Prospects for fungus identification using *CO1* DNA barcodes, with *Penicillium* as a test case

Keith A. Seifert*†, Robert A. Samson‡, Jeremy R. deWaard§, Jos Houbraken‡, C. André Lévesque*, Jean-Marc Moncalvo¶, Gerry Louis-Seize*, and Paul D. N. Hebert§

*Biodiversity (Mycology and Botany), Environmental Sciences, Agriculture and Agri-Food Canada, Ottawa, ON, Canada K1A 0C6; ‡Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, 3508 AD, Utrecht, The Netherlands; §Biodiversity Institute of Ontario, University of Guelph, Guelph, ON, Canada N1G 2W1; and ¶Department of Natural History, Royal Ontario Museum, and Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 2C6

DNA barcoding systems employ a short, standardized gene region to identify species. A 648-bp segment of mitochondrial cytochrome *c* oxidase 1 (*CO1*) is the core barcode region for animals, but its utility has not been tested in fungi. This study began with an examination of patterns of sequence divergences in this gene region for 38 fungal taxa with full *CO1* sequences. Because these results suggested that *CO1* could be effective in species recognition, we designed primers for a 545-bp fragment of *CO1* and generated sequences for multiple strains from 58 species of *Penicillium* subgenus *Penicillium* and 12 allied species. Despite the frequent literature reports of introns in fungal mitochondrial genomes, we detected introns in only 2 of 370 *Penicillium* strains. Representatives from 38 of 58 species formed cohesive assemblages with distinct *CO1* sequences, and all cases of sequence sharing involved known species complexes. *CO1* sequence divergences averaged 0.06% within species, less than for internal transcribed spacer nrDNA or β-tubulin sequences (*BenA*). *CO1* divergences between species averaged 5.6%, comparable to internal transcribed spacer, but less than values for *BenA* (14.4%). Although the latter gene delivered higher taxonomic resolution, the amplification and alignment of *CO1* was simpler. The development of a barcoding system for fungi that shares a common gene target with other kingdoms would be a significant advance.

β-tubulin | cytochrome *c* oxidase I | DNA barcoding | internal transcribed spacer | species identification

The identification of species is a critical first step in all biological research. Correct identifications unlock the body of information known about each organism, its ecological roles, its physiological and biochemical properties, and its societal risks or benefits. Precise identifications of species have historically been the realm of taxonomic experts. Each taxonomic group has an attendant body of specialized literature and terminology that evolved to describe morphological characters so that species can be recognized by scientists skilled in the art. The rise of identification systems based on diversity in DNA sequences represents a significant advance. Because DNA methods use shared techniques and language, it is possible (at least in theory) for practitioners to correctly identify species from all kingdoms. As a consequence, the identification and enumeration of all organisms in any environmental setting is now a possibility.

An accurate, rapid, cost-effective, and universally accessible identification system is needed for fungi. Recent estimates suggest that 1.5 million species of fungi exist, but <10% are formally described (1). The frequent lack of distinctive morphological characters, the preponderance of microscopic species, and the considerable socioeconomic importance of this kingdom reinforce the need for a DNA-based identification system. Identifications of species with molecular techniques are now routine in some groups of fungi, but little attention has been paid to standardization. DNA-based systems for species of fungi have variously used a barcode-like 400- to 600-bp region of the nuclear large ribosomal subunit (2), the internal transcribed spacer (ITS) cistron (e.g., refs. 3 and 4), partial β-tubulin A (*BenA*) gene sequences (5), or partial elongation factor 1-α (EF-1α) sequences (6), and sometimes other protein-coding genes.

The concept of DNA barcoding proposes that effective, broad spectrum identification systems can be based on sequence diversity in short, standardized gene regions (7–9). To date, this premise has been tested most extensively in the animal kingdom, where a 648-bp region of the cytochrome *c* oxidase 1 (*CO1*) gene consistently delivered species-level resolution in more than 95% of taxa from test sets of different animal lineages (10–12). The few cases of incomplete resolution involved closely allied species and were offset by the revelation of new species (13, 14). This success has now provoked several large-scale studies on animals to develop barcode libraries for complete taxonomic groups, such as birds and fishes (15). The effectiveness of DNA barcoding in animals raises the question of how universal a *CO1*-based system might be. Early work on marine algae (16) revealed that *CO1*-based systems are promising for this group. In contrast, this gene will be less effective for land plants because of their slower mitochondrial evolution, and efforts are underway to identify alternate gene regions that will deliver species-level resolution (17–19). The potential effectiveness of *CO1* in species identification of fungi has not yet been evaluated, and the present study represents a first step to address this gap.

There is limited but growing information on fungal mitochondrial genomes (20). Existing studies of complete fungal mitochondrial genomes reveal a potential complication to PCR-based surveys of *CO1* sequence diversity, i.e., the prevalence of mobile introns (21). Therefore, we assessed the incidence and sizes of introns in the barcode region of *CO1* from different fungal lineages, based on data in GenBank, to see whether their presence might lengthen target amplicons beyond what is easily recoverable by conventional PCR.

