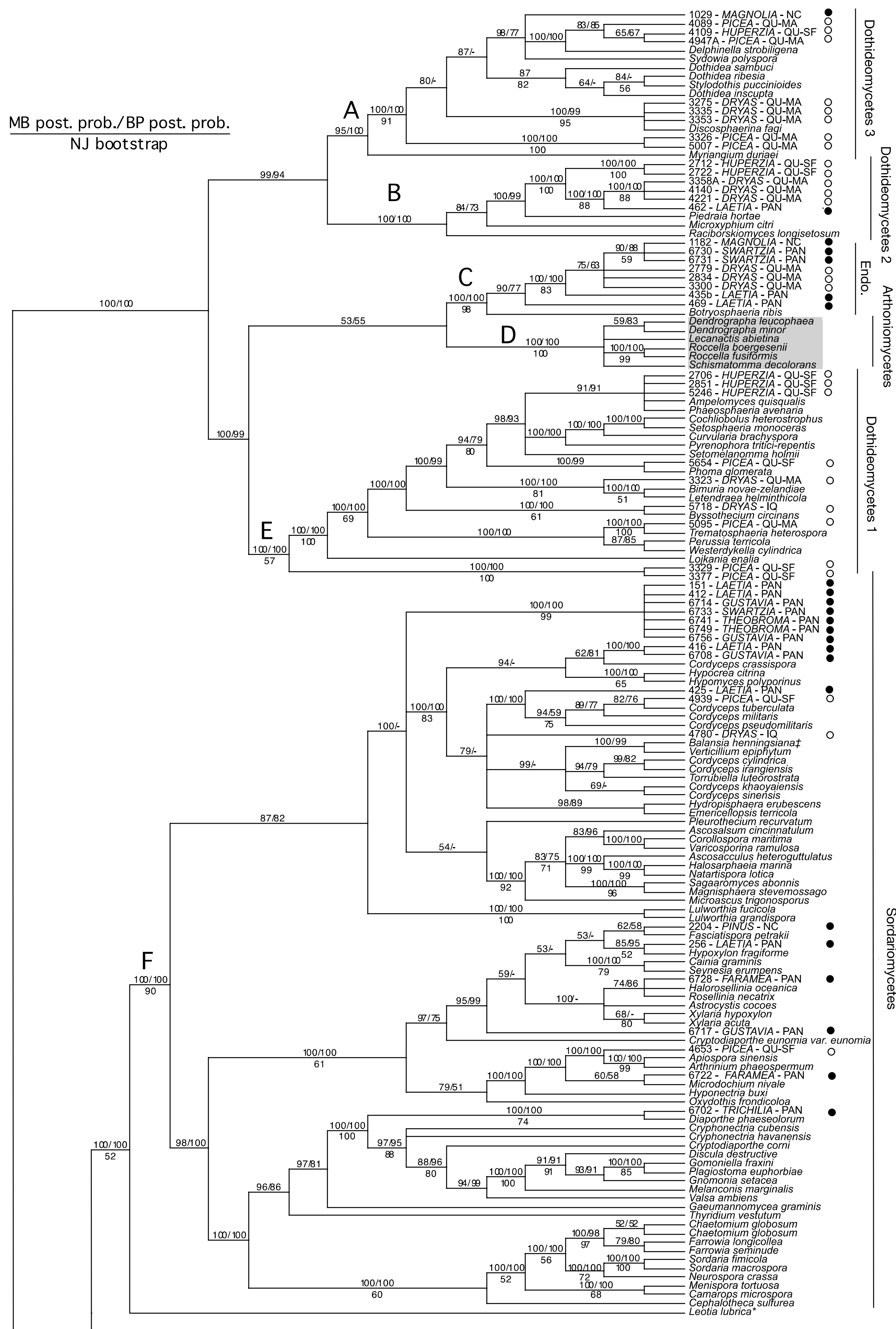

Fig.3 (*Continued*)

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

in Panama, and three isolates from *Dryas*) (Fig. 3). These phylotypes may represent endophytic lineages of fungi that were previously unknown, or lineages of recognized taxa that are not yet represented by LSUrDNA and SSUrDNA sequences in GenBank. The presence of endophyte-exclusive and endophyte-dominated phylotypes argues strongly for the importance of these and other little-known symbionts in defining the shape and structure of the Ascomycota tree of life. Moreover, the observation that ITS genotype groups can contain multiple, well-supported phylotypes argues that true species diversity may exceed our conservative estimate.

