INTRODUCTION: Where things stand currently, and how we got here

The last complete monograph of Aspergillus was written in 1965 (Raper & Fennell 1965). They recognised 132 species and 18 varieties. This was a pragmatic treatise. Although they provided Latin descriptions for new species, they did not designate formal type specimens, instead using type strains. They described teleomorphs under the anamorph (Aspergillus) name. They divided the species into 18 informal "groups" based on the authors' opinions of probable relationships. The "groups" they established, which largely reflected groups defined in previous treatises (Thom & Church 1926; Thom & Raper 1945), have been amazingly stable through a great deal of morphological, physiological and molecular scrutiny over the subsequent 42 years. In an effort to bring the taxonomy of the genus in line with the International Code of Botanical Nomenclature, the species were typified (Samson & Gams 1985), the groups were revised and given formal taxonomic status as sections, and subgenera were added (Gams et al. 1985).

A number of changes have been made to the infrageneric taxa proposed by Gams et al. (1985). In his revision of the genus based on rDNA sequences, Peterson (2000) proposed eliminating three of the six subgenera established by Gams et al. (1985), retaining 12 of the 18 sections, modifying three of the sections and deleting the other three. Frisvad et al. (2005) proposed Section Ochraceorosei to accommodate the species A. ochraceoroseus and A. rambelli. The genus Neopetromyces was proposed by the same authors for the teleomorph of A. muricatus (Frisvad & Samson 2000).

Species named since the turn of the century are listed in Table 1. Raper and Fennell (1965) described 150 taxa in their monograph; the latest compilation of names in current use (Pitt et al. 2000) listed 182. Samson (2000) listed another 36 published between 1992 and 1999. More than 40 new species descriptions have been published since then (Table 1), bringing the total number to ~250. The number will continue to grow as we discover new species and refine species concepts.

Aspergillus is one of the most economically important genera of microfungi, so the rigor and stability of its taxonomy is of significant practical concern. We joke that everyone trying to break a patent is a "splitter", and everyone trying to retain a patent is a "lumper", but there is some truth in the joke. Some aspergilli produce metabolites toxic to animals that are highly regulated in many countries while other aspergilli cause allergies or mycoses. It is important that we carefully consider any taxonomic changes in order to keep the taxonomic system practical for economic and regulatory reasons. This has been accomplished through open discussion and consensus-building in meetings such as the Aspergillus Workshops, and in efforts such as the lists of accepted species and synonyms (Pitt et al. 2000).

Aspergillus taxonomists have a long tradition of an eclectic approach to our discipline. This is reflected in the methods used by many researchers to delineate new species (Table 1). In addition to morphological characters - the shapes and sizes of various structures, we have traditionally used physiological characters such as colony diam and production of coloured metabolites in our taxonomic systems. We now also use data on other metabolites as measured by chromatography and mass spectroscopy. As a rule, molecular data have supported relationships previously inferred based on morphological and physiological characters.
Table 1. Characters used in delineating new species of Aspergillus since 2000.

| Section     | Species              | Morphology | Physiology | Molecular     | Reference                                      |
|-------------|----------------------|------------|------------|---------------|-----------------------------------------------|
| Circumdati  | A. persii            | i          | x          | BenA          | Zotti & Corti (2002)                          |
|             | A. cretensis         | i          | x          | BenA          | Frisvad et al. (2004)                         |
|             | A. flocculosus       | i          | x          | BenA          | Frisvad et al. (2004)                         |
|             | A. neobridgeri       | i          | g          | BenA          | Frisvad et al. (2004)                         |
|             | A. pseudoelegans     | i          | x          | BenA          | Frisvad et al. (2004)                         |
|             | A. roseoglobosus     | i          | x          | BenA          | Frisvad et al. (2004)                         |
|             | A. steynii           | i          | g x        | BenA          | Frisvad et al. (2004)                         |
|             | A. westerdijkiae     | i          | x          | BenA          | Frisvad et al. (2004)                         |
| Nigri       | A. costaricaensis    | i          | g x        | BenA          | Samson et al. (2004)                          |
|             | A. homomorphus       | i          | g x        | BenA          | Samson et al. (2004)                          |
|             | A. lactocoffeatus    | i          | g x        | BenA          | Samson et al. (2004)                          |
|             | A. piperis           | i          | g x        | BenA          | Samson et al. (2004)                          |
|             | A. scleroctioniger   | i          | g x        | BenA          | Samson et al. (2004)                          |
|             | A. vadensis          | i          | g x        | multi         | Vries et al. (2005)                           |
|             | A. ibericus          | i          | g x        | multi         | Serra et al. (2006)                           |
|             | A. brasiliensis      | i          | g x        | multi         | Varga et al. (2007)                           |
|             | A. uvarum            | i          | g x        | multi         | Pertone et al. (2007)                         |
|             | A. aculeatinus       | i          | g x        | multi         | Noonim et al. (2007)                          |
|             | A. scleroticonarbonarius | i          | g x        | multi         | Noonim et al. (2007)                          |
| Flavi       | A. bombysis          | i          | g x        | multi         | Peterson et al. (2001)                        |
|             | A. pseudotamarii     | i          | g x        | multi         | Ito et al. (2001)                             |
|             | A. arachidicola      | i          | g x        | multi         | Pildain et al. (2007)                         |
|             | A. miniscerotigenes  | i          | g x        | multi         | Pildain et al. (2007)                         |
|             | A. pariscerotigenus  | i          | g x        | BenA          | Frisvad et al. (2005)                         |
| Ochraceorosei| A. rambellii        | i          | g x        | ITS, RAPD     | Frisvad et al. (2005)                         |
| Fumigati    | A. lentulus          | i          | multi      |                | Balajee et al. (2005)                         |
|             | A. fumigatiatifinis  | i          | x          | multi         | Hong et al. (2005)                            |
|             | A. novofumigatus     | i          | x          | multi         | Hong et al. (2005)                            |
|             | A. turcosus          | i          | g x        | multi         | Hong et al. (2007)                            |
| Neoantorya  | N. nishimurae        | a          | g          |                | Takada et al. (2001)                          |
|             | N. otanii            | a          | l          | g             | Takada et al. (2001)                          |
|             | N. takakii           | a          |            |                | Horie et al. (2001)                           |
|             | N. indohii           | a          |            |                | Horie et al. (2003)                           |
|             | N. tsurufae          | a          |            |                | Horie et al. (2003)                           |
|             | N. coreana           | a          |            | BenA/Calm      | Hong et al. (2006)                            |
|             | N. laciniosa         | a          |            | BenA/Calm      | Hong et al. (2006)                            |
|             | N. assulata          | a          | i          | g x           | multi                                        | Hong et al. (2007) |
|             | N. denticulata       | a          | i          | g x           | multi                                        | Hong et al. (2007) |
|             | N. galapagensis      | a          | i          | g x           | multi                                        | Hong et al. (2007) |
|             | N. australensis      | a          | i          | g x           | multi                                        | Samson et al. (2007) |
|             | N. ferenczii         | a          | i          | g x           | multi                                        | Samson et al. (2007) |
|             | N. papuensis         | a          | i          | g x           | multi                                        | Samson et al. (2007) |
|             | N. warcupii          | a          | i          | g x           | multi                                        | Samson et al. (2007) |
| Emericella  | Em. qinqixianii      | a          |            |                | Horie et al. (2000)                           |
|             | Em. venezuelensis    | a          | l          | BenA          | Frisvad & Samson (2004)                       |
| Eurotium    | Eu. taiilimakanense  | a          | i          | g             | Ablt & Horie (2001)                           |

