

Molecular Identification of Unusual Pathogenic Yeast Isolates by Large Ribosomal Subunit Gene Sequencing: 2 Years of Experience at the United Kingdom Mycology Reference Laboratory[∇]

Christopher J. Linton,^{1*} Andrew M. Borman,¹ Grace Cheung,² Ann D. Holmes,¹ Adrien Szekely,¹ Michael D. Palmer,¹ Paul D. Bridge,³ Colin K. Campbell,¹ and Elizabeth M. Johnson¹

Mycology Reference Laboratory, Health Protection Agency, Bristol,¹ Department of Pathology and Microbiology, University of Bristol, Bristol,² and British Antarctic Survey, Cambridge,³ United Kingdom

Received 6 October 2006/Returned for modification 7 November 2006/Accepted 13 January 2007

Rapid identification of yeast isolates from clinical samples is particularly important given their innately variable antifungal susceptibility profiles. We present here an analysis of the utility of PCR amplification and sequence analysis of the hypervariable D1/D2 region of the 26S rRNA gene for the identification of yeast species submitted to the United Kingdom Mycology Reference Laboratory over a 2-year period. A total of 3,033 clinical isolates were received from 2004 to 2006 encompassing 50 different yeast species. While more than 90% of the isolates, corresponding to the most common *Candida* species, could be identified by using the AUXACOLOR2 yeast identification kit, 153 isolates (5%), comprised of 47 species, could not be identified by using this system and were subjected to molecular identification via 26S rRNA gene sequencing. These isolates included some common species that exhibited atypical biochemical and phenotypic profiles and also many rarer yeast species that are infrequently encountered in the clinical setting. All 47 species requiring molecular identification were unambiguously identified on the basis of D1/D2 sequences, and the molecular identities correlated well with the observed biochemical profiles of the various organisms. Together, our data underscore the utility of molecular techniques as a reference adjunct to conventional methods of yeast identification. Further, we show that PCR amplification and sequencing of the D1/D2 region reliably identifies more than 45 species of clinically significant yeasts and can also potentially identify new pathogenic yeast species.

Invasive fungal infections, especially those caused by *Candida* spp., remain a significant cause of mortality in immunocompromised patients and in those undergoing invasive procedures (reviewed in references 27 and 31). Although *Candida albicans* remains the principal agent of nosocomial infections, more than 150 yeast spp. have now been associated with human pathologies (17), at least partly because human pathogenic yeasts are part of the human commensal flora. The patterns of antifungal susceptibilities of *Candida* spp. vary almost as greatly as the organisms themselves. In addition, the widespread use of antifungal agents may have contributed to a shift in species distributions via the emergence of inherently resistant species as significant pathogens (21, 22). Thus, informed therapeutic decisions increasingly require correct and rapid identification of an ever-increasing number of potential pathogens.

Conventional identification of pathogenic yeasts in the clinical laboratory relies upon a combination of morphological and biochemical features using tests that are labor-intensive and time-consuming. In addition, the currently available commercial yeast identification systems frequently fail to identify the less common pathogens (6) or to discriminate between closely related species (32). Molecular methods for the identification potentially represent rapid and sensitive alternatives to con-

ventional identification for yeasts. Among the approaches evaluated, randomly amplified polymorphic DNA analysis (5), TaqMan PCR (18), DNA microarrays (23), amplified DNA length polymorphisms (12, 13), and species-specific PCR primer combinations (24) have shown promise. Although these approaches could distinguish among the top 10 to 15 most common yeast species, they have not been rigorously applied to the more unusual isolates from human infections.

An alternative molecular diagnostic approach involves PCR amplification of genomic DNA, followed by sequencing of the resulting amplicons. Target genes for amplification have included the mitochondrial cytochrome *b* gene (33) and the regions coding for the nuclear rRNA genes (12, 13). Such regions evolve slowly, show high degrees of conservation among fungi, and are thus potentially suited for molecular identification and the establishment of phylogenetic relationships. Indeed, in limited analyses, sequencing of the internal transcribed regions (ITS) of the nuclear rRNA gene cassette proved sufficient to discriminate between some 40 species of clinically important yeasts (12, 13).