4.1. Identities of boreal and arctic endophytes

Johnson and Whitney (1992) examined spore morphology of endophytes from *P. mariana* in New Brunswick, Canada. Among 914 isolates from 4800 needle segments, *Cryptocline abietina* (241 isolates), *Aureobasidium pullulans* (186 isolates), and *Phaeococcus catenatus* (137 isolates) were most common. Similarly, Fisher et al. (1995) most frequently recovered *Clypeopycnis* sp. and *Wettsteinina dryadis* among endophytes of *Dryas octopetala* in Norway. *Cryptocline abietina, P. catenatus, Clypeopycnis* sp., and *Wettsteinina* are not presently represented by ITS data in GenBank, and thus could not be recovered by our BLAST searches. However, our BLAST analyses did not yield close matches to taxa closely related to these species. One endophyte in the present study showed highest BLAST affinity for *A. pullulans* based on ITS data (endophyte 3357, from *Dryas*; Appendix 2 and 3). However, LSUrDNA and SSUrDNA sequences for *A. pullulans* are not currently available, and thus were not included in our phylogenetic analyses. These observations are representative of an ongoing challenge in fungal biodiversity exploration: relating traditionally identified fungi to genotypes of sterile isolates.

Because our study did not recover all endophyte taxa present in our focal host species (Fig. 1), it is possible that these fungal species remain to be recovered in our study sites. Alternatively, spatial or geographic variation in endophyte assemblages (Table 3) suggests that these host species may affiliate with different symbionts over their geographic ranges (see Fisher et al., 1995). The present study, which suggests an interaction between geographic structure and host affinity (Table 4), provides evidence for this view. This hypothesis is gaining further support from other comparative studies (see Fisher et al., 1995; Fröhlich and Hyde, 1999; Arnold and Lutzoni, 2006) and suggests remarkable complexity in the evolution and ecology of endophyte-plant associations.

4.2. Utility of BLAST identification compared to phylogenetic analysis

Phylogenetic affinities of isolates based on SSUrDNA and LSUrDNA data are useful for assessing the accuracy and value of BLAST identifications of ITS sequences, which may be limited by misidentified, unidentified, and unrepresented fungi in GenBank. In this study, phylogenetic placement of endophytes generally agreed with higher taxonomic levels (order, family) based on BLAST identification of ITS sequences. One exception was observed: endophyte 4466 displayed BLAST affinity to the Sordariomycetes, but phylogenetic analysis placed it within the Chaetothyriomycetidae (Eurotiomycetes). Phylogenetic inference determined specific placement for isolates 3377, 5718 (Dothideomycetes 1), and 5627 (Leotiomycetes 1), for which the best BLAST matches were unidentified fungi (e.g., ''ectomycorrhizal root tip,'' mitosporic Ascomycota, and ''uncultured mycorrhizal fungus''). When BLAST results were compared with results of a FASTA search (Pearson, 2000), only one endophyte was placed more precisely: endophyte 5627 showed high affinity to *Lachnum virgineum* under FASTA (http://iab-devel.arsc.edu/metagenomics) which was confirmed by phylogenetic analysis.

As more sequences become available in GenBank and misidentified sequences are corrected, the accuracy of BLAST and FASTA identifications will improve and these techniques will become more useful for identifying unknown fungi. At present, phylogenetic analysis of DNA sequences remains a more effective method for determining taxonomic placement of sterile endophytes. Phylogenetic analyses also provide complementary information about the evolution of endophytic fungi, which cannot be inferred from BLAST or FASTA searches alone. The phylogeny recovered here (Fig. 3) is suggestive of multiple, unique origins of the endophyte symbiosis, providing a testable hypothesis for future studies with more intensive and geographically comprehensive sampling.

Endophytes recovered here displayed high affinity for all lineages of non-lichen forming, filamentous Ascomycota, with the exception of the Eurotiomycetidae and the Leotiomycetes 2. Of particular interest is the presence of two endophytes within the Chaetothyriomycetidae (Eurotiomycetes), which contains both lichen-forming and free-living fungi. Lutzoni et al. (2001) showed that the Chaetothyriales and Eurotiomycetidae (Eurotiomycetes) are descended from lichen-forming ancestors, which suggests an ancestral ability to form close symbiotic associations with photosynthetic organisms. Yet, endophytes are relatively rare in these lineages, indicating that such transitions may be more difficult than would be expected given that ancestral lichen symbiosis.

Sordariomycetes are frequently recovered as endophytes of tropical plants, and represent a significant proportion of tropical and temperate endophyte diversity (e.g., Arnold, 2006; Arnold and Lutzoni, 2006). However, few arctic and boreal endophytes displayed affinity to the Sordariomycetes, with highest frequency observed instead among the Dothideomycetes and Leotiomycetes. To our knowledge, the effects of latitude on the composition and phylogenetic diversity of endophyte communities has not been characterized, and presents an important question for

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.

understanding the evolutionary interactions between endophytic fungi and the plant lineages in which they are found. Integration of endophytes into a phylogenetic context provides a much-needed tool to address broad ecological and evolutionary questions in the study of endophytic associations. As the present study demonstrates, fungal endophytes from high-latitude and extreme environments are likely to be key components in disentangling the origins and roles of the endophyte symbiosis.

Acknowledgments

Funding support was provided by an NSF Postdoctoral Fellowship in Microbial Biology to AEA (DEB-0200413), NSF DEB-0228668 (Assembling the Tree of Life [AToL] grant) to FL, NSF CAREER Grant DEB-0133891 to FL, and NSF DEB-0343953 to AEA and J. Dalling. KLH was supported by the NSF Research Experience for Undergraduates program in Bioinformatics and Phylogenetics at Duke University with mentorship from AEA, JM, and FL, and gratefully acknowledges the support of Duke's Howard Hughes Biology Forum and the Undergraduate Research Support Program. We thank Nancy Dénommée from Parks Canada for logistical support at the Mingan Archipelago; Oksana Choulik and the McGill Subarctic Research Station for logistical support at Schefferville; the Lutzoni lab group at Duke University and the Arnold lab group at the University of Arizona for comments on the manuscript; and Frank Kauff at Duke University for the use of his BioPython scripts, which facilitated our analyses.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2006.07.012.

References

Arnold, A.E., 2006. Endophytic fungi: hidden components of tropical forest community ecology. In: Schnitzer, S., Carson, W. (Eds.), Tropical Forest Community Ecology. Blackwell Scientific, Inc. (in press).

Arnold, A.E., Henk, D.A., Eells, R.A., Lutzoni, F., Vilgalys, R., 2006. Diversity and phylogenetic affinities of foliar fungal endophytes in loblolly pine inferred by culturing and environmental PCR. Mycologia (in press).

Arnold, A.E., Lewis, L.C., 2005. Evolution of fungal endophytes, and their roles against insects. In: Vega, F., Blackwell, M. (Eds.), Ecological and Evolutionary Advances in Insect–Fungus Associations. Oxford University Press, Oxford, pp. 74–96.

Arnold, A.E., Lutzoni, F.L., 2006. Diversity and host range of foliar fungal endophytes: Are tropical leaves biodiversity hotspots? Ecology (in press).

Arnold, A.E., Maynard, Z., Gilbert, G.S., Coley, P.D., Kursar, T.A., 2000. Are tropical fungal endophytes hyperdiverse? Ecol. Lett. 3, 267–274.

Arnold, A.E., Mejía, L.C., Kyllo, D., Rojas, E.I., Maynard, Z., Robbins, N., Herre, E.A., 2003. Fungal endophytes limit pathogen damage in a tropical tree. Proc. Nat. Acad. Sci. USA 100, 15649–15654.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M., Pande, N., Shang, Z., Yu, N., Gutell, R.R., 2002. The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BioMed. Central Bioinf. 3, 2.

Carroll, G.C., Carroll, F.E., 1978. Studies on the incidence of coniferous needle endophytes in the Pacific Northwest. Can. J. Bot. 56, 3034–3043.

Clay, K., Schardl, C., 2002. Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. Am. Nat. 160, S99–S127.

Davis, E.C., Franklin, J.B., Shaw, A.J., Vilgalys, R., 2003. Endophytic *Xylaria* (Xylariaceae) among liverworts and angiosperms: Phylogenetics, distribution, and symbiosis. Am. J. Bot. 90, 1661–1667.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. Genome Res. 8, 186–194.

Ewing, B., Hillier, L., Wendl, M., Green, P., 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. Genome Res. 8, 175–185.

Fisher, P.J., Graf, F., Petrini, L.E., Sutton, B.C., Wookey, P.A., 1995. Fungal endophytes of *Dryas octopetala* from a high arctic polar semidesert and from the Swiss Alps. Mycologia 87, 319–323.

Fröhlich, J., Hyde, K.D., 1999. Biodiversity of palm fungi in the tropics: are global fungal diversity estimates realistic? Biodiversity Conserv. 8, 977–1004.

Fröhlich, J., Hyde, K.D., Petrini, O., 2000. Endophytic fungi associated with palms. Mycol. Res. 104, 1202–1212.

Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity of basidiomycetes: application to the identification of mycorrhizae and rusts. Mol. Ecol. 2, 113–118.

Gargas, A., Taylor, J.W., 1992. Polymerase chain reaction (PCR) primers for amplifying and sequencing nuclear 18S rDNA from lichenized fungi. Mycologia 84, 589–592.

Grondin, P., Melançon, M., 1980. Étude phyto-écologique de la Grosse île au Marteau et de l'île à Samuel archipel de Mingan, Québec. Laboratoire d'écologie forestière, Universitè Laval, Québec, Canada.

Hambleton, S., Currah, R.S., 1997. Fungal endophytes from the roots of alpine and boreal Ericaceae. Can. J. Bot. 75, 1570–1581.

Hawksworth, D.L., 1991. The fungal dimension of biodiversity: magnitude, significance, and conservation. Mycol. Res. 95, 641–655.

Hawksworth, D.L., 2001. The magnitude of fungal diversity: the 1.5 million species estimate revisited. Mycol. Res. 105, 1422–1432.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Johnson, J.A., Whitney, N.J., 1992. Isolation of fungal endophytes from black spruce (*Picea mariana*) dormant buds and needles from New Brunswick, Canada. Can. J. Bot. 70, 1754–1757.

Jumpponen, A., 2001. Dark septate endophytes—are they mycorrhizal? Mycorrhiza 11, 207–211.

Kartesz, J.T., Meacham, C.A., 1994. Synthesis of the North American Flora. University of North Carolina at Chapel Hill, North Carolina Botanical Garden.

Kauff, F., Lutzoni, F., 2002. Phylogeny of Gyalectales and Ostropales (Ascomycota, Fungi): among and within order relationships based on nuclear ribosomal RNA small and large subunits. Mol. Phylogenet. Evol. 25, 138–156.

Kjer, K.M., 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. J. Mol. Evol. 4, 315–330.

Lauriault, J., 1989. Identification Guide to the Trees of Canada. National Museum of Natural Sciences, Ottawa.

Leigh Jr., E.G., 1999. Tropical Forest Ecology. Oxford University Press, Oxford.

Lumbsch, H.T., Schmitt, I., Lindemuth, R., Miller, A., Mangold, A., Fernandez, F., Huhndorf, S., 2005. Performance of four ribosomal DNA regions to infer higher-level phylogenetic relationships of inoperculate euascomycetes (Leotiomycota). Mol. Phylogenet. Evol. 34, 512–524.

ARTICLE IN PRESS

Lutzoni, F., Kauff, F., Cox, C.J., McLaughlin, D., Celio, G., Dentinger, B., Padamsee, M., Hibbett, D., James, T.Y., Baloch, E., Grube, M., Reeb, V., Hofstetter, V., Schoch, C., Arnold, A.E., Miadlikowska, J.M., Spatafora, J., Johnson, D., Hambleton, S., Crockett, M., Shoemaker, R., Sung, G.H., Lücking, R., Lumbsch, T., O'Donnell, K., Binder, M., Diederich, P., Ertz, D., Gueidan, C., Hansen, K., Harris, R.K., Hosaka, K., Lim, Y.W., Matheny, B., Nishida, H., Pfister, D., Rogers, J., Rossman, A., Schmitt, I., Sipman, H., Stone, J., Sugiyama, J., Yahr, R., Vilgalys, R., 2004. Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits. Am. J. Bot. 91, 1446–1480.

Lutzoni, F., Pagel, M., Reeb, V., 2001. Major fungal lineages are derived from lichen symbiotic ancestors. Nature 411, 937–940.

Lutzoni, F., Wagner, P., Reeb, V., Zoller, S., 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst. Biol. 49, 628–651.

Maddison, D.R., Maddison, W.P., 2003. MacClade 4.06: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA.

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

O'Brien, H.E., Parrent, J.L., Jackson, J.A., Moncalvo, J.-M., Vilgalys, R., 2005. Fungal community analysis by large-scale sequencing of environmental samples. Appl. Environ. Microbio. 71, 5544–5550.

Olsson, P.A., Eriksen, B., Dahlberg, A., 2004. Colonization by arbuscular mychorrhizal and fine endophytic fungi in herbaceous vegetation in the Canadian High Arctic. Can. J. Bot. 82, 1547–1556.

Pagel, M., Meade, A., 2004a. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. Syst. Biol. 55, 571–581.

Pagel, M., Meade, A., 2004b. Mixture models in phylogenetic inference. In: Gascuel, O. (Ed.), Mathematics of Evolution and Phylogeny. Clarendon Press, Oxford.

Pearson, W.R., 2000. Flexible sequence similarity searching with the FASTA3 program package. Methods Mol. Biol. 132, 185–219.

Petrini, O., 1986. Taxonomy of endophytic fungi of aerial plant tissues. In: Fokkema, N.J., van den Huevel, J. (Eds.), Microbiology of the Phyllosphere. Cambridge University Press, Cambridge, England, pp. 175–187.

Porsild, A.E., Cody, W.J., 1980. Vascular Plants of Continental Northwest Territories, Canada. National Museums of Canada, Ottawa.

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

Redman, R.S., Roossinck, M.J., Maher, S., Andrews, Q.C., Schneider, W.L., Rodriguez, R.J., 2002. Field performance of cucurbit and tomato plants colonized with a nonpathogenic, mutualistic mutant (path-1) of *Colletotrichum magna* (Teleomorph: *Glomerella magna*; Jenkins and Winstead). Symbiosis 32, 55–70.

Reeb, V., Lutzoni, F., Roux, C., 2004. Contribution of RPB2 to multilocus phylogenetic studies of the euascomycetes (Pezizomycotina, Fungi) with special emphasis on the lichen-forming Acarosporaceae and evolution of polyspory. Mol. Phylogenet. Evol. 32, 1036–1060.

Saikkonen, K., Faeth, S.H., Helander, M., Sullivan, T.J., 1998. Fungal endophytes: a continuum of interactions with host plants. Annu. Rev. Ecol. Syst. 29, 319–343.

Spatafora, J.W., Mitchell, T.G., Vilgalys, R., 1995. Analysis of genes coding for small-subunit rRNA sequences in studying phylogenetics of dematiaceous fungal pathogens. J. Clin. Microbiol. 33, 1322–1326.

Stone, J.K., 1987. Initiation and development of latent infections by *Rhabdocline parkeri* on Douglas-fir. Can. J. Bot. 65, 2614–2621.

Stone, J.K., Bacon, C.W., White Jr., J.F., 2000. An overview of endophytic microbes: Endophytism defined. In: Bacon, C.W., White, J.F. (Eds.), Microbial Endophytes. Marcel Dekker, Inc., New York, pp. 3–29.

Swofford, D.L., 2004. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4.0b8a Sinauer Associates, Sunderland, MA.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR protocols: A Guide to Methods and Applications. Academic Press, San Diego, pp. 315–322.

Vilgalys, R., Hester, M., 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J. Bacteriol. 172, 4238–4246.

Zheng, R.Y., Jiang, H., 1995. *Rhizomucor endophyticus* sp. nov., an endophytic zygomycetes from higher plants. Mycotaxon 56, 455–466.

Please cite this article as: K. Lindsay Higgins et al., Phylogenetic relationships, host affinity, and geographic structure ..., Molecular Phylogenetics and Evolution (2006), doi:10.1016/j.ympev.2006.07.012.