We also examined a second, more critical issue, i.e., the nature and extent of *CO1* sequence diversity among closely related species of fungi. This analysis focused on species of *Penicillium* subgenus *Penicillium* (Trichocomaceae, Eurotiales, Eurotiomycetes, Ascomycota), a monophyletic group of moulds that rep-

EVOLUTION

resent 58 of the ≈250 accepted species in the genus *Penicillium* (22). The species of this subgenus include many sources of antibiotics, producers of mycotoxins, agents of plant disease and food spoilage, and beneficial species used in cheese manufacture [supporting information (SI) Dataset 1]. These species reproduce asexually and are phylogenetically related to (but independently named from) species with sexual states classified in *Eupenicillium*. The polyphasic species concept applied in the most recent monograph (23) used micromorphology, mycotoxin profiles, physiological tests (such as growth on diagnostic agar media), and ecological behavior. These species have not yet been subjected to the multilocus sequencing necessary to propose phylogenetic species concepts (24). Tests of ITS showed that it was insufficiently variable to reliably discriminate species (25). The more variable *BenA* was adopted as a species marker in some studies (5, 26) and is the only gene so far sequenced for all species of the subgenus. Despite its overall superior performance compared with the ITS, there are still several species complexes whose members cannot be distinguished by *BenA* sequences. This is one of the few parts of the fungal kingdom where phenotypically defined species diversity exceeds that revealed by molecular analysis. For these reasons, we consider subgenus *Penicillium* a robust test for *CO1* barcoding.

This article begins with a broad-scale analysis of the structure of the *CO1* gene region in fungi, with particular emphasis on intron position and size. We developed primers for amplifying *CO1* from *Penicillium* species, and generated *CO1* sequences for 58 species in subgenus *Penicillium*, as well as 12 allied taxa. Levels of inter- and intraspecific variation in *CO1* within this subgenus were subsequently compared with variation in other genes and with variation in other groups of organisms.

Results

Coarse Scale Analysis. Analyses of complete fungal *CO1* sequences demonstrated that the length of this gene varied from 1,584 to 22,006 bp. This size variation largely reflected the varying number and length of introns; the coding region of *CO1* varied from just 1,584 to 1,905 bp. These introns occurred at different positions in the *CO1* gene, and from one to seven of these sites were in the region used for barcoding of animals. Fig. 1*a* shows that the representatives of some groups, such as the Oomycetes (fungus-like organisms now classified in the kingdom Straminipila), have no reported introns, whereas some fungal classes have as many as seven. For the Eurotiomycetes, which includes the family Trichochomaceae and the genus *Penicillium*, 0–2 introns were reported in the barcode region (27). The length of the introns also varies considerably (Fig. 1*b*), ranging from 134 bp to 3.1 kb. As a result, the barcode region potentially ranges from 642 bp in *Hypocrea jecorina* to 12.3 kb in *Podospora anserina*. Despite the few taxa available for analysis, the neighbor-joining dendrogram of *CO1* divergences showed high cohesion for representatives of major fungal lineages. More significantly, the dendrogram revealed deep divergences between species in the seven genera where two or more species were analyzed (mean Kimura two-parameter percent divergence, 12.9 ± 1.2%; range, 0.6–34.8).

Fine Scale Analysis. Forty-two species of *Penicillium* subgenus *Penicillium* had invariant *CO1* sequences among all sequenced strains (Fig. 2). Fourteen species exhibited variation, namely *P. aurantiogriseum* (5 strains in 2 groups), *P. nordicum* (8/2), *P. venetum* (5/2), *P. clavigerum* (5/2), *P. coprobium* (5/2), *P. coprophilum* (5/2), *P. expansum* (5/3), *P. glandicola* (5/3), *P. carneum* (4/2), *P. mononematosum* (6/2, paraphyletic with the single strain of *P. confertum*), *P. nalgiovense* (5/3, paraphyletic with *P. dipodomyis*), *P. vulpinum* (6/2), *P. atramentosum* (7/4), and *P. olsonii* (5/2). For *P. persicinum* and *P. formosanum*, only single strains were available. Among the outgroup species, there was

variation among strains of *P. citrinum*, *P. spinulosum*, and *P. glabrum*.

Thirty-eight of the 58 species in subgenus *Penicillium* formed cohesive assemblages that had distinct *CO1* sequences from all other species. Most cases of sequence sharing involved species complexes (4 complexes, a total of 17 species). The largest of these groups included the seven members of the *P. aurantiogriseum* complex (*P. aurantiogriseum*, *P. cyclopium*, *P. freii*, *P. melanoconidium*, *P. neoechinulatum*, *P. polonicum*, and *P. viridicatum*). The five members of the *P. camemberti* complex, consisting of that species and *P. caseifulvum*, *P. cavernicola*, *P. commune*, and *P. palitans*, also had identical *CO1* sequences. Finally, the three members of the *P. hirsutum* complex (*P. albocoremium*, *P. allii*, and *P. hirsutum*) shared sequences, as did the two plant pathogenic species, *P. radicicola* and *P. tulipae*.
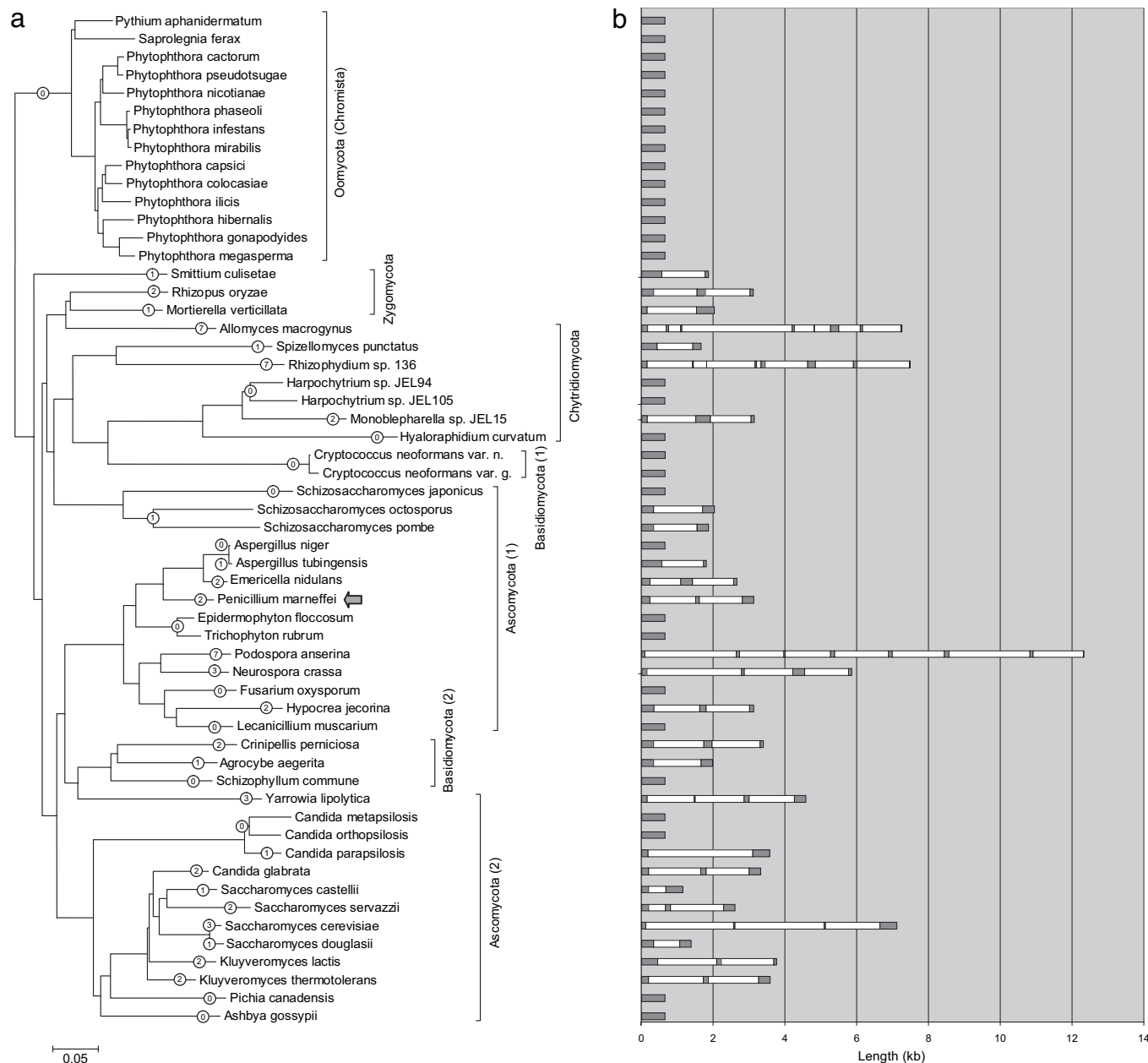
An 1,183-bp intron in *Penicillium pinophilum* KAS 1773 was amplified by PCR using PenF2 and PenR1. BLASTn searches on GenBank showed that a 682-bp stretch of this intron had 92% similarity with the *CO1* intron of *Penicillium marneffei* [GI33860251 (27)]. Five segments varying from 51 to 114 bp had from 80% to 87% similarity with a *CO1* intron of *Aspergillus japonicus* [GI4894617 (28)]. Another intron (≈1,100–1,200 bp) was detected in *P. arenicola* DAOM 229808 by using primers PenF1 and PenR2, but we were unable to obtain a complete sequence for it. An ≈150-bp segment of the single-stranded sequence had 90–94% homology with *CO1* introns in four species of *Aspergillus* (*A. oryzae*, *A. niger*, *A. tubingensis*, and *A. nidulans*) and 88% similarity with *P. marneffei*.
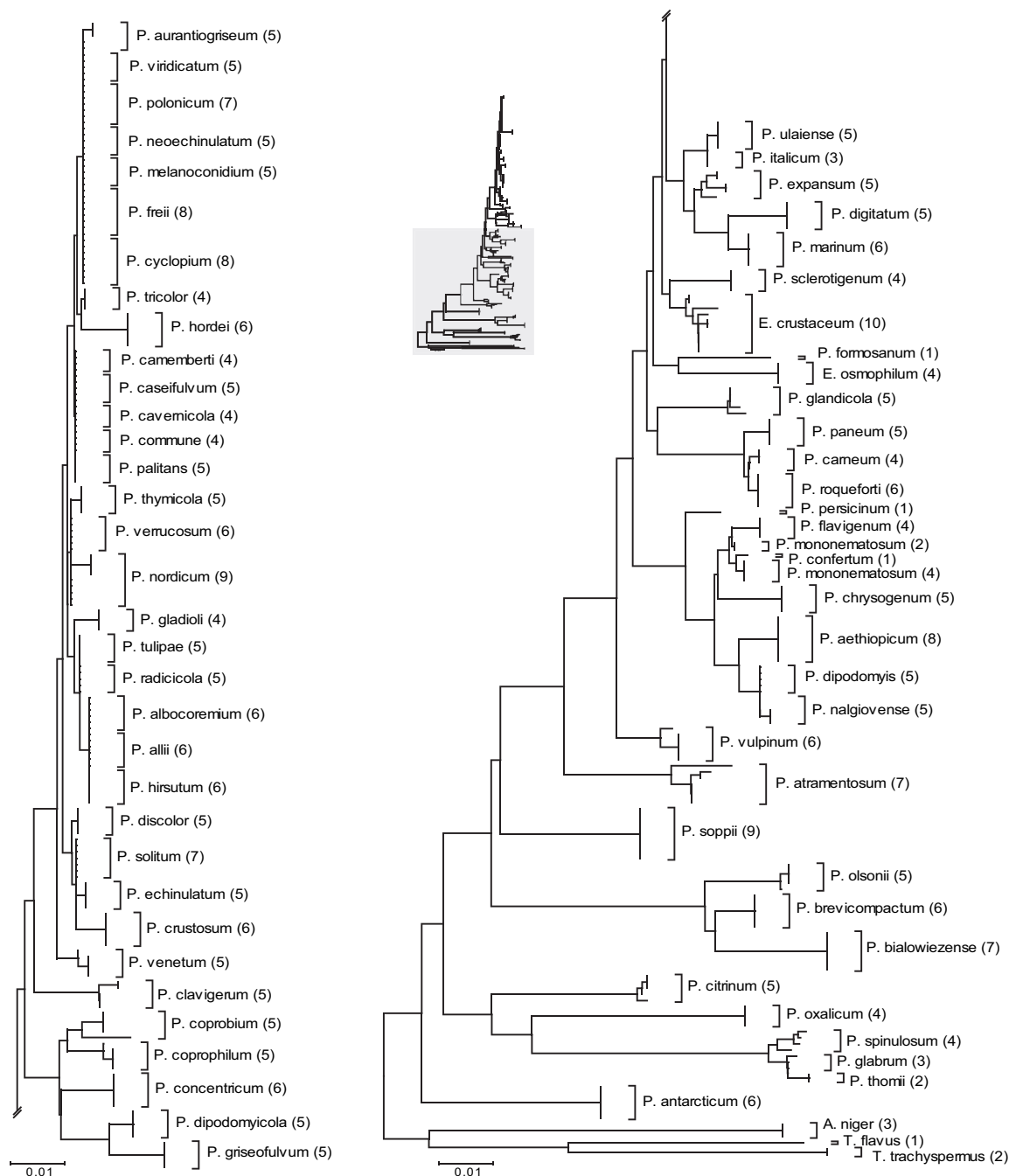
Comparison with Other Candidate Markers. Table 1 compares some key aspects of the sequence variation in three gene regions (*BenA*, ITS, and *CO1*) that we considered as the basis for a barcode system in *Penicillium*. These results indicate that both *BenA* and ITS show a high incidence of indels that complicate alignments within the subgenus, whereas no indels were detected in *CO1*. The taxonomic resolution of the regions varied, with *BenA* providing greater resolution than ITS or *CO1* (SI Fig. 3).

Discussion

Past work has shown that fungal mitochondrial genomes have many introns and our analyses confirm that some occur within the barcode region of the *CO1* gene. Based on their prevalence and size in fungal lineages, we anticipated that introns would regularly complicate PCR-based recovery of the target region of *CO1* from *Penicillium*, but we encountered *CO1* introns in only two *Penicillium* species. It is possible that members of subgenus *Penicillium* have an unusual scarcity of introns, but we note that introns have been revealed mostly through sequencing studies on whole mitochondrial DNA molecules harvested in bulk by density gradient centrifugation. Perhaps our failure to detect introns indicates that a small fraction of *CO1* sequences are stripped of introns for part of their life cycle, providing templates that can be amplified by PCR. Some members of the *Aspergillus niger* complex, in the same ascomycete family as *Penicillium*, have mobile mitochondrial introns that might be excised from some copies (28, 29). Furthermore, many reports of mobile introns in the mitochondrial genomes of fungi originate from laboratory strains subjected to mutagenesis or protoplast fusion to provoke mitochondrial recombination and thus may not be representative of wild-type strains. A concentrated effort to PCR amplify and sequence introns where they are expected (i.e., Fig. 1*b*) would address this possibility. In addition, this would allow an assessment of the variation in the intronic regions among closely related species, to determine whether *CO1* introns contain some phylogenetic signal and might themselves be useful species markers.

The *CO1* data generated for *Penicillium* subgenus *Penicillium* resulted in species-specific sequences (invariant among the

Fig. 1. Number, size, and position of introns in *CO1* of fungi and fungus-like Oomycota. (*a*) Neighbor-joining tree of 56 fungal species for which complete mitochondrial genomes and/or complete *CO1* genes are available in GenBank. The tree is constructed with the 648-bp *CO1* barcode region. The circled numbers indicate the number of introns in the barcode region for all taxa to the right of the symbol, and the arrow indicates the approximate taxonomic location of the fine-scale analysis. (*b*) Exon and intron lengths in the 648-bp *CO1* barcode region for the 56 fungal species. Exons are shaded dark gray, and introns are shaded white. Exon–intron boundaries are based on annotations provided by the GenBank records.

strains, or with some terminal branching within the species) for ≈2/3 of the species. There were four complexes of species where all strains either had identical sequences or only minor variations that were not species-specific. Our comparisons of the species-resolving power of *CO1*, *BenA*, and ITS are qualified by the few ITS studies available where multiple strains of *Penicillium* were sequenced. There are several closely related species assemblages where the results from *CO1* and *BenA* are concordant. For example, the *P. roqueforti* complex includes three species that can easily be distinguished by *BenA* (5) and ITS (30), and with *CO1*. Similarly, the six fruit-pathogenic species of *Penicillium* all have unique *CO1* barcodes and unique *BenA* sequences. However, there were other cases where *CO1* did not discriminate lineages that could be distinguished with *BenA*, as discussed below.

The *P. aurantiogriseum* complex includes seven species. *BenA* distinguishes all except *P. freii* and *P. neoechinulatum*, although some of the other species are separated by only a few base differences. The results from ITS are unclear. Skouboe *et al.* (25) identified six nucleotide substitutions and one insertion among the ITS of the single strains of each species in this complex, but Peterson (31) found only one substitution in the ITS of four type strains. The virtually invariant *CO1* sequences for this complex reveal less variation than either *BenA* or ITS but lend support to the conclusion of Pitt (32, 33) that many of these species should be considered synonyms of *P. aurantiogriseum*. The *P. verrucosum* complex presently includes three species, each with distinct *BenA* sequences, but no ITS data are available for two of the

Fig. 2. Neighbor-joining tree of 70 species and 354 isolates from the subgenus *Penicillium* and related species. The tree is constructed with 545 bp of *CO1*. Bracketed numbers represent the number of strains sequenced for each species.

species. Of these three species, *P. thymicola* has a unique *CO1* barcode. *P. verrucosum* and three strains of *P. nordicum*, both species producing the regulated mycotoxin ochratoxin A (OA), have identical *CO1* barcodes. However, four *P. nordicum* strains emerge from this cluster, with their own unique barcode. Interestingly, this same dichotomy among *P. nordicum* strains is seen in *BenA*. Frisvad and Samson (23) considered designating this group as a distinct species from *P. nordicum* but did not describe it. The *P. camemberti* complex includes four species. There has been no ITS analysis of this complex, but *BenA* sequences distinguish the cheese contaminants *P. commune* and *P. palitans* but not the two cheese-producing species *P. camemberti* and *P. caseifulvum*. These latter two species are very similar,

differing in metabolites and conidial color (white in *P. camemberti*, and grayish in *P. caseifulvum*). By contrast, *CO1* sequences from all 22 strains of this complex were identical. There has been speculation that *P. commune* is the ancestral species of *P. camemberti*, which could then be considered a domesticated variant (23).

Our results suggest that a broader exploration of *CO1* diversity in fungal lineages is justified. Our analysis of *CO1* sequences from whole-genome studies indicates that sequence divergences in this gene are deep for fungi, mirroring the pattern seen in animals and protists rather than the shallower divergence seen in plants. Although introns large enough to complicate PCR-based approaches were prevalent in whole-mitochondria studies,

Table 1. Comparison of three potential markers for fungal DNA barcoding

Marker	No. of isolates	No of. species	Mean sequence length	Mean intraspecific divergence, %	Range of means, %	Mean interspecific divergence, %	Range of means, %	Species resolution,* %	Mean no. of introns (approx.)	Mean no. of gaps in alignment
CO1	354	70	545	0.06	0–0.6	5.6	0–15.4		67.1	0
BenA	249	64	446	0.83	0–4.9	14.4	0.5–67.2	81.2	3	258
ITS	282	49	555	0.51	0–5.3	5.1	0–26.8		24.5	0

*A species is considered resolved if all of its constituent sequences form a monophyletic cluster and are distinct from other sequences.

we rarely encountered problems in recovering a PCR product for members of the subgenus *Penicillium*, perhaps reflecting the excision of introns from some portion of the mitochondrial molecules. However, even if introns prove common in other groups, alternate methodologies, such as reverse transcription PCR, could overcome their impact on amplicon size. Fungal *BenA* typically contains numerous introns (34), as do the 18S and 28S ribosomal genes that flank the ITS region and contain the priming sites for its amplification (35). A final factor supporting further work on *CO1* is the simple demonstration of its effectiveness in discriminating species of *Penicillium*, a taxonomically challenging group of fungi. Sequence divergences were lower in *CO1* than in *BenA*, but the difference in resolution was not dramatic and cases of compromised resolution invariably involved closely allied taxa. In mycological studies to date, phylogenetic species concepts (24) have been applied mostly to the exploration of species complexes rather than genera or subgenera. Further comparisons of *CO1* with the level of variation detected in such multigene studies will be an important further test for this kind of barcode for fungi.

CO1 has two important attributes that favor its candidacy as a fungal barcode marker. First, *CO1* sequences can be aligned with certainty across all fungal lineages, which facilitates development and vetting of data. Alignments for ITS and protein-coding genes such as *BenA* are often problematic, even within a single subgenus. Robust alignments mean that sequence information need not simply be the basis of an identification system but can also provide insights into rates of molecular or protein evolution. Sequencing errors, namely unexpected deletions or double bases, are also obvious. Second, many benefits, for both methodological and environmental applications, will arise if *CO1* can be used as a core identification system for fungal, animal, and protist lineages. Although genetic minimalism and standardization are central principles of DNA barcoding, there is no proscription against employing additional genetic information to resolve identifications or questions of species boundaries. Diagnosticians know that a single assay rarely answers all questions, and the results of one standardized test often provoke a more specific test. If our results on *Penicillium* prove representative of those for other fungi, *CO1* will deliver a reliable species identification for most taxa. When needed, its resolution can be augmented by the sequencing of other genes.

Materials and Methods

Coarse Scale Analysis. We downloaded all complete *CO1* genes (primarily from complete mitochondrial genomes) for *Ascomycota* (27) (mostly Saccharomycotina), Basidiomycota (5), Chytridiomycota (7), Zygomycota (3) (Eumycota), and Oomycota (14) (Straminipila) available in GenBank release 150 (November 2005). The sequences were cropped to the 648-bp barcode region: amino acids 19–234 of the bovine (*Bos taurus*) *CO1* gene to allow for a determination of levels of sequence divergence. We then determined the position and size of introns within this gene region from annotations in the GenBank flat files. The taxonomic sample and GenBank accessions are given

in SI Dataset 2 and are available in the public project "GenBank Fungi (*CO1*)" in the Barcode of Life Data System (www.barcodinglife.org).

Fine Scale Analysis. To investigate *CO1* variation at a finer scale, we concentrated on *Penicillium* subgenus *Penicillium* and some closely related species. Strains were obtained from several sources including the Canadian Collection of Fungal Cultures and Centraalbureau voor Schimmelcultures. Strains were grown on malt peptone broth by using 10% (vol/vol) of malt extract (Brix 10) and 0.1% (wt/vol) bacto peptone (Difco) in 2 ml of medium in 15-ml tubes. Growth conditions were 24°C for 7 d in darkness.

We included between 1 and 10 strains of the 58 species for which *BenA* sequences were analyzed by Samson *et al.* (5), including the three to four strains of each species included in that study, along with 12 additional species representing other subgenera of *Penicillium* or its sister genus *Aspergillus*. Although additional strains are available for some common, widespread species (e.g., *P. chrysogenum*, *P. commune*, *P. crustosum*, and *P. expansum*), in general our sampling in this subgenus exceeds that seen in most species-level studies in fungi. The taxonomic sampling, collection accessions, and GenBank accession numbers are given in SI Dataset 3 and are also available online in the public project "Penicillium subgenus Penicillium" in the Barcode of Life Data System. DNA was extracted for each strain by using the FastDNA kit (BIO 101) according to the manufacturer's instructions.

PCR primers were designed from a ClustalW alignment for the only three species of the family Trichocomaceae with *CO1* sequences available from GenBank, *Penicillium marneffei* (GI33860251), a member of subgenus *Biverticillium*, and two species of *Aspergillus*, i.e., *A. tubingensis* (GI55925030, GI12751551) and *A. niger* (GI9971736, GI9957432, GI9957434, GI9957436). The conserved regions of interest were rather AT-rich. One forward primer (PenF1, 5′-GACAAGAAAGGT-GATTTTTATCTTC) was anchored at site 104 (numbering relative to the intron-free mRNA *CO1* sequence from *P. marneffei*). Two reverse primers were anchored at site 674, one designed to match *P. marneffei* (PenR1, 5′-GGTAA*A*GATAATAATA-ATAACAC*T*GCTG) and the other to match the *Aspergillus* species (reverse primer AspR1 5′-GGTAA*T*GATAATAATA-ATAATAC*A*GCTG). Two sequencing primers were positioned at site 410 (PenF2 5′-TWAGTTTCTGATTATTAGTAC-CTAGTTT) and its reverse complement (PenR2 5′-AAAC-TAGGTACTAATAATCAGAAACTWA), between two sites noted to have introns in the *CO1* sequence of *P. marneffei*.

We tested the performance of PenF1 with both PenR1 and AspR1 in PCR amplifications using genomic DNA isolated from ten species of the family *Trichocomaceae*, namely *Aspergillus flavus* DAOM 144287, A. *puniceus* 03_892_6O, *Paecilomyces variotii* KAS1238, *Penicillium arenicola* DAOM 229808, *P. dendriticum* DAOM 233861 (subgenus *Biverticillium*), *P. digitatum* DAOM 226855 (subgenus *Penicillium*), *P. italicum* DAOM 226891 (subgenus *Penicillium*), *P. paxilli* DAOM 233863 (subgenus *Furcatum*), *P. pinophilum* KAS1773 (subgenus *Biverticil-*

EVOLUTION

lium) and *P. spinulosum* KAS1780 (subgenus *Aspergilloides*). Preliminary PCR amplications using PenF1 and PenR1 were sporadically successful, whereas amplications using AspR1 had a 100% success rate. This reflects the fact that the species of *Aspergillus* used to design the primers, *A. niger* and *A. tubingenense*, are phylogenetically more closely related to *Penicillium* subgenus *Penicillium* than is *P. marneffei*, which is related to the subgenus of *Penicillium* with *Talaromyces* sexual states (36).

Amplification, sequencing, and sequence analysis closely followed standard methods (37). PCRs were performed in 11.5-μl volumes containing 8.25 μl of PCR-grade water, 1.75 μl of 10× PCR buffer (New England BioLabs), 0.625 μl of MgCl$_2$ (50 mM), 0.125 μl of each primer (10 μM), 0.0625 μl of dNTPs (10 mM), 0.0625 μl of Taq polymerase (5 units/μl), and 1.0 μl of DNA extract. The cycling conditions were an initial step of 3 min at 95°C, 35 cycles of 60 s at 95°C, 45 s at 56°C, and 90 s at 72°C, followed by 10 min at 72°C. The primers *PenF1 and AspR1* were used to amplify a 545-bp fragment of *CO1* for all samples.

PCR products were directly sequenced with the same primers in 10-μl reactions containing 5 μl of 10% trehalose, 1.875 μl of 5× sequencing buffer, 1.0 μl of primer (10 μM), 0.875 μl of PCR-grade water, 0.25 μl of BigDye v3.1 terminator mix (Applied Biosystems), and 1.0 μl of PCR product. Primers PenF2 or PenR2 were also used when an intron was found. The cycling conditions were an initial step of 2 min at 96°C, followed by 30 cycles of 30 s at 96°C, 15 s at 55°C, and 4 min at 60°C. Sequencing reactions were performed in both directions by using the PCR primers. Sequencing products were purified with Sephadex G-50 (Sigma) columns in multiscreen HV filter plates (Millipore) and then run on an Applied Biosystems 3730 DNA analyzer. Resultant sequences were assembled, edited, and aligned in SeqScape V.3.0 (Applied Biosystems) before being uploaded to the Barcode of Life Data System.

Comparison with Other Candidate Markers. We downloaded *BenA* and ITS sequences from GenBank, if available, for the 70 species that were investigated for *CO1*. Only sequences >300 bp in length were retained for analysis. In total, 249 sequences of *BenA* (for 64 species) and 282 sequences of ITS (for 49 species) were downloaded, aligned by using ClustalW with BioEdit (38), and analyzed in MEGA3 (39). For all analyses, nucleotide-sequence divergences and dendrograms were calculated with the neighbor-joining algorithm and the Kimura two-parameter model; the performance of this method is comparable with other techniques when distances are low (40) and large species assemblages are analyzed (41).

1. Hawksworth DL (1991) *Mycol Res* 95:641–655.
2. Kurtzman CP, Robnett CJ (1998) *Ant Leeuwenhoek* 73:331–371.
3. Druzhinina I, Kopchinskiy AG, Komon M, Bissett J, Szakacs G, Kubicek CP (2005) *Fungal Genet Biol* 42:813–828.
4. Kõljalg U, Larsson K-H, Abarenkov K, Nilsson RH, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, *et al.* (2005) *New Phytol* 166:1063–1068.
5. Samson RA, Seifert KA, Kuijpers AFA, Houbraken JAMP, Frisvad JC (2004) *Stud Mycol* 49:175–200.
6. Geiser DM, Jiménez-Gasco M, Kang S, Makalowska I, Veerarahavan N, Ward T, Zhang N, Kuldau G, O'donnell K (2004) *Eur J Plant Pathol* 110:473–479.
7. Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) *Proc R Soc London Ser B* 270:313–321.
8. Blaxter M (2003) *Nature* 421:122–124.
9. Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) *Philos Trans R Soc London B* 360:1805–1811.
10. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) *Philos Trans R Soc London B* 360:1847–1857.
11. Smith MA, Fisher BL, Hebert PDN (2005) *Philos Trans R Soc London B* 360:1825–1834.
12. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) *Proc Natl Acad Sci USA* 103:968–971.
13. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) *PLoS Biol* 2:e312.
14. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) *Proc Natl Acad Sci USA* 101:14812–14817.
15. Marshall E (2005) *Science* 307:1037.
16. Saunders GW (2005) *Philos Trans R Soc London B* 360:1879–1888.
17. Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, Savolainen V (2005) *Phil Trans R Soc London B* 360:1889–1895.
18. Kress JW, Wurdack KJ, Zimmer EAC, Weigt LA, Janzen DH (2005) *Proc Natl Acad Sci USA* 102:8369–8374.
19. Newmaster SG, Fazekas A, Ragupathy S (2006) *Can J Bot* 84:335–341.
20. Clark TA, Moncalvo JM (2005) in *The Biodiversity of Fungi—Their Role in Human Life*, eds Deshmukh SK, Rai MK (Science Publishing, Enfield, NH), pp 15–32.
21. Hamari Z, Juhász Á, Kevei F (2002) *Acta Microbiol Immunol Hung* 49:331–335.
22. Pitt JI, Samson RA (2000) in *Integration of Modern Taxonomic Methods for Penicillium and Aspergillus*, eds Samson RA, Pitt JI (Harwood Academic, Amsterdam), pp 51–72.
23. Frisvad JC, Samson RA (2004) *Stud Mycol* 49:1–174.
24. Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, Fisher MC (2000) *Fungal Genet Biol* 31:21–32.
25. Skouboe P, Frisvad JC, Taylor JW, Lauritsen D, Boysen M, Rossen L (1999) *Mycol Res* 103:873–881.
26. Seifert KA, Louis-Seize G (2000) in *Integration of Modern Taxonomic Methods for Penicillium and Aspergillus*, eds Samson RA, Pitt JI (Harwood Academic, Amsterdam), pp 189–198.
27. Woo PC, Zhen H, Cai JJ, Yu J, Lau SK, Wang J, Teng JL, Wong SS, Tse RS, Chen R, *et al.* (2003) *FEBS Lett* 555:469–477.
28. Hamari Z, Juhász Á, Gácser A, Kucsera J, Pfeiffer I, Kevei F (2001) *Fungal Genet Biol* 33:83–95.
29. Juhász A, Láday M, Gácser A, Kucsera J, Pfeiffer I, Kevei F, Hamari Z (2004) *FEMS Microbiol Lett* 241:119–126.
30. Boysen M, Skouboe P, Frisvad J, Rossen L (1996) *Microbiology* 142:541–549.
31. Peterson SW (2000) in *Integration of Modern Taxonomic Methods for Penicillium and Aspergillus*, eds Samson RA, Pitt JI (Harwood Academic, Amsterdam), pp 163–178.
32. Pitt JI (1979) *The Genus Penicillium and its Teleomorphic States Eupenicillium and Talaromyces* (Academic, London).
33. Pitt JI (1988) *A Laboratory Guide to Common Penicillium Species* (Food Science Australia, North Ryde, Australia), 2nd Ed.
34. Donaldson GC, Ball LA, Axelrood PE, Glass NL (1995) *Appl Environ Microbiol* 61:1331–1340.
35. Martin KJ, Rygiewicz PT (2005) *BMC Microbiol* 5:28.
36. Berbee ML, Yoshimura A, Sugiyama J, Taylor JW (1995) *Mycologia* 87:210–222.
37. Hajibabaei M, deWaard JR, Ivanova NV, Ratnasingham S, Dooh R, Kirk SL, Mackie PM, Hebert PDN (2005) *Philos Trans R Soc London B* 360:1959–1967.
38. Hall TA (1999) *Nucleic Acids Symp Ser* 41:95–98.
39. Kumar S, Tamura K, Nei M (2004) *Brief Bioinform* 5:150–163.
40. Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics* (Oxford Univ Press, New York).
41. Kumar S, Gadagkar SR (2000) *J Mol Evol* 51:544–553.