Abbreviations: a = ascospore characters; l = anamorph characters; g = Growth rate; x = Extrolites; BenA = β-tubulin; Calm = calmodulin; multi = three or more molecular probes.
There is no one method (morphological, physiological or molecular) that works flawlessly in recognising species. Perhaps this is why *Aspergillus* taxonomists have so readily embraced the eclectic approach. Morphological characters can vary. For instance, sclerotia which are "characteristic" of some species are not always present in all isolates of a species, and their production can vary among cultures of the same isolate. Currently we are beginning to question the past emphasis on ascospore wall characteristics in systematics. A case in point is the *Emericella nidulans* variants that possess rough-walled ascospores normally characteristic of *Em. rugulosa* (Klich et al. 2001). Physiological characters may vary, or in the case of metabolites, be absent altogether in some isolates. Multilocus DNA sequence data are extremely useful for recognising species boundaries, but we do not have strict criteria as to where to draw the line between phylogenetic species and well-differentiated species boundaries, but we do not have strict criteria as to where to draw the line between phylogenetic species and well-differentiated populations that are potentially capable of interbreeding, at the same time integrating them into a coherent species concept. Therefore, we advocate the use of morphological, physiological and molecular data in circumscribing *Aspergillus* species wherever possible.

Although the majority of fungal systematists now utilise molecular data to some extent, researchers give varying levels of weight to molecular versus other sources of information in making taxonomic decisions. Indeed, *Aspergillus* systematists are no exception, and a rather large diversity of opinions on this matter exists even among the authors of this paper! Rather than dwelling on these disagreements, we emphasise the considerable common ground among us, which is leading to a taxonomic system that reflects morphological, physiological, ecological and phylogenetic relationships. Dobzhansky (1951) noted that the discontinuous distribution of variation among organisms is a matter of observation consistent with species being real entities, and that discontinuity is as evident in the genus *Aspergillus* as it is anywhere. While there are inherent problems of logic with placing these discontinuous units into species using consistent rules (Hey 2001) we have to do the best job we can, and the approaches employed over recent decades appear to have served us well. We advocate the use of as much data from as many sources as possible in making taxonomic decisions (Samson et al. 2006).

The roles of molecular, morphological and physiological characters in species recognition

It is important to consider carefully the roles of different kinds of characters in drawing boundaries between species. Variable DNA sequence characters provide the best means for inferring relationships among organisms, simply because it is possible to sample very large numbers of variable characters, and for the most part, those characters vary because they are under little or very weak selection. Within and between sister species, the genealogies inferred from single genes do not reflect the underlying organisinal genealogies, because of the stochastic effects of the segregation of variation during speciation, and recombination. However, the point where different gene genealogies become concordant is a useful place to assign a species boundary, as it is likely to reflect a lack of historical recombination among species (Dykhuizen & Green 1991; Koufopanou et al. 1997; Geiser et al. 2000; Taylor et al. 2000). In practice, this has proven to be a powerful tool in the fungi and in the genus *Aspergillus*, as discussed below.

An inference of a species boundary based on multilocus data provides an objectively determined line of demarcation, but as a rule in fungi, it appears that reproductive isolation precedes morphological differentiation in speciation (Taylor et al. 2006). This can lead to some unsettling observations. On the one hand, it is comforting that there is usually a very strong correlation between biological species in fungi defined based on laboratory mating tests and those defined based on genealogical concordance principles (O'Donnell et al. 1998; O'Donnell 2000). In some cases, closely related phylogenetic species show some degree of cross-fertility (Dettman et al. 2006), allowing a window into the forces of hybridisation that affect fungal evolution as much as they do animals and plants, yet have gone mostly unobserved because fungal morphological species concepts tend to be too broad to allow hybridisation to be considered (O’Donnell et al. 2004). What is less satisfying about the genealogical concordance approach is that it tends to yield species that are morphologically and otherwise biologically cryptic (Koufopanou et al. 1997; Geiser et al. 1998; O’Donnell et al. 2004; Pringle et al. 2005; Balajee et al. 2005), in some cases leaving authors with little choice but to describe species using nucleotide characters as the primary descriptors (Fisher et al. 2002; O’Donnell et al. 2004). However, this may be a simple matter of not looking hard enough. Indeed, initial observations of the morphology and physiology of the cryptic human pathogenic species *Coccidioides immitis* and *C. posadasii* uncovered no morphological or physiological differences other than growth rates on media with high salt concentration (Fisher et al. 2002), but molecular evolutionary studies of a proline-rich antigen in these species showed different patterns of positive selection that need to be considered in vaccine design (Johannesson et al. 2004). While these species may be cryptic to the human eye, perhaps they can be distinguished by the human immune system, and we are served well by recognising them as separate species.

Most *Aspergillus* species were defined based on morphology, with the additional consideration of molecular and extrolite data used in recent years (Hong et al. 2005). Molecular phylogenetics has uncovered cryptic speciation in a number of taxa (Geiser et al. 1998; Pringle et al. 2005; Balajee et al. 2005), suggesting that morphological characters provide a very broad species concept that does not reflect the true extent of evolutionary divergence and reproductive isolation, as appears to be the rule in fungi (Geiser 2004). However, physiological characters, including growth rates and the production of extrolites, often show differences that reflect phylogenetic species boundaries (Geiser et al. 2000). Considering this, greater emphasis should be placed on extrolite profiles and growth characteristics in species descriptions.

Molecular characters provide the greatest number of variable characters for fungal taxonomy, they can be generated using a widely available technology, that technology comes with an extremely well-developed bioinformatic infrastructure that allows worldwide communication and comparison of results, and they produce results that generally correlate with reproductive barriers and physiological differences. This utility ensures that molecular characters will have a primary role in recognising fungal taxa. However, good taxonomy does not end with the recognition of a species and a Latin binomial. Species descriptions should include data from as many sources as possible, comprising morphology, physiology and molecular data, which can be used not only as tools for identifying an isolate, but understanding its biology.
Some insights on *Aspergillus* species recognition based on multilocus phylogenetics

In recent years, molecular tools such as RFLP's, RAPD's, AFLP, MLTEE, ribosomal RNA sequences, and protein-coding gene sequences have been applied to taxonomic questions in the genus. Multilocus DNA sequence studies of some anamorphic species have shown that the patterns of polymorphisms in different genes are consistent with recombination in these asexual species (Geiser et al. 1998). Clonal lineages accumulate large numbers of deleterious mutations over time but perhaps aspergilli survive because they are not as clonal as we once thought. The discovery of MAT idiomorphs in the complete genome sequence of *A. fumigatus* and the subsequent discovery of both MAT idiomorphs in populations of *A. fumigatus* (Paoletti et al. 2005; Dyer & Paoletti 2005; Nierman et al. 2006; Rydholm et al. 2007) strongly indicate that this species is heterothallic, consistent with population genetic patterns suggesting some level of recombination (Pringle et al. 2005). These indications that putatively asexual *Aspergillus* species are actually recombining allow the use of genealogical concordance methods for delimiting species (Taylor et al. 2000).

Single locus DNA sequence studies have been conducted in *Aspergillus* using different loci, and there are extensive databases available for nuclear ribosomal RNA genes (large subunit, internal transcribed spacers) and β-tubulin. Any of these gene regions alone may serve as an effective tool for identifying well-defined species, but a weakness of the single locus approach is that not all species can be identified from DNA polymorphisms therein.

In order to resolve the central question of species boundaries in *Aspergillus* and detection of those species using methodologies based on DNA sequences, a four locus DNA sequence study was undertaken covering all major lineages in *Aspergillus* and have included most of the known and accepted species. Interpretation of the results using genealogical concordance theory results in a species recognition system that agrees in part with phenotypic studies and reveals the presence of many undescribed species not resolved by phenotype.

A primary question of this study is whether the *Aspergillus* anamorph is a reliable marker for species that belong in the *Aspergillus* clade. The species of Sclerocleista, *S. ornata* and *S. thaxteri*, are vastly different from most *Aspergillus* species and form a distinct group distantly related to the main *Aspergillus* clade. *A. zonatus*, *A. clavatoflavus* (section Flavi) and *Warcupiella spinosa* (section Ornati), like the Sclerocleista species are distant from the main body of species with *Aspergillus* anamorphs and should be placed in their own groups too. *Hemicarpenteles paradoxus*, *A. malodoratus* and *A. crystallinus* are phylogenetically outside of all other *Aspergillus* species (Stevenson, Samson & Varga, unpubl. data).

In examining close relationships, it was found that *Aspergillus niveus* ex type is not the same species as *Fennellia nivea* ex type; *Chaetosartorya stromatoides* ex type is not the same species as *Aspergillus stromatoides* ex type (Peterson 1995) and *Fennellia flavipes* ex type is not the same species as *A. flavipes* ex type. This raises a nomenclatural challenge, because multilocus DNA analysis showed that the assignments of anamorph-teleomorph connections were incorrect. Additionally, among isolates identified as *A. flavipes* there are at least three new species.

In the *Eurotium* lineage, a teleomorphic state was identified that belongs with *A. proliferans*. This species was placed in the *Eurotium* clade even though it was anamorphic (Raper & Fennell, 1965). The isolates identified as telemorphs of *A. proliferans* had previously been described by Raper & Fennell (1965) as a colonial variant of *A. ruber* (NRRL 71) or a transitional strain somewhere between *A. ruber* and *A. mangini* in taxonomic terms (NRRL 114). Because these isolates are conspecific with *A. proliferans* they should have a distinct *Eurotium* name and description. Separately, about 80 isolates of *A. restrictus* were sequenced at a single locus. Three of those isolates were identical with known *Eurotium* species at that locus. Subsequent sequencing of the other three loci conclusively showed that the anamorphs of *E. intermedium*, *E. repens* and *E. amastelodami* can be found in nature apart from their teleomorphic state.

*Aspergillus terreus* isolates contain a greater amount of intraspecific variation than in many other species, and additionally, among the many *A. terreus* isolates examined were several new species. In this species, protein coding genes seem to provide a good set of loci for strain typing. It remains to be seen whether the typing possible in this group with these loci will have relevance for medical treatments.

The internet provides a means for the rapid distribution of data, but not every site on the web contains data of equal quality. GenBank contains sequences of questionable quality and probable errors in the identity of source organisms. Proposals from the MLST community would create curated sites with a more narrow scope than GenBank. At such a site, researchers with a putative new sequence type submit their DNA sequence along with corroborating information, such as the tracings from the DNA sequence, to the curator of the site. The curator verifies the quality of the data and either asks for more information from the submitting scientist or posts the new sequence type to the web site. In this way, data that has gone through a third party quality check quickly appears at a central site accessible from all over the globe.

From species recognition to species identification

DNA barcoding is a taxonomic method which uses a short genetic marker in an organism’s DNA to quickly and easily identify it as belonging to a particular species. A DNA sequence should meet several criteria to be used successfully for species identification. DNA sequences should be orthologous in the examined organisms, and variable enough to allow species identification, with low levels of intraspecific variation (Hebert et al. 2003). A DNA barcode should be easily accessible (universally amplified/sequenced by standardised primers from a wide set of organisms), relatively short (≤ ~ 500–600 bp), simple to sequence, easily alignable [although this problem can be overcome by using Composition Vector Tree analysis (Chu et al. 2006), or other non-alignment based algorithms developed recently (Little & Stevenson 2007)], with no recombination. An ~ 600 bp region of the 5' end of the mitochondrial cytochrome oxidase subunit 1 (*cox1*, usually referred to as CO1 in barcoding studies) was proposed to be a good candidate for barcoding animal species including birds (Hebert et al. 2004b), fishes (Ward et al. 2005), and Lepidopteran insects (Hebert et al. 2004a; Hajibabaei et al. 2006). The *cox1* region was also used successfully to develop DNA barcodes for red algae (Saunders 2005). However, recent studies have indicated that mtDNA based barcoding region on its own is not suitable for species identification in several cases. Factors such as interspecific hybridisation (Hurst & Jiggins 2005; Bachtrg et al. 2006), presence of mtDNA derived genes in the nuclear genome (Thalmann et al. 2004; Bensasson et al. 2001), and infection by maternally transmitted endosymbionts
such as *Wolbachia* are known to cause flow of mitochondrial genes between biological species, so species groupings created using mtDNA can differ from the true species groupings (Hurst & Jiggins 2005; Withworth et al. 2007). Intra-individual heterogeneity of mitochondrial sequences can also cause problems (Sword et al. 2007). Besides, data presented by Wiemers & Fiedler (2007) indicate that the “barcoding gap” (i.e. the difference between inter- and intraspecific variability) is an artifact of insufficient sampling, and suggested that other characters should be examined to identify new species. Hickerson et al. (2006) claimed that single-locus mtDNA-based barcodes can only consistently discover new species if their populations have been isolated for more than 4 million generations. Other authors also suggested that mtDNA-based barcodes should be supplemented with nuclear barcodes (Moritz & Cicero 2004; Daschamapata & Mallet 2006). Attempts have been made to evaluate various regions of the nuclear ribosomal RNA gene cluster for barcoding animals (Sonnenberg et al. 2007; Schill & Steinbrück 2007; Vences et al. 2006a).

For some animal groups, other (usually nuclear) genomic regions have been proposed to be used for species identification. In nematodes, a small region of the 18S rRNA gene was proposed as a DNA barcode (Bhadury et al. 2005). For dinoflagellates and sponges the intergenic transcribed spacer (ITS) region was found to be the most promising for species identification (Park et al. 2007; Litaker et al. 2007), while for Cephalopoda rRNA sequences were suggested to be used as DNA barcodes (Strugnell & Lindgren 2007). Regarding plants, the *cox1* region was found to be inappropriate for species identification because of a much slower rate of evolution of *cox1* (and other mitochondrial genes) in higher plants than in animals (Kress et al. 2005; Chase et al. 2005). A variety of loci have been suggested as DNA barcodes for plants, including coding genes and non-coding spacers in the nuclear and plastid genomes. In flowering plants, the nuclear ITS region and the plastid *tmH-psbA* intergenic spacer are two of the leading candidates (Kress et al. 2005). More recently, the non-coding *tmH-psbA* spacer region coupled with the coding *rbcL* gene have been suggested as a two-locus global barcode for land plants (Kress & Erickson 2007). In protists, Scicluna et al. (2006) used successfully part of the ssu rRNA gene for species identification, while spliced leader RNA sequences have been used as barcodes in Trypanosomatidae (Maslov et al. 2007).

Regarding fungi, ITS and translation elongation factor 1-alpha (*tef1*) based DNA barcodes have been developed recently for identification of *Trichoderma* and *Hypocrea* species (Druzhinina et al. 2005). The ITS region was also found to be useful for species identification in other fungal groups including Zygomycetes (Schwarz et al. 2006), dematiaceous fungi (Desnos-Ollivier et al. 2006) and *Trichophyton* species (Summerbell et al. 2007). For the identification of *Fusarium* species, Geiser et al. (2004) developed *FUSAR. RIUM-ID* v.1.0, a publicly available sequence database of partial *tef1* sequences. Recently, the applicability of the *cox1* region for species identification in *Penicillium* subgenus *Penicillium* has been examined by Seifert et al. (2007). They found that representatives from 38 of 58 species could be distinguished from each other using *cox1* sequences. Work is in progress in several laboratories to develop DNA barcodes for various fungal groups.

In this study, we wished to evaluate the usefulness of *cox1* for species identification in the *Aspergillus* genus. Our studies were focused on *Aspergillus* section *Nigri*. Our previous work on mitochondrial genetics of aspergilli indicated that the *cox1* gene, or more generally any genes located on the mtDNA of aspergilli does not meet all criteria needed for a DNA barcode. Regarding intraspecific variability, several studies found high levels of intraspecific variability among black aspergilli (Varga et al. 1993, 1994; Hamari et al. 1997; Kevei et al. 1996), which is manifested not only in the presence or absence of intronic sequences, but also in exonic regions (Hamari et al. 2001; Juhász et al. 2007). Both inter- and intraspecific recombination has been detected in aspergilli even under non-selective conditions (Earl et al. 1981; Tóth et al. 1998; Kevei et al. 1997; Hamari et al. 2003; Juhász et al. 2004). Additionally, the *cox1* gene of *Aspergillus* species examined so far carries numerous introns which could make further work cumbersome (Hamari et al. 2003). Several studies also indicated that phylogenetic trees based on mitochondrial and nuclear sequence data are incongruent (Geiser et al. 1996; Wang et al. 2000). Although DNA barcodes are not meant to be used for phylogenetic analyses, the standard short barcode sequences (ca. 600 bp) were found to be unsuitable for inferring accurate phylogenetic relationships among fungi (Min & Hickey 2007).

The authors proposed to extend the barcoding region from 600 to 1200 bp to be able to distinguish between these closely related species; however, longer sequences are impractical for use in high-throughput screening programs.

During this study, we concentrated on the *A. niger* species complex, which includes eight species: *A. niger*, *A. tubingensis*, *A. foetidus*, *A. piperis*, *A. brasiliensis*, *A. vadensis*, *A. costaricensis* and *A. lacticoffeatus* (Samson et al. 2004). According to our previous studies, all these species can be distinguished from each other using calmodulin sequence data, and all except one could be distinguished using *β*-tubulin sequence data (A. lacticoffeatus had identical *β*-tubulin sequences to some *A. niger* isolates; Samson et al. 2004; Varga et al. 2007). The ITS data set delimited 4 groups: 1. *A. niger* and *A. lacticoffeatus* isolates; 2. *A. brasiliensis*; 3. *A. costaricensis*; 4. *A. tubingensis*, *A. foetidus*, *A. vadensis* and *A. piperis* (Varga et al. 2007). We also examined the applicability of the IGS (intergenic spacer region) for species identification; our data indicate that this region exhibits too high intraspecific variability to be useful for DNA barcoding. Other genomic regions examined by other research groups could also distinguish at least 2–5 species in the *A. niger* species complex, including pyruvate kinase, pectin lyase, polygalacturonase, arabinoylanarabinofuranohydrolase and several other genes (Gielkens et al. 1997; de Vries et al. 2005; Parenicova et al. 2001), translation initiation factor 2, pyruvate carboxylase, 70 kD heat shock protein, chaperonin complex component (TCP-1), ATPase (Witiak et al. 2007), and translation elongation factor 1-α, RNA polymerase II and actin gene sequences (S.W. Peterson, pers. comm.).

To evaluate the usefulness of the *cox1* gene for DNA barcoding, we amplified and sequenced part of the *cox1* gene from about 45 isolates of the *A. niger* species complex using the primer pairs developed by Seifert et al. (2007), and examined their properties for species delimitation. Sequence alignments and phylogenetic analysis were made by the MEGA 3 software package (Kumar et al. 2004). For parsimony analysis, the PAUP v. 4.0 software was used (Swofford 2003). Alignment gaps were treated as a fifth character, chaperonin complex component (TCP-1), ATPase (Witiak et al. 2007), and translation elongation factor 1-α, RNA polymerase II and actin gene sequences (S.W. Peterson, pers. comm.).
Fig. 1. Neighbour-joining tree based on cox1 sequences of the A. niger aggregate. Aspergillus ellipticus CBS 482.65 was used as outgroup. Bootstrap values above 50% are shown on the branches. Partial cox1 sequences of the examined isolates have been deposited in the GenBank database under accession numbers EU021012–EU021046.

which was used as outgroup, 8 of the 12 variable characters were parsimony informative). The topology of the Neighbour-joining tree shown in Fig. 1 is the same as one of the 76 maximum parsimony trees (length: 56 steps, CI = 0.9038, RI = 0.8936, RC = 0.8077). The intraspecific genetic distances calculated using Kimura’s 2-parameter model (Kimura 1980) were 0.003 both for A. niger and A. tubingensis, 0.007 for A. brasiliensis, and 0.002 for A. foetidus. The interspecific diversities were in the same range: 0.003 between A. niger and A. tubingensis or A. foetidus, 0.004 between A. tubingensis and A. foetidus, and 0.005–0.006 between A. brasiliensis and the other species examined. The genetic distance from A. ellipticus varied between 0.084–0.090 for the different species of the A. niger aggregate. Although most isolates of either A. tubingensis or A. niger were found to have identical cox1 sequences, several isolates of these species did not fit into these clades. We could not identify a single site which could be used without ambiguity for distinguishing A. niger from A. tubingensis using a character-based approach as suggested by DeSalle et al. (2005). Although at position 328 all A. tubingensis isolates contained an A nucleotide and most A. niger isolates contained C, some A. niger isolates also had A in this position (data not shown). This nonsynonymous substitution resulted in an amino acid change from leucine to isoleucine. Regarding other amino acid sequences, altogether 5 variable sites have been found, 3 of which
were parsimony informative. All these amino acid substitutions represented intraspecific variability within the species (data not shown).

Our data indicate that cox1 is not appropriate to be used as DNA barcode in black aspergilli since none of the eight species of the A. niger species complex could be identified unequivocally (Fig. 1). The cox1 gene sequences of the A. niger species complex do not obey either of the two most frequently used methods for species delimitation, reciprocal monophyly as proposed by Wiens & Penkroft (2002), or the “10× rule”, the observation of 10 times greater average pairwise genetic difference between the n1 individuals of the candidate species and the reference species than the average within-species pairwise differences found in the particular taxonomic group (Hebert et al. 2004). The phylogenetic tree constructed based on the cox1 sequences shows an overlap between intra- and interspecific variation possibly due to past mitochondrial DNA recombination events as suggested earlier (Tóth et al. 1998). Although a high degree of heterokaryon incompatibility was observed among isolates of the A. niger species complex (van Diepeningen et al. 1997), mtDNA transfers occur readily even between incompatible isolates (Tóth et al. 1998). Our data are in agreement with those of Min & Hickey (2007), who recently analyzed the cox1 sequences of A. niger and A. tubingensis isolates available from the GenBank database. The intra- and interspecific genetic diversity was found to be in the same range, and the amino acid sequences of the cox1 barcode region of isolates of the two species were the same. In the barcoding region, 4 variant sites have been found, of which 2 represent intraspecific variability within A. niger, one substitution was present in an A. niger and an A. tubingensis isolate, and one substitution represented interspecific variability between A. niger and A. tubingensis. This very low level of interspecific variability compared to the similar level of intraspecific diversity in our view is unsatisfactory for a DNA barcode to be used successfully for species identification.

Recent attempts to use the cox1 gene for species identification in other fungal groups including Fusarium species and basidiomycetes have also met with limited success (K. Seifert, pers. comm.). Comparing the phylogenies based on cox1, ITS, β-tubulin and calmodulin sequences, either β-tubulin and calmodulin could serve as a suitable region for species identification among black aspergilli. Recently, participants of the All Fungi DNA Barcoding Planning Workshop (Smithsonian Conservation and Research Center, Front Royal, Virginia, 13–15 May, 2007) accepted the ITS region as the first choice for DNA barcoding of the Fungal Kingdom (www.allfungi.org). However, ITS does not always resolve very closely related phylogenetic species (Bruns 2001), whereas intron-rich protein coding genes generally do much better (Geiser 2004). In Fusarium, partial translation elongation factor 1-alpha sequences have emerged as the most useful single-locus identification tool (Geiser et al. 2004). In the case of aspergilli, our opinion is to use either β-tubulin or calmodulin sequences for accurate species identification because of their prevalence in public databases universality of application, and relative resolving power. However, studies in progress indicate that other regions of the genome might serve better both for species identification and for phylogenetic studies (V. Roberts, pers. comm.). Further studies are in progress to evaluate the usefulness of cox1 for DNA barcoding in another Aspergillus section, and work is in progress to examine the applicability of other loci for species identification.

Looking ahead: genomics, many-locus phylogenies, and multilocus sequence typing

With the complete genome sequences of eight Aspergillus species (Em. nidulans, A. oryzae, A. fumigatus, N. fischeri, A. terreus, A. clavatus, A. niger, A. flavus) in advanced states of release, opportunities are upon us to generate data from far more loci in characterising Aspergillus species. One application would be to produce a many-locus phylogeny, with the goal of having a firm understanding of the evolutionary relationships among species across the entire genus. Previous phylogenetic studies across the genus have utilised one or a few markers, and unsurprisingly, do not resolve backbone relationships among sections and subgenera (Berbee et al. 1995; Geiser et al. 1996; Ogawa & Sugiyama, 2000; Peterson, 2000). Rokas et al. (2003) showed that strongly inferred nodes could be inferred consistently in the genus Saccharomyces, when approximately 20 genes were used, or about 8 000 orthologous nucleotide sites. With the goal of producing a well-supported phylogeny for the genus, we are using the available complete genome sequences to design new sequence markers that can be applied across the genus. Primers were designed to amplify 500–1 300 bp regions of protein coding genes, using conserved stretches of amino acids as primer sites. Our goal is to generate data from up to 20 genes across the entire genus. Initial results based on ten genes and 36 taxa indicate that backbone node support remains elusive across the genus, but excellent bootstrap support is generally obtained for relationships among species within sections and subgenera (Witiak & Geiser, unpubl.). Poor backbone support may be due to the overwhelming amount of sequence diversity in the genus, leading to saturation of third codon positions and long branch attraction artifacts. These problems may be averted by selecting genes exhibiting appropriate levels of resolving power (Townsend 2007) or perhaps with full taxon sampling breaking up long branches. Regardless of any inability to resolve early evolutionary events in the evolution of the genus, these data will provide the basis for multilocus sequence typing schemes that will allow for the precise identification of unknowns and the discovery of new taxa. With advances in the utility of DNA chip technologies accompanied by lower cost, we should be looking toward the application of an affordable and accurate multilocus barcoding chip for the genus Aspergillus in the foreseeable future.

REFERENCES

Abitz P, Horie Y, Hui Y, Nishimura K, Li R (2001). New and interesting species of Eurotium from Chinese soil. Mycoscience 42: 289–294.
Bachthor D, Thornton K, Clark A, Andolfatto P (2006). Extensive introgression of mitochondrial DNA relative to nuclear genes in the Drosophila yakuba species group. Evolution 60: 292–302.
Balajee SA, Gribskov J, Hanley E, Nickle D, Marr KA (2005). Aspergillus lentulus sp. nov., a new sibling species of A. fumigatus. Eukaryotic Cell 4: 625–632.
Benassio D, Zhang DX, Hart DL, Hewitt GM (2001). Mitochondrial pseudogene: evolution’s misplaced witnesses. Trends in Ecology and Evolution 16: 314–321.
Berbee ML, Yoshimura A, Sugiyama J, Taylor JW (1995). Is Penicillium monophyletic? An evaluation of phylogeny in the family Trichocomaceae from 18S, 5.8S and ITS ribosomal DNA sequence data. Mycologia 87: 210–222.
Bhutury P, Austen MC, Bilton DT, Lambohead PJD, Rogers AD, Smerdon GR (2008). Development and evaluation of a DNA-barcoding approach for the rapid identification of nematodes. Marine Ecology 320: 1–9.
Bruns TD (2001). ITS Reality. Inoculum 52: 2–3.
Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, Savolainen V (2005). Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transactions of the Royal Society of London B: Biological Sciences 360: 1869–1895.
Chu KH, Li CP, Qi J (2006). Ribosomal RNA as molecular barcodes: a simple correlation analysis without sequence alignment. Bioinformatics 22: 1690–1701.

Dasmahapatra KK, Mallet and Applied Microbiology 27: 442–453.

DeSalle R, Egan MG, Siddall M (2005). The unholy trinity, taxonomy, species delimitation and DNA barcoding. Proceedings of the Royal Society of London B: Biological Sciences 360: 1905–1916.

Desnos-Ollivier M, Lange L, Tkacz J, eds. Current Genetics 3: 221–225.

Dobzhansky TG (1951). Genetics and the Origin of Species. New York: Columbia University Press.

Druzhinin IS, Korphininsky AG, Komon M, Bissett J, Szakacs G, Kubicek CP (2005). An oligonucleotide barcode for species identification in Trichoderma and Hypocrea. Fungal Genetics and Biology 42: 813–828.

Dykhuizen DE, Green L (1991). Recombination in Escherichia coli and the definition of biological species. Journal of Bacteriology 173: 2757–2768.

Earl AJ, Turner G, Croft JH, Dales RBG, Lazarus CM, Lünsdorf H, Küntzel H (1981). High frequency transfer of species specific mitochondrial DNA sequences from members of the Aspergillaceae. Current Genetics 3: 221–225.

Frisvad JC, Frank JM, Houbraken JAMP, Kuijpers AFA, Samson RA (2004). New aflatoxin-producing fungus. In: Molecular and General Genetics 254: 379–388.

Frisvad JC, Samson RA (2004). Emericella venezuelensis, a new species with stellate ascospores producing sterigmatocystin and aflatoxin B,. Systematic and Applied Microbiology 27: 672–680.

Frisvad JC, Skouboe P and Samson RA (2005). Taxonomic comparison of three Chu KH, Li CP, Qi J (2006). Ribosomal RNA as molecular barcodes: a simple correlation analysis without sequence alignment. Bioinformatics 22: 1690–1701. Dasmahapatra KK, Mallet and Applied Microbiology 27: 442–453.

DeSalle R, Egan MG, Siddall M (2005). The unholy trinity, taxonomy, species delimitation and DNA barcoding. Proceedings of the Royal Society of London B: Biological Sciences 360: 1905–1916. Desnos-Ollivier M, Lange L, Tkacz J, eds. Current Genetics 3: 221–225.

Dobzhansky TG (1951). Genetics and the Origin of Species. New York: Columbia University Press. Druzhinin IS, Korphininsky AG, Komon M, Bissett J, Szakacs G, Kubicek CP (2005). An oligonucleotide barcode for species identification in Trichoderma and Hypocrea. Fungal Genetics and Biology 42: 813–828. Dykhuizen DE, Green L (1991). Recombination in Escherichia coli and the definition of biological species. Journal of Bacteriology 173: 2757–2768.

Earl AJ, Turner G, Croft JH, Dales RBG, Lazarus CM, Lünsdorf H, Küntzel H (1981). High frequency transfer of species specific mitochondrial DNA sequences from members of the Aspergillaceae. Current Genetics 3: 221–225. Fisher MO, Peckert M (2006). Reproduction in Aspergillus fumigatus: sexuality in a supposedly asexual species? Medical Mycology 43: S7–S14.

Dykhuizen DE, Green L (1991). Recombination in Escherichia coli and the definition of biological species. Journal of Bacteriology 173: 2757–2768.

Earl AJ, Turner G, Croft JH, Dales RBG, Lazarus CM, Lünsdorf H, Küntzel H (1981). High frequency transfer of species specific mitochondrial DNA sequences from members of the Aspergillaceae. Current Genetics 3: 221–225. Fisher MO, Peckert M (2006). Reproduction in Aspergillus fumigatus: sexuality in a supposedly asexual species? Medical Mycology 43: S7–S14. Dykhuizen DE. (1991). Recombination in Escherichia coli and the definition of biological species. Journal of Bacteriology 173: 2757–2768.

Earl AJ, Turner G, Croft JH, Dales RBG, Lazarus CM, Lünsdorf H, Küntzel H (1981). High frequency transfer of species specific mitochondrial DNA sequences from members of the Aspergillaceae. Current Genetics 3: 221–225. Fisher MO, Peckert M (2006). Reproduction in Aspergillus fumigatus: sexuality in a supposedly asexual species? Medical Mycology 43: S7–S14. Dykhuizen DE. (1991). Recombination in Escherichia coli and the definition of biological species. Journal of Bacteriology 173: 2757–2768. Earl AJ, Turner G, Croft JH, Dales RBG, Lazarus CM, Lünsdorf H, Küntzel H (1981). High frequency transfer of species specific mitochondrial DNA sequences from members of the Aspergillaceae. Current Genetics 3: 221–225. Fisher MO, Peckert M (2006). Reproduction in Aspergillus fumigatus: sexuality in a supposedly asexual species? Medical Mycology 43: S7–S14.
Varga J, Kevei F, Fekete C, Coenen A, Kozakiewicz Z, Croft JH (1993). Restriction fragment length polymorphisms in the mitochondrial DNAs of the Aspergillus niger aggregate. Mycological Research 97: 1207–1212.

Varga J, Kevei F, Vriesema A, Debets F, Kozakiewicz Z, Croft JH (1994). Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the Aspergillus niger aggregate. Canadian Journal of Microbiology 40: 612–621.

Varga J, Kocsubé S, Tóth B, Frisvad JC, Perrone G, Susca A, Meijer M, Samson RA (2007). Aspergillus brasiliensis sp. nov., a biseriate black Aspergillus species with worldwide distribution. International Journal of Systematic and Evolutionary Microbiology 57: 1925–1932.

Vences M, Thomas M, Bonett RM, Vlieg DR (2005). Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philosophical Transactions of the Royal Society B: Biological Sciences 360: 1859–1868.

Wang L, Yokoyama K, Miyaji M, Nishimura K (2000). Mitochondrial cytochrome b gene analysis of Aspergillus fumigatus and related species. Journal of Clinical Microbiology 38: 1352–1358.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005). DNA barcoding Australia’s fish species. Proceedings of the Royal Society of London B: Biological Sciences 360: 1847–1857.

Whitworth TL, Dawson RD, Magallon H, Baudry E (2007). DNA barcoding cannot reliably identify species of the blowfly genus Protocalliphora (Diptera: Calliphoridae). Proceedings of the Royal Society of London B: Biological Sciences 274: 1731–1739.

Vares M, Fiedler K (2007). Does the DNA barcoding gap exist? Frontiers in Zoology 4: 8.

Wiens JJ, Penkroft TA (2002). Delimiting species limits in spiny lizards (Sceloporus). Systematic Biology 51: 69–91.

Witiak SM, Samson RA, Varga J, Rokas A, Geiser DM (2007). Phylogenetic markers for the genus Aspergillus developed from complete genome sequences. 24th Fungal Genetics Conference, Asilomar, Abstract No. 130.

Zotti M, Corti AM (2002). Aspergillus persii: A new species in section Circumdati. Mycotaxon 83: 269–278.