In the present study, we assessed the possibility of using PCR amplification and sequencing of the D1/D2 region of the large ribosomal subunit gene for the identification of clinically important species. All yeast isolates submitted to the United Kingdom National Mycology Reference Laboratory (MRL), Bristol, over a period of 24 months that failed to be definitively identified using commercial identification kits were subjected to PCR sequencing analysis targeting the D1/D2 large ribosomal subunit region. This approach, with more than 150 isolates, combined with conventional identification methods al-

* Corresponding author. Mailing address: Mycology Reference Laboratory, Health Protection Agency South-West Regional Laboratory, Myrtle Road, Bristol BS2 8EL, United Kingdom. Phone: 117 926 8683. Fax: 117 922 6611. E-mail: Chris.Linton@bris.ac.uk.

[∇] Published ahead of print on 24 January 2007.

TABLE 1. Yeast isolates identified at the United Kingdom MRL from March 2004 through March 2006^a

Organism ^b	Total	% Total	AUX (no. of isolates)	Code (no. of isolates)	API 20C (no. of isolates)	Molecular ID ^c (no. of isolates)	% Molecular ID
<i>Candida albicans</i>	1,192 (germ tube = 1,126)	39.3	56	7145207 (30) 7144207 (15) 7141207 (4), PCR ^d ; 7144007 (2); 7145205 (3); 7144204 (2)	3	7	0.6
<i>Candida glabrata</i>	659	21.7	649	1040004 (649)	4	6	0.9
<i>Candida parapsilosis</i>	430	14.2	415	7143105 (134); 7145105 (124); 7147105 (120); 7146105 (10); 7102105 (10); 7142105 (8)	12	3	0.7
<i>Candida tropicalis</i>	156	5.2	154	7147005 (154)	2		0
<i>Candida lusitanae</i>	99	3.3	45	7067005 (17); 7062005 (11); 7167005 (12); 7163005 (3); 7162005 (2)	35	19	19.2
<i>Candida krusei</i>	74	2.4	70	1000005 (70)		4	5.4
<i>Cryptococcus neoformans</i>	64	2.1	61	7577444 (15); 7557444 (10); 7576444 (6); 7157444 (5); 7572544 (4); 7577544 (4)	1	3 ^e	4.7
<i>Saccharomyces cerevisiae</i>	60	2.0	29	7540004 (7); 7140004 (6); 7500004 (6); 7040004 (5); 7540004 (1); 7542004 (1)	27	4	6.7
<i>Candida guilliermondii</i>	53	1.7	25	7567105 (25)	11	19*	35.8
<i>Malassezia pachydermatis</i>	36 (34 direct ID)	1.2	NA ^h	NA		2	5.5
<i>Candida dubliniensis</i>	33 (germ tube = 15)	1.1	30	7141207 (22), PCR; 7141007 (2); 7141205 (2); 7143207 (1); 7146607 (1); 7140207 (1)		3	9.1
<i>Trichosporon</i> sp.	23	0.8	17	7364325 (5); 7777325 (5); 7767325 (2); 3364325 (2); 7374325 (1); 7366325 (1)	2	4	17.4
<i>Rhodotorula</i> sp.	19	0.7	NA	NA	19		0
<i>Candida inconspicua</i>	17	0.6	15	1000004 (15)		5*	29.4
<i>Candida kefyri</i>	15	0.5	5	5700005 (5)	6	4	26.7
<i>Candida famata</i>	12	0.4	7	7567104 (7) ^f	2	3	25.0
<i>Candida</i> sp. nov.	11	0.4	0	NA		11	100
<i>Geotrichum</i> sp.	10	0.3	5	1000025 (4) 1104021 (1)	2	3	30.0
<i>Candida pelliculosa</i>	8	0.3	0	NA	6	2	25.0
<i>Candida fabianii</i>	5	0.2	0	NA		5 ^g	100
<i>Candida lipolytica</i>	5	0.2	2	1100005 (2)		3	60.0
<i>Candida blankii</i>	4	0.1	0	NA		4	100
<i>Candida utilis</i>	4	0.1	0	NA	3	1	25.0
<i>Cryptococcus albidus</i>	4	0.1	4	7776140 (2); 7676140 (1); 7066140 (1)			0
<i>Candida rugosa</i>	3	0.1	0	NA	2	1	33.3
<i>Candida norvegensis</i>	3	0.1	0	NA		3	100
<i>Candida pararugosa</i>	3	0.1	0	NA		3	100
<i>Candida catenulata</i>	3	0.1	0	NA		3	100
<i>Kloeckera</i> sp.	3	0.1	0	NA	1	2	66.7
<i>Saccharomyces elongasporus</i>	2		0	NA		2	100
<i>Candida zeylanoides</i>	2	0.3	0	NA		2	100
<i>Candida eremophila</i>	2		0	NA		2	100
<i>Candida lambica</i>	2		0	NA		2	100
Total	3,033				138	153	5.0

^a Figures are given for the total number of isolates received for each species, the percentage of total isolates represented by each species, and the number of isolates identified by germ tube (*C. albicans* and *C. dubliniensis* only), AUXACOLOR2 (AUX; with the most frequent AUXACOLOR2 codes obtained), API 20C, and D1D2 sequencing (molecular ID, with the percentage of isolates of each species that required sequence analysis) analyses.

^b The following 17 organisms (0.5% of all isolates) were single isolates whose identification required molecular approaches: *Candida ciferrii*, *Candida boidinii*, *Candida palmiophila*, *Candida freyschussii*, *Kazachstania pintolopesii*, *Candida magnoliae*, *Candida viswanathii*, *Metschnikowia pulcherrima*, *Pichia onychis*, *Sporobolomyces* sp., *Zygosaccharomyces lentus*, *Issatchenkia terricola*, *Candida haemulonii*, *Candida pseudointermedia*, *Candida pseudoglebosa*, *Sporopachydermia cereana*, and *Kazachstania bovina*. *Candida ciferrii*, *Candida boidinii*, and *Candida magnoliae* were identified as part of quality control programs and were not isolates from clinical cases. Sequencing did not reliably resolve the *Trichosporon*, *Kloeckera*, and *Geotrichum* isolates to the species level.

^c Proportions of the molecular identifications indicated with an asterisk were used to confirm AUXACOLOR2 presumptive identities.

^d Four isolates of *C. albicans* gave AUXACOLOR2 profiles consistent with *C. dubliniensis* but were confirmed as *C. albicans* by intron-specific PCR (see the text).

^e *C. neoformans* isolates were identified by sequencing of the ITS1 region.

^f The published AUXACOLOR2 code of 7567100 for *C. famata* is incorrect and should read 7567104.

^g *C. fabianii* and *C. veronae* were initially reported to be indistinguishable on the basis of D1D2 sequencing (13).

^h NA, not applicable.

TABLE 2. List of isolation sites, where known, for representative isolates of the less frequently encountered yeast species

Organism	Isolation site(s) ^a
<i>Candida</i> sp. nov.	Exit site swab, pelvic collection, pelvic abscess fluid, ascitic fluid , and mouth (three times)
<i>Candida fabianii</i>	Environmental, gastric lavage, blood culture (sepsis open abdomen), urine and central line tip, and ear swab
<i>Candida blankii</i>	Blood culture (twice) , sputum, and endocarditis (repeatedly isolated from valves)
<i>Candida catenulata</i>	Blood culture , urinary tract infection, and sputum
<i>Candida eremophila</i>	Abdominal drain (post-ileac surgery)
<i>Candida utilis</i>	Collection around kidney (human)
<i>Kloekera</i> sp.	Vaginal swab
<i>Saccharomyces elongasporus</i>	Blood culture
<i>Candida freyschussii</i>	Medicolegal genitourinary sample
<i>Candida lambica</i>	Yeast infection of skin not responding to treatment
<i>Candida palmioleophila</i>	Urine from baby with fungal urinary tract infection
<i>Candida pintolopesii</i>	Blood culture
<i>Candida rugosa</i>	Blood culture
<i>Issatchenkia terricola</i>	Ascitic fluid
<i>Kazachstania bovina</i>	Mouth
<i>Metschnikowia pulcherrima</i>	Blood culture
<i>Pichia onychis</i>	Sputum
<i>Sporobolomyces</i> sp.	Vitreous humor, endophthalmitis
<i>Zygosaccharomyces lentus</i>	Soft drink

^a Sites considered to be most likely to be clinically significant are indicated in boldface type.

lowed the identification of 50 different species of potentially pathogenic yeasts. Moreover, in all cases, molecular identifications correlated well with observed phenotypic characteristics of the isolates.

MATERIALS AND METHODS

Test isolates. A total of 3,033 clinical isolates that had been submitted to the MRL for identification were tested. Isolates were subcultured twice on plates of Oxoid Sabouraud dextrose agar containing 0.5% (wt/vol) chloramphenicol (Unipath Limited, Basingstoke, England). Cultures were incubated for 24 h at 35°C prior to testing. All procedures used coded isolates, and their correct identities were only revealed after full conventional and molecular analyses.

Conventional identification methods. The clinical isolates included in the present study had all failed to be identified by the conventional identification methods used by the MRL. These conventional methods were as follows. After initial germ tube testing, isolates were subjected to the AUXACOLOR2 identification kit (Sanofi Diagnostics Pasteur, Paris, France) exactly as described previously (11). Isolates were also cultured on Dalmau plates (Oxoid cornmeal agar, supplemented with 1% Tween 80, with a sterile coverslip over a single streak of the organism) to establish the additional morphological characteristics required for complete AUXACOLOR2 profiles. After the failure of this system to identify the isolates (i.e., the generation of AUXACOLOR2 profiles that did not match any profiles in the database), isolates were tested in the API 20C system (bioMérieux, Marcy-l'Étoile, France) exactly as described previously (15).

Molecular methods. Genomic DNA was prepared from yeast cultures after 2 days of incubation on Sabouraud agar using Whatman FTA filter paper technology exactly as described previously (8). Briefly, yeast aqueous suspensions (200 μ l) were applied directly to Whatman FTA microcards, which were then placed, open, in a Panasonic microwave (800 W) while still damp and subjected to two cycles of 30 s on full power, with a pause of at least 30 s between each cycle. Punches (2 mm in diameter) were then removed from dried FTA filters by using a Harris micropunch and washed and used to program PCRs. The washing conditions consisted of two washes for 1 min each with 100 μ l of Whatman FTA wash reagent, followed by two washes for 1 min each with 100 μ l of TE buffer (10 mM Tris-HCl [pH 7.5], 0.1 mM EDTA). Washed filters were then dried for 5 min at 55°C on a dry heat block, and PCR mixes were added directly to the washed and dried FTA filter punches.

Amplification of a region of the large subunit gene (LSU) was performed with the primers described previously (20). For PCR, reactions (100 μ l) were performed in the presence of 200 μ M concentrations of each deoxynucleoside triphosphate, 250 nM concentrations of the appropriate primers, 2 U of HotStarTaq polymerase (QIAGEN, Valencia, CA), and either 5 μ l of Magna-

Pure-extracted total genomic DNA or a single filter punch. After enzyme activation at 94°C for 15 min, reactions were subjected to 40 thermal cycles of 94°C for 15 s, 55°C for 15 s, and 72°C for 90 s on a GeneAmp PCR Systems 9700 thermocycler (Applied Biosystems, Foster City, CA). Amplification success was evaluated by electrophoresis of a fraction of total amplification products in 1.2% (wt/vol) agarose gels run for 45 min at 120 V in Tris-borate buffer. For sequencing of PCR products, the remainder of the PCRs were adjusted to final concentrations of 10% (wt/vol) PEG 8000 and 10 mM MgCl₂ and then centrifuged for 10 min at 12,000 rpm in a bench-top centrifuge. The resulting DNA pellets were washed in 75% ethanol, air dried, resuspended in sterile water, and subjected to automatic sequencing by using the commercial service available at the Advanced Biotechnology Centre (Imperial College, London, United Kingdom). For clinical isolates, organisms were formally identified by using BLAST searches against fungal sequences in existing DNA databases (2) and multiple sequence alignments (CLUSTAL W, version 1.82 [30]) using a database of formally identified organisms compiled by the MRL.

For principal component analysis (PCA), DNA sequences were aligned in CLUSTAL W (30) and manually edited by using JalView (14). A similarity matrix was obtained from the aligned sequences through the DNAdist program in the PHYLIP package (16). The similarity matrix was reformatted manually, and PCA was undertaken through the MSVP package (Kovachs Computing, Ltd.).

RESULTS

From March 2004 to March 2006, the United Kingdom MRL received 3,033 clinical isolates of yeast for identification. Of these, 2,742 (90.5%) were identified by using a combination of germ tube testing and the AUXACOLOR2 system. A further 138 (4.5%) required use of the API 20C kit in addition to, or instead of, AUXACOLOR2, and the remaining 153 (5.0%) could not be satisfactorily identified by either of these commercially available conventional identification methods (Table 1). These remaining 153 isolates were thus identified by sequence analysis of the LSU D1D2 region.

Results of biochemical identification of pathogenic yeast isolates. Identification of the 3,033 clinical isolates received from 2004 to 2006 indicated that these isolates encompassed a total of 50 different yeast species. *C. albicans* was by far the most prevalent organism, accounting for nearly 40% of all

isolates (Table 1). Indeed, *C. albicans*, together with *C. glabrata*, *C. parapsilosis*, and *C. tropicalis*, represented more than 80% of all isolates received. More than 95% of the *C. albicans* isolates were positive in the germ tube test and were not further analyzed. More than 95% of germ tube-negative strains of *C. albicans* and of the isolates of *C. glabrata*, *C. parapsilosis*, and *C. tropicalis* were successfully identified by using the AUXACOLOR2 system (Table 1), with the result that <1% of isolates of these common species required molecular identification. Interestingly, four isolates yielded AUXACOLOR2 profiles corresponding to *C. dubliniensis*, and correct identification as *C. albicans* required PCR amplification of a discriminatory type 1 intron region (9).

A second group of eight species of potentially pathogenic yeast (*C. lusitaniae*, *C. krusei*, *C. guilliermondii*, *C. dubliniensis*, *Cryptococcus neoformans*, *Saccharomyces cerevisiae*, *Malassezia pachydermatis*, and *Trichosporon* sp.) accounted for 15% of total organisms received. At least 50% of these isolates were successfully identified by AUXACOLOR2. However, >5% of these organisms could not be identified by using either AUXACOLOR2 or API 20C kits and thus required molecular identification (Table 1). This situation was exaggerated with *C. lusitaniae* and *C. guilliermondii*, partly due to the failure of these organisms to produce the pseudomycelia on Dalmau cultures (data not shown) required for correct AUXACOLOR2 profiles, with the result that 20% or more of all isolates required molecular identification.

A further 153 yeast isolates, encompassing another 38 species, comprised the remaining 5% of all organisms identified during the study period (Table 1). With the exception of *Rhodotorula* sