Endophytes of grasses

Culturing and direct PCR suggest prevalent host generalism among diverse fungal endophytes of tropical forest grasses

K. Lindsay Higgins
Phyllis D. Coley
Thomas A. Kursar

Division of Plant Pathology and Microbiology, School of Plant Sciences, University of Arizona, Tucson, Arizona 85721

Abstract: Most studies examining endophytic fungi associated with grasses (Poaceae) have focused on agronomically important species in managed ecosystems or on wild grasses in subtropical, temperate and boreal grasslands. However grasses first arose in tropical forests, where they remain a significant and diverse component of understory and forest-edge communities. To provide a broader context for understanding grass-endophyte associations we characterized fungal endophyte communities inhabiting foliage of 11 species of phylogenetically diverse C₃ grasses in the understory of a lowland tropical forest at Barro Colorado Island, Panama. Our sample included members of early-arising subfamilies of Poaceae that are endemic to forests, as well as more recently arising subfamilies that transitioned to open environments. Isolation on culture media and direct PCR and cloning revealed that these grasses harbor species-rich and phylogenetically diverse communities that lack the endophytic Clavicipitaceae known from diverse woodland and pasture grasses in the temperate zone. Both the incidence and
diversity of endophytes was consistent among grass species regardless of subfamily, clade affiliation or ancestral habitat use. Genotype and phylogenetic analyses suggest that these endophytic fungi are predominantly host generalists, shared not only among distinctive lineages of Poaceae but also with non-grass plants at the same site.

**Key words:** Ascomycota, Barro Colorado Island, diversity, endophytes, environmental sampling, internal transcribed spacer, Poaceae, tropical forest

**INTRODUCTION**


Such studies traditionally have focused on graminicolous Clavicipitaceae, including *Atkinsonella, Balansia, Balansiopsis, Epichloë, Myriogenospora* and *Neotyphodium* (reviewed by Siegel et al. 1987, Saikkonen et al. 1998, Clay and Schardl 2002). Fungal endophytes in these genera are referred to as “clavicipitaceous endophytes” and recently were categorized as Class 1
endophytes by Rodriguez et al. (2009). They grow systemically within aerial tissues and share the general trait of alkaloid production, yet exhibit a diversity of transmission modes and effects on host ecology (Siegel et al. 1987, Clay and Scharald 2002, Rodriguez et al. 2009). As a group they vary in host specificity; some genera specialize on grasses in particular subfamilies or with a particular photosynthetic strategy (i.e. C_3 vs. C_4 pathway), whereas others associate with several subfamilies and/or infect hosts that differ in photosynthetic pathways and associated leaf architecture (Clay 1984, 1989, 1990; Siegel et al. 1987; Clay and Scharald 2002; Faeth and Fagan 2002).

Complementing these Class 1 endophytes are the phylogenetically diverse fungi that occur within above-ground tissues of all major lineages of land plants. These Class 3 endophytes (Rodriguez et al. 2009) occur in mosses and liverworts, ferns and allies, and diverse seed plants in agro ecosystems and natural plant communities ranging from tundra to tropical forests (e.g. Arnold and Lutzoni 2007, Arnold et al. 2009). Class 3 endophytes form localized instead of systemic infections, typically spread by horizontal transmission, and often are species rich within individual hosts and host tissues (e.g. Lodge et al. 1996, Herre et al. 2007). Their high diversity in planta presents a challenge for interpreting their costs and benefits to hosts, but investigations of several endophyte associations suggest that individual species and mixed-species assemblages can alter host physiology and defense against antagonists (e.g. Arnold et al. 2003, Arnold and Engelbrecht 2007, Herre et al. 2007). Some authors have suggested that their diversity might play an indirectly host-protective role (see Carroll 1991); dense infections by diverse fungi might generate a heterogeneous chemical landscape within and among plant tissues and individuals,
potentially constraining the evolution of specialist herbivores or pathogens or defending against diverse antagonists in species-rich settings such as tropical forests (Arnold 2008).

We are broadly interested in understanding the suite of biotic and abiotic factors that shape fungal endophyte communities. Even though most studies have focused on the incidence and importance of clavicipitaceous endophytes in grasses (reviewed by Siegel et al. 1987, Clay and Schardl 2002, Rudgers and Clay 2005), some Poaceae harbor only Class 3 endophytes, whereas others simultaneously host endophytes from both Class 1 and Class 3 (e.g. Schulthess and Faeth 1998). The relative importance of subfamily placement and associated evolutionary history of hosts, photosynthetic pathway and associated leaf architecture, introduced versus native status, pressure from antagonists and abiotic features of habitat in affecting the colonization of grasses by Class 1 and Class 3 endophytes has not been fully resolved, in part reflecting a paucity of studies examining endophytes of diverse Poaceae in their native environments. To date most studies of wild grasses have focused on grassland ecosystems, especially in the subtropics, temperate zone and boreal regions (e.g. Saikkonen et al. 2000, Zabalgogeazcoa et al. 2003). However Poaceae first arose and diversified in the shaded margins of tropical forests (Kellogg 2001), where they remain a significant and diverse component of understory and forest-edge communities (e.g. Croat 1978).

The goal of our study was to provide a broader context for understanding grass-endophyte associations by examining foliar endophytic fungi in phylogenetically diverse grasses in the understory of a lowland tropical forest. Here we used culture-based and direct-PCR approaches to examine endophytes associated with 11 co-occurring species in six subfamilies of Poaceae at Barro Colorado Island, Panama. Our sample included early-arising subfamilies that are endemic
to forest environments, as well as more recently arising subfamilies that transitioned to grassland environments. Our study provided a framework for evaluating currently available tools for designating operational taxonomic units (OTUs) from sequence data and revealed that these forest grasses harbor diverse Class 3 endophytes that are (i) predominantly host generalists with regard to grass subfamilies, clades and ancestral habitat, and (ii) shared with diverse non-grass hosts at the same site.

**MATERIALS AND METHODS**

*Study site and host taxa.*—This study was conducted at Barro Colorado Island, Panama (BCI; ~ 9°9′N, 79°51′W), which was isolated by the creation of Gatun Lake in 1914 for construction of the Panama Canal. BCI is composed of secondary forest (approximately 100 y old) and mature forest (> 400 y old) and currently is maintained as part of a forest reserve by the government of Panama. Since 1946 research on the ca. 1400 ha island has been coordinated by the Smithsonian Tropical Research Institute.

Eleven locally common species representing six subfamilies of Poaceae (sensu Barker et al. 2001, Kellogg 2001) were chosen for this study (TABLE I). All are perennial C₃ grasses that occur frequently in the understory of primary and secondary forest (Croat 1978). Anomochlooidae and Pharoideae are early-arising lineages of ancestrally forest-dwelling grasses (Kellogg 2001, Edwards and Smith 2010). Bambusoideae and Ehrhartioideae comprise forest grasses and are sister of the Pooideae, which transitioned to open habitats (Kellogg 2001). Together the three families form the ancestrally forest-dwelling BEP clade (sensu Barker et al. 2001). Centothecoideae and Panicoideae are members of the PACCAD clade (sensu Barker et al. 2001) and represent lineages that transitioned to open habitats (Kellogg 2001, Edwards and Smith 2010) (TABLE I).

*Field sampling.*—Sampling was conducted in the early rainy season, May-Jun 2006, 2007, in sites with light to medium canopy cover and a relatively open understory along BCI’s island-wide trail system. Each species was collected at six sites (except *Rhipidocladum*, which was collected at five; SUPPLEMENTARY TABLE I). Sites were spaced as widely as possible around the island for the purposes of a concurrent study (Higgins 2008); the closest sites were separated by less than 3 m and the farthest sites were 5.4 km distant (on opposite sides of the island).
Multiple host species were sampled within sites when possible. Data regarding spatial distributions of endophytes are evaluated in Higgins (2008) and Higgins et al. (in review).

Isolation of endophytes.—One overtly healthy, mature leaf was collected from each of three plants per species per site (except for a few cases in which only two individuals were available per site; Table I), stored in a clean plastic bag and transported to the lab. Within 6 h leaves were washed with running tap water and cut into 2 mm² pieces, which were surface-sterilized with sequential washes of 70% EtOH (2 min), 10% bleach (0.5% NaOCl; 2 min) and 95% EtOH (30 s) (Arnold and Lutzoni 2007). Fifteen segments per leaf (2925 segments total) were selected at random and plated onto 2% malt extract agar (MEA), which promotes growth by diverse endophytes (Fröhlich and Hyde 1999, Arnold 2002). Plates were sealed with Parafilm® and incubated at room temperature 12 wk with approximately 12 h light/dark cycles. Emergent hyphae were isolated immediately into pure culture on 2% MEA, allowed to grow 7–14 d and sorted into morphotypes based on culture morphology and pigmentation (Arnold 2002). Vouchers of all isolates were deposited in the culture collection of the International Cooperative Biodiversity Group at the Smithsonian Tropical Research Institute (Panama).

DNA extraction and sequencing.—Small pellets of mycelium were stored in 500 μL 2% SDS extraction buffer before DNA extraction following Arnold and Lutzoni (2007). DNA was extracted from all cultures; from these we selected 402 isolates for sequencing with the goal of maximizing morphological diversity in our sample. Representatives of all morphotypes were sequenced, and morphotypes were sequenced in proportion to their abundance. Isolates selected for sequencing represented all host species and study sites (Supplementary Table I).

We amplified the nuclear ribosomal internal transcribed spacers and 5.8s gene (nrITS) and ca. 600 bp of the large ribosomal subunit (LSU) as a single fragment with primers ITS5 or ITS1F (White et al. 1990, Gardes and Bruns 1993) and LR3 (Vilgalys and Hester 1990). Each 25 μL PCR reaction mixture contained 12.5 μL Sigma RedTaq (Sigma-Aldrich, St Louis, Missouri), 10 μL water, 1 μL each primer (10 μM) and 0.5 μL DNA template (Hoffman and Arnold 2008). Cycling conditions followed Higgins et al. (2007) with a 54 C annealing temperature and 90 s extension. Gel electrophoresis (1% agarose in TAE) and staining with SYBR green revealed single bands for each product.

Amplicons were cleaned, normalized and sequenced in both directions at the Genomics and Technology Core facility at the University of Arizona on an AB3730xl DNA Analyzer (Applied Biosystems, Foster City, California).
with PCR primers (5 μM). Forward and reverse reads were assembled and bases called by *phred* and *phrap* (Ewing et al. 1998) with automation provided by Mesquite (Maddison and Maddison 2007). Assembled reads were edited manually in Sequencher 4.5 (Gene Codes Corp., Ann Arbor, Michigan) and consensus sequences submitted to BLAST queries of the NCBI GenBank database for preliminary identification at higher taxonomic levels. Sequences from 402 cultured endophytes and 46 clones (below) were submitted to GenBank under accession numbers EU686744–EU687191.

*Direct PCR.*—One leaf from each of two individuals of *Panicum*, *Oplismenus*, *Orthoclada*, *Streptochaeta* and *Streptogyna*, and three individuals of *Pharus* and *Rhipidocladum*, was selected arbitrarily for direct PCR (TABLE II). Leaf tissue was surface-sterilized as described above. One ca. 0.5 cm² segment per leaf was placed in 750 μL CTAB buffer and homogenized by grinding with miniature pestle and sterile sand (Arnold et al. 2007). DNA extraction and amplification followed the methods listed above, except that the volume of template DNA was increased to 1.5 μL, the volume of water decreased correspondingly and the PCR cycle extension increased to 2 min. Sterile water was used in place of template for controls in each reaction. Studies have shown that this surface-sterilization method prevents accidental amplification of fungal DNA from the leaf surface (Arnold et al. 2007, Gallery et al. 2007). We found no evidence that ITS1F and LR3 amplified plant DNA.

Gel electrophoresis of PCR products with SYBR Green rarely yielded visible bands. To recover fungal amplicons we used a Strataclone PCR Cloning Kit (Stratagene, La Jolla, California) according to the manufacturer's instructions, except that one-half the recommended reagent volumes were used for each reaction. After blue/white screening successfully transformed colonies were transferred to new plates and incubated an additional 24 h to increase colony size. Eight positive clones per leaf were amplified in secondary PCR with primers ITS1F and LR3. Based on relative gel position after electrophoresis, up to five amplicons of different lengths were selected for sequencing as described above.

*Inference of genotype groups.*—We designated taxonomic units on the basis of percent sequence similarity (O'Brien et al. 2005, Arnold and Lutzoni 2007, Gallery et al. 2007, Hoffman and Arnold 2008). This approach provides a metric for estimating richness, diversity and composition in the absence of robust phylogenetic species definitions, which are difficult to achieve in single-site, single-guild biodiversity surveys (Arnold et al. 2007, U'Ren et al. 2009). A threshold of 95% similarity at the nrITS locus has been shown to correspond to phylogenetically delimited species
in four genera of Sordariomycetes and Dothideomycetes that are commonly recovered as tropical endophytes (Botryosphaeria, Colletotrichum, Mycosphaerella, Xylaria; U'Ren et al. 2009), all of which were recovered in this study. Incorporation of the first ca. 600 bp of LSU does not affect the quality of these species-boundary estimates (U'Ren et al. in press). Accordingly we used a threshold of 95% nrITS-partial LSU sequence to designate putative species. Because our study examined confamilial host species over a small geographic scale we used operational taxonomic units based on 99% sequence similarity (1% divergence) to obtain a more detailed view of the distribution of strains among hosts while still allowing for small amounts of sequencing error (Gallery et al. 2007).

OTUs initially were assembled with Sequencher 4.6 (99% similarity, ≥ 40% overlap; Arnold et al. 2007, 2009, U'Ren et al. 2009) and the dereplication program FastGroupII (Yu et al. 2006; http://biome.sdsu.edu/fastgroup/; percent sequence identity algorithm; PSI = 99%). The number and composition of OTUs differed markedly between Sequencher and FastGroupII outputs (see RESULTS), prompting us to validate their output through phylogenetic analyses (below). On the basis of strong phylogenetic support in two focal genera we used the much more reliable OTUs obtained from Sequencher to determine richness, diversity and host affinity, with OTUs defined by 99% nrITS-partial LSU sequence similarity.

Genotype analyses.—Taxon accumulation curves were constructed in EstimateS (Colwell 2009) with bootstrap estimates of total richness and randomized resampling to ensure comparable sample sizes and sampling intensity of hosts. Diversity was measured by Fisher's $\alpha$, which is robust to differences in sample size (see Arnold and Lutzoni 2007). Isolation frequency and diversity were compared among subfamilies, clades (defined functionally as the BEP clade, PACCAD clade and the polyphyletic group defined by early-arising lineages) and with regard to ancestral habitat use (forest, open habitat; TABLE I) by ANOVA or $t$-tests. Similarity indices based on presence-absence data (Jaccard's index) and isolation frequencies (Morisita-Horn index) of nonsingleton OTU were calculated for each pairwise combination of host species, and within versus between subfamily, clade or ancestral habitat values were compared by ANOVA. A goodness-of-fit test was used to examine the observed versus expected distribution of fungal OTU from grasses versus woody plants after comparison of sequences obtained in this study against an existing database of 1135 isolates from woody plants at BCI (Arnold and Lutzoni 2007, Arnold et al. 2009, Arnold unpubl data). Statistical analyses were conducted in R 2.6.2 (R Development Core Team 2008) with the vegan and LabDSV packages (http://www.cran.r-project.org) and in JMP 7.0.2 (SAS Institute 2007).
Phylogenetic analyses.—OTU-level analyses were complemented with phylogenetic analyses to assess the quality of genotype groupings and BLAST-based identification, as well as the relationships of endophytes obtained by culturing and direct PCR. We focused on *Anthostomella* (Xylariaceae, Xylariales, Sordariomycetes) and *Colletotrichum* (Phyllachoraceae, Phyllachorales, Sordariomycetes), two genera that were common among cultures and clones from grasses (Supplementary Table II) and as endophytes of non-grass plant hosts at BCI (Supplementary Table III).

For each genus representative sequences from GenBank were aligned in Clustal X 2.0.12 (Larkin et al. 2007) with novel sequences from endophytes at BCI. Taxon sampling of currently recognized species was based on recent studies (e.g. graminicolous *Colletotrichum*, Crouch et al. 2009) and/or drew from the available diversity of species with vouchers or high-quality GenBank records (*Anthostomella*). For *Anthostomella* the dataset comprised 11 sequences from GenBank and 43 sequences from BCI (eight isolates from non-grass hosts, 22 isolates from grasses and 13 clones from grasses). For *Colletotrichum* the dataset comprised 24 sequences from GenBank and 92 sequences from BCI (33 isolates from woody plants, 50 isolates from grasses and nine clones from grasses).

Endophytes from non-grass hosts were isolated, vouchered and sequenced as per above (host information provided in Supplementary Table III).

The alignment for each dataset was adjusted manually in Mesquite (Maddison and Maddison 2007) and evaluated in jModelTest (Posada 2008, Guindon and Gascuel 2003) to infer the appropriate model of evolution (in each case GTR + I + gamma) for Bayesian and maximum likelihood analyses. In the former two sets of four independent runs were conducted for each dataset in MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003), in each case for 3 000 000 generations with sampling every 1000th tree. Runs were terminated after the average standard deviation of split frequencies fell below 0.01. Two thousand trees from the posterior of each Bayesian analysis were used to infer a majority rule consensus tree in Mesquite (Maddison and Maddison 2007) after discarding 1002 trees per run as burn-in based on examination of −ln li values. Branch support was determined from Bayesian posterior probabilities. GARLI (Zwickl 2006) was used both for maximum likelihood inference and bootstrap analyses of branch support (100 replicates). Consensus trees from Bayesian analyses were congruent with the most likely trees obtained in GARLI and are included (Figs. 2, 3) with taxon names annotated to indicate host genus and subfamily, OTU based on Sequencher and FastGroupII and BLAST-based identification.
RESULTS

Endophytic fungi were recovered in culture from every individual examined. In sum 2264 isolates were obtained from 2925 leaf segments. Isolation frequency was ca. 44–91% of 2mm² segments per leaf per species and did not differ significantly with regard to grass subfamily, clade or ancestral habitat (Table I).

Among 402 representative isolates we recovered 245 distinct nrITS-partial LSU sequence types (based on 100% sequence similarity), of which 195 were found only once. Analysis in Sequencher indicated that these comprise 94 putative species and 124 OTUs based on 95% and 99% sequence similarity respectively (Table I). Analysis of the same dataset in FastGroupII enumerated twice as many OTUs (Table I). Cumulative diversity (Fisher's α) for the sample was 4.6-fold greater based on FastGroupII versus Sequencher-delimited OTU, and mean diversity per host species was estimated to be sixfold greater with FastGroupII (Table I).

Phylogenetic analyses for two focal genera show that Sequencher-delimited OTUs generally agreed with phylogenetic results, whereas FastGroupII markedly overestimated the number of OTUs and produced groups that were incongruent with phylogenies (Figs. 2, 3). Accordingly we used OTUs inferred in Sequencher (based on 99% sequence similarity) for further analyses.

Taxon accumulation curves were non-asymptotic, with bootstrap estimates of OTU richness falling outside the 95% confidence intervals around observed richness (Fig. 1). Overall 62.8% of OTUs were found only once. Diversity of endophytes did not differ significantly as a function of grass subfamily, clade or ancestral habitat (Table I).

Host specificity of cultured isolates.—Ninety percent of OTUs recovered more than once (i.e. nonsingletons) occurred in multiple host genera. When analyses were repeated with genotype
groups based on 100% sequence similarity 94% nonsingletons were recovered from more than one genus. Jaccard's index of similarity was significantly higher among species within the same subfamily than among hosts in different subfamilies, but the difference (JI = 0.06) is not likely to be biologically significant (Table III). We found no evidence of greater similarity based on shared membership in clades or similar ancestral habitat (data not shown). When isolation frequency was taken into account, there was no apparent structure based on host subfamily (Morisita-Horn values, Table III) or of clade or ancestral habitat (data not shown). OTUs belonging to Anthostomella and Colletotrichum showed no evidence of phylogenetic structure based on the species, subfamily, clade or ancestral habitat of hosts (FIGS. 2, 3; TABLE I).

**Taxonomic composition of cultured isolates.**—Top BLAST matches for all 402 sequences were to members of the Pezizomycotina, including 348 isolates with closest matches to fungi that were identified unequivocally to class. Examination of these results indicated that > 90% were Sordariomycetes, representing 10 orders and approximately 14–16 families. Xylariales and Phyllachorales were particularly common. The remainder included diverse lineages of Dothideomycetes and a small number of Eurotiomycetes (TABLE IV). Four isolates matched *Claviceps fusiformis* as their top BLAST hit (SUPPLEMENTARY TABLE II); however in these cases both the query coverage and maximum identity value were below 90%, non-clavicipitaceous fungi yielded similar query coverage and maximum identity values and phylogenetic analyses could not place these strains with certainty within the Clavicipitaceae (data not shown). No other sequences were consistent with clavicipitaceous fungi. Phylogenetic analyses of Anthostomella and Colletotrichum demonstrate that BLAST matches to named species markedly
 underestimated endophyte richness and frequently misidentified isolates at the species level (Figs. 2, 3).

**Direct PCR.**—Forty-six clones obtained from direct PCR yielded 29 OTUs (Table IV, Fig. 1). Sequencing of up to five clones per leaf yielded an average of 2.9 ± 1.3 OTU per 0.5 cm² leaf tissue (1–5 OTU). No sequence obtained by direct PCR was consistent with a clavicipitaceous species (Supplementary Table II).

Cloning and culturing approaches yielded similar OTU richness given a similar sampling effort. The number of OTUs obtained by cloning did not differ significantly from the number represented in random draws of 46 cultured isolates (mean ± SE from 1000 draws, each representing 46 isolates from two to three leaves of the seven host species used for direct PCR: 27.8 ± 2.7, 95% CI = 26.3–29.2 OTUs; Fig. 1). However increasing our sequencing sample by 46 clones yielded an overall increase in Fisher's α from 60.1 (124 OTUs, 402 sequences; Table I) to 70.5 (141 OTUs, 448 sequences). Despite a ninefold larger sample from culturing, > 55% of OTU obtained by cloning never were found in our much larger culturing effort (Supplementary Table II). Two classes (Pezizomycetes, Leotiomycetes) were found only by direct PCR (Table IV).

Phylogenetic analyses of *Anthostomella* and *Colletotrichum* illustrate that sequences obtained by direct PCR frequently represented either novel OTUs within clades also recovered by our larger culturing effort or OTUs that matched those recovered by culturing (Figs. 2, 3). However in *Colletotrichum* several clones represented distinct lineages relative to sequences available in GenBank and our cultured isolates (Fig. 3).
Comparison with endophyte communities of non-grass hosts.—We compared 402 sequences from grasses against 1135 nrITS sequences for endophytes of 37 non-grass species in 23 families surveyed contemporaneously for endophytic fungi at BCI (Supplementary Table IV). The resulting dataset of 1537 sequences represented 252 OTUs. Among 113 nonsingletons 34% were found only in woody plants, 54% in woody plants and grasses and 12% only in grasses. The percent of OTUs containing endophytes from grasses (75 of 113 nonsingleton genotypes, 66.4%) was significantly higher than expected given the representation of grass endophytes in the sample (28% of isolates; G₁ = 12.68, P = 0.0004), consistent with a high degree of overlap among endophytes of grasses and woody plants.

Phylogenetic analyses of Anthostomella and Colletotrichum frequently reconstructed isolates from grasses as close relatives of isolates from non-grass hosts. In general topologies do not reflect a strong signal of host taxonomy or phylogenetic relationships (Figs. 2, 3). Several well supported clades comprising only grass endophytes were recovered, but in the absence of further sampling and in light of our OTU-level analyses it is premature to conclude that these are specialist genotypes or clades that occur only in Poaceae.

Discussion
In contrast to hundreds of papers focusing on Class 1 endophytes (i.e. clavicipitaceous endophytes; Web of Science, Mar 2009), relatively few have examined the diversity of non-clavicipitaceous endophytes (Class 3) in grasses, particularly in wild species (but see Schultness and Faeth 1998, Chiang et al. 2001, Wirsel et al. 2001, Hyde et al. 2002, Mármuez et al. 2007). In grassland ecosystems, studies of undomesticated grasses frequently detect Class 1 endophytes, in contrast to the prevalence of Class 3 endophytes in related but cultivated grasses, such as maize,
rice, wheat (Tian et al. 2004, Larran et al. 2007, Pan et al. 2008, Saunders and Kohn 2009) and certain bamboos (e.g. in temperate Japan, Morakotkarn et al. 2007). Because grasses arose and first diversified in association with the shaded margins of tropical forests (Kellogg 2001) the goal of our study was to examine the endophyte communities associated with phylogenetically diverse, wild grasses in a tropical forest understory.

The few studies examining fungal endophytes of wild grasses in forests have recovered Class 1 endophytes in diverse Poaceae from temperate woodlands (e.g. Clay and Leuchtmann 1989, Schulthess and Faeth 1998). Studies have isolated clavicipitaceous endophytes on MEA (e.g. Schulthess and Faeth 1998, Marshall et al. 1999, Gentile et al. 2005, Wei et al. 2006, Moon et al. 2007), and the primers used in our direct-PCR analyses successfully amplified Class 1 endophytes in a study (U'Ren et al. in press). Although taxon-accumulation curves were non-asymptotic and only 80% of estimated genotypic richness was recovered (TABLE I, FIG. 1), our results provide strong evidence that clavicipitaceous endophytes are not a significant component of the fungal community of these forest grasses. Instead Class 3 endophytes were consistently abundant and diverse among grasses that differed in subfamily placement, clade affiliation and ancestral habitat use (TABLE I).

Consistent with studies of endophytes in woody plants at BCI (Arnold et al. 2000, 2001, 2003), more than 62% of OTUs found in this study were recovered only once. The large number of rare taxa constrains community-wide inferences of host specificity because host affiliations of only a minority of OTU can be considered. Even so we found that at least 90% of nonsingleton OTUs occurred in more than one genus of Poaceae. This host-generalism is strongly supported not only by OTU analyses based on 99% sequence similarity (TABLE III) but also by genotype
analyses based on 100% sequence identity (data not shown) and phylogenetic analyses in two focal genera (Figs. 2, 3). It is possible that nrITS-partial LSU data are insufficient to diagnose host specificity, especially in closely related and geographically proximate hosts; the limitations of this locus for fine-scale analyses in genera such as *Colletotrichum* are becoming clearer (Rojas et al. 2010) and future work will require additional loci for confirmation. However, given the tools currently available in large-scale surveys of fungal diversity, our data argue strongly for host generalism among endophytes of these forest grasses. This observation is consistent with the suggestion by Arnold and Lutzoni (2007) that the most common tropical forest endophytes appear to be host generalists (see also Cannon and Simmons 2002).

Because our study was focused on confamilial plant species within a small geographic area it might be predisposed to conclusions of host generalism; many other organisms that associate with tropical plants tend to be specialized at the level of plant family instead of species or subfamily (e.g. herbivorous insects; Barone 1998). However comparisons between communities of endophytes in grass- and non-grass hosts at BCI suggest little to no specificity at the family level. When viewed through the lens of genotype and phylogenetic composition of their endophyte communities these grasses are indistinguishable from non-grass hosts at the same site. Although phylogenetic analyses suggest that some clades might consist of only grass-inhabiting taxa, the overall dataset suggests that further sampling will reveal that these clades like others reconstructed here comprise taxa with an apparently wide host range.

When used in concert, culture-based isolation and direct amplification of fungal DNA from host tissue each reveal a unique but complementary view of the overall endophyte community (see also Allen et al. 2003, Arnold et al. 2007). Here direct PCR indicated that ca. three
genetically distinct fungi reside in each 0.5cm² piece of leaf tissue in our focal grasses. Considering the limited sampling effort per leaf and potential limitations of our primers, it is likely that species richness was underestimated. Direct PCR did not recover a greater diversity of fungi than for culture-based isolation with a comparable sampling effort but instead enabled detection of distinctive fungal OTUs that increased the overall diversity of the sample. Direct PCR also recovered two classes, Pezizomycetes and Leotiomycetes, not found in culture-based studies of diverse angiosperms at this forest site (Arnold et al. 2000, 2001, 2003; Arnold and Lutzoni 2007).

One of the major challenges in understanding fungal diversity lies in delimiting species or OTUs for meaningful ecological analyses. Discussions of such issues frequently center on two topics: first, the shortcomings of BLAST-based identification or delimitation, illustrated here by our phylogenetic analyses (Figs. 2, 3) and the misidentification of some isolates as species of *Claviceps* on the basis of BLAST hits alone; and second, the degree of sequence divergence or similarity that should be used to delimit biologically meaningful OTUs (e.g. U'Ren et al. 2009). A further complication that has received relatively little attention is that currently available algorithms differ markedly in assembling genotype groups. Comparisons between OTUs assembled with FastGroupII and Sequencher highlighted striking differences, with the former over-estimating richness and recovering OTUs that were inconsistent with phylogenetic analyses (TABLE I; Figs. 2, 3). Studies have contrasted Sequencher-based OTUs with those inferred from needle and DOTUR (Arnold et al. 2007, U'Ren et al. 2009); in each case Sequencher has provided genotype groups that match the topology of phylogenetic analyses.
Even with conservative estimates provided by Sequencher, our study recovered a high richness of Class 3 endophytes in tropical forest grasses and revealed the overlap of these fungi with co-occurring non-grass hosts. High diversity in planta is a consistent feature of endophyte communities in this species-rich forest (Arnold et al. 2000, 2001, 2003), prompting us to revisit the implications of Carroll's (1991) perspective that diverse endophyte communities might play a protective role by thwarting the evolution of specialists or impeding attacks by diverse, antagonistic generalists. If endophytes do play such a role, our data could suggest an alternative host-protection strategy for tropical forest versus grassland Poaceae, with forest grasses exploiting the same defense strategy as their dicotyledonous counterparts. Much work is needed to evaluate such a broad hypothesis and will be informed by studies of Class 3 endophytes in wild grasses across the diversity of ecosystems in which they occur.

ACKNOWLEDGMENTS
We are grateful for the help of the Smithsonian Tropical Research Institute, which provided logistical support and permission to carry out research at BCI; Panama's Autoridad Nacional del Ambiente (ANAM) for permission to export samples; M. Gunatilaka, M. Hoffman, J. U'Ren, C. Gibson, D. Mahana, B. Klein and other members of the Arnold lab group at the University of Arizona for technical assistance and helpful discussion; J. U'Ren for scriptwriting that facilitated our analyses; F. Adler and D. Feener for guidance and statistical support; F. Adler for critical evaluation of the manuscript and E.A. Kellogg for her perspective on the evolutionary origins and history of Poaceae. Financial support from the Smithsonian Tropical Research Institute (Ernst Mayr Short-Term Fellowship, KLH), a Sigma Xi Grant-in-Aid of Research (KLH), the Garden Club of America Anne S. Chatham Fellowship in Medicinal Botany (KLH), the College of Agriculture and Life Sciences (UA) and NSF DEB-0640996 (AEA) is gratefully acknowledged. This paper fulfills a portion of the MS requirements for KLH at the University of Utah. All work was done in compliance with ANAM and the laws of Panama.

LITERATURE CITED


Schulthess FM, Faeth SH. 1998. Distribution, abundance and associations of the endophytic fungal community of


LEGENDS

**Fig. 1.** Richness of endophytic fungi associated with 11 species of grasses at BCI, Panama. A. Accumulation of putative species among cultures (estimated in Sequencher with 95% sequence similarity; solid black line), 95% confidence interval (fine dotted lines) and bootstrap estimate of species richness (solid gray line). B. Accumulation of OTUs among cultures estimated in Sequencher (99% sequence similarity; observed, solid black line; confidence interval, fine dotted lines; bootstrap estimate, solid gray line). C. Accumulation of genotype groups among cultures estimated in FastGroupII (99% sequence similarity; observed, solid black line; confidence interval, fine dotted lines; bootstrap estimate, solid gray line). D. Accumulation of OTUs from culturing (large dashed line) and direct PCR
(solid black line) given a sample size of 46 sequences, with values for cultures inferred with means from 50 random draws of 46 sequences from the observed richness of cultured genotypes; 95% confidence intervals (fine dotted lines) and bootstrap estimate (gray solid line) shown for direct PCR.

**Fig 2.** Majority rule consensus based on Bayesian and maximum likelihood analyses of *Anthostomella* spp. obtained by cloning from surface-sterilized leaves of grasses (numeric codes with C-prefix; boldface italics), or by culturing from (a) grasses (numeric codes; boldface) or (b) non-grass hosts (numeric codes with P-prefix) at BCI, with exemplar sequences representing the diversity of *Anthostomella* in GenBank (marked with accession numbers). Sequences of endophytes from non-grass hosts at BCI indicate host plant species or lichen genus (*Parmelia*). Sequences of endophytes from grasses indicate host species and subfamily abbreviation (first three letters of subfamily name; see **Table I**) followed in parentheses by OTUs (based on 99% sequence similarity, inferred with Sequencher; number); group based on 99% sequence similarity, inferred with FastGroup II (letter); and top match in GenBank (species name or other taxonomic information). Support values indicate Bayesian posterior probability (before slash) and maximum likelihood bootstrap proportions (after slash); tree is rooted with endophyte 1884 (putative *Xylaria*).

**Fig. 3.** Majority rule consensus based on Bayesian and maximum likelihood analyses of putative *Colletotrichum* spp. obtained by cloning from surface-sterilized leaves of grasses (numeric codes with C-prefix; boldface italics) or by culturing from (a) grasses (number, boldface) or (b) woody plants (numeric code with P-prefix) at BCI, with exemplar sequences representing the diversity of *Colletotrichum* in GenBank (marked with accession numbers). Sequences of endophytes from non-grass hosts indicate host plant species. Sequences of endophytes from grasses indicate host species and subfamily abbreviation (first three letters of subfamily name; see **Table I**) followed in parentheses by OTUs (based on 99% sequence similarity, inferred with Sequencher; number); group based on 99% sequence similarity, inferred with FastGroup II (letter); and top match in GenBank (species name or other taxonomic information). Support values are Bayesian posterior probabilities (before slash) and maximum likelihood bootstrap proportions (after slash). Tree is rooted internally with *C. acutatum* (Crouch et al. 2009).

**FOOTNOTES**

Submitted 25 Jun 2009; accepted for publication 2 Aug 2010.
Corresponding author: E-mail: Arnold@ag.arizona.edu. Phone: (520)621-7212.
Table I. Results of a culture-based survey of endophytic fungi associated with healthy foliage of 11 species of grasses in the forest understory at Barro Colorado Island, Panama

<table>
<thead>
<tr>
<th>Host clade</th>
<th>Host subfamily</th>
<th>Ancestral habitat</th>
<th>Host species</th>
<th>Leaves</th>
<th>Isolation frequency</th>
<th>Sequenced isolates</th>
<th>OTUs Seq</th>
<th>OTUs FGII</th>
<th>Bootstrap estimate (percent recovered)</th>
<th>Fisher's α²</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Anomochlooideae</td>
<td>Forest</td>
<td>Streptochaeta spicata</td>
<td>16</td>
<td>0.51 ± 0.27</td>
<td>43</td>
<td>24</td>
<td>34</td>
<td>31 (77.4)</td>
<td>22.2</td>
</tr>
<tr>
<td>N/A</td>
<td>Pharoideae</td>
<td>Forest</td>
<td>Pharus latifolius</td>
<td>18</td>
<td>0.58 ± 0.38</td>
<td>34</td>
<td>19</td>
<td>27</td>
<td>24 (79.2)</td>
<td>17.5</td>
</tr>
<tr>
<td>BEP</td>
<td>Bambusoideae</td>
<td>Forest</td>
<td>Chusquea simpliciflora</td>
<td>18</td>
<td>0.80 ± 0.16</td>
<td>50</td>
<td>31</td>
<td>46</td>
<td>40 (77.5)</td>
<td>34.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lithachne pauciflora</td>
<td>18</td>
<td>0.44 ± 0.31</td>
<td>27</td>
<td>21</td>
<td>26</td>
<td>27 (77.8)</td>
<td>43.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Olyra latifolia</td>
<td>18</td>
<td>0.66 ± 0.26</td>
<td>54</td>
<td>24</td>
<td>43</td>
<td>29 (82.8)</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rhipidoocladium racemiflorum</td>
<td>15</td>
<td>0.91 ± 0.10</td>
<td>29</td>
<td>20</td>
<td>26</td>
<td>25 (80.0)</td>
<td>28.4</td>
</tr>
<tr>
<td>Ehrhartoideae</td>
<td>Forest</td>
<td></td>
<td>Streptogyna americana</td>
<td>18</td>
<td>0.73 ± 0.15</td>
<td>37</td>
<td>23</td>
<td>31</td>
<td>29 (79.3)</td>
<td>26.2</td>
</tr>
<tr>
<td>PACCAD</td>
<td>Centothecoideae</td>
<td>Open</td>
<td>Orthooclada laxa</td>
<td>17</td>
<td>0.53 ± 0.19</td>
<td>26</td>
<td>15</td>
<td>21</td>
<td>19 (78.9)</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>Panicoideae</td>
<td>Open</td>
<td>Ichnanthus pallens</td>
<td>18</td>
<td>0.67 ± 0.29</td>
<td>38</td>
<td>19</td>
<td>34</td>
<td>24 (79.2)</td>
<td>15.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oplismenhus hirtellus</td>
<td>18</td>
<td>0.90 ± 0.16</td>
<td>33</td>
<td>25</td>
<td>31</td>
<td>32 (78.1)</td>
<td>47.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Panicum pilosum</td>
<td>18</td>
<td>0.61 ± 0.29</td>
<td>31</td>
<td>18</td>
<td>25</td>
<td>23 (78.3)</td>
<td>18.2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>192</td>
<td>0.67 ± 0.16</td>
<td>402</td>
<td>124</td>
<td>248</td>
<td>155 (80.0)</td>
<td>60.1</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td></td>
<td></td>
<td></td>
<td>–</td>
<td>0.67 ± 0.16</td>
<td>36.5 ± 9.2</td>
<td>21.7 ± 4.3</td>
<td>31.3 ± 7.7</td>
<td>27.5 ± 5.6 (78.9 ± 1.5)</td>
<td>25.9 ± 11.5</td>
</tr>
</tbody>
</table>

Notes: Columns indicate currently recognized clades (after Barker et al. 2001), subfamilies (Barker et al. 2001), ancestral habitat use by each lineage (Kellogg 2001, Edwards and Smith 2010) and host species; the number of leaves sampled; isolation frequency (mean ± SE, calculated from the proportion of leaf segments per individual from which a fungus was isolated in culture); the number of isolates sequenced; the number of putative species (95% nrITS sequence similarity) and OTUs (99% nrITS-partial LSU sequence similarity) obtained from analysis in Sequencer (Seq) or FastGroupII (FGII); bootstrap estimates of total richness, calculated in EstimateS with Sequencer-based OTUs; and diversity (Fisher's α) based on OTUs inferred in Sequencer or FastGroupII.
b Isolation frequency did not differ significantly among subfamilies (comparisons based only on Bambusoideae and Panicoideae, from which multiple species were sampled; $F_{1,5} = 0.0294, P = 0.8705$), clades (BEP vs. PACCAD vs. early lineages, $F_{2,8} = 0.0129, P = 0.9872$), or ancestral habitat use (forest vs. open habitat; $t_9 = 0.1570, P = 0.8787$).

c Diversity did not differ significantly among subfamilies (Bambusoideae and Panicoideae only; $F_{1,5} = 0.1197, P = 0.7434$), clades ($F_{2,8} = 0.3451, P = 0.7182$), or ancestral habitat use ($t_9 = -0.4131, P = 0.6892$).
<table>
<thead>
<tr>
<th>Host</th>
<th>Leaves</th>
<th>Sequences</th>
<th>Genotype groups</th>
<th>Seq</th>
<th>FGII</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Oplismenus</em></td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><em>Orthoclada</em></td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td><em>Panicum</em></td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td><em>Pharus</em></td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><em>Rhipidocladum</em></td>
<td>3</td>
<td>13</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td><em>Streptochaeta</em></td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td><em>Streptogyna</em></td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Cumulative</td>
<td>16</td>
<td>46</td>
<td>29</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>

*aCumulative values for richness represent pooled analyses of all 46 sequences in each analysis program, instead of a sum of groups found per host species.*

**TABLE II.** Results of direct-PCR survey of endophytic fungi associated with seven focal species of grasses at BCI: number of clones sequenced for nrITS and OTUs groups based on 99% nrITS-partial LSU sequence similarity inferred in Sequencher (Seq) and FastGroupII (FGII).
TABLE III. Mean similarity values for endophyte communities of eleven grass species, calculated on the basis of presence/absence only (Jaccard's index, JI) and isolation frequency (Morisita-Horn index, MH) for nonsingleton OTUs only, as a function of host subfamily.

<table>
<thead>
<tr>
<th>Index</th>
<th>Within-subfamily</th>
<th>Between-subfamily</th>
<th>F</th>
<th>DF</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>JI</td>
<td>0.31 (0.27–0.34)</td>
<td>0.25 (0.22–0.27)</td>
<td>8.92</td>
<td>1, 53</td>
<td>0.0043*</td>
</tr>
<tr>
<td>MH</td>
<td>0.69 (0.61–0.76)</td>
<td>0.64 (0.61–0.68)</td>
<td>1.08</td>
<td>1, 53</td>
<td>0.3039</td>
</tr>
</tbody>
</table>

*aAlthough analyses of JI indicate that species in the same subfamily have significantly more similar endophyte communities than species in different subfamilies, the difference is not likely to be biologically significant.*
TABLE IV. Estimated class and ordinal placement based on verified BLAST queries for endophytic fungi isolated in culture (“isolates”) or identified with direct PCR (“clones”) from foliage of 11 grass species at BCI, Panama.

<table>
<thead>
<tr>
<th>Class</th>
<th>Order</th>
<th>Isolates</th>
<th>Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dothideomycetes</td>
<td>Botryosphaeriales</td>
<td>2.3</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Capnodiales</td>
<td>0.5</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>Dothideales</td>
<td>0.0</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Pleosporales</td>
<td>1.9</td>
<td>0</td>
</tr>
<tr>
<td>Eurotiomycetes</td>
<td>Chaetothyriales</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Eurotiales</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>Sordariomycetes</td>
<td>Boliniales</td>
<td>0.9</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Calosphaeriales</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Chaetosphaeriales</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Diaporthales</td>
<td>0.9</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Halosphaeriales</td>
<td>16.6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Hypocreales</td>
<td>3.2</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Phyllachorales</td>
<td>15.2</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>Sordariales</td>
<td>2.0</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Xylariales</td>
<td>48.0</td>
<td>38.9</td>
</tr>
<tr>
<td>Pezizomycetes</td>
<td>Pezizales</td>
<td>0</td>
<td>2.8</td>
</tr>
<tr>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>0</td>
<td>2.8</td>
</tr>
</tbody>
</table>

*Only those matches for which reliable taxonomic placement was available are included. Numbers indicate the percent of named isolates or clones represented by each order.*
Putative species (A) or genotype groups (B, C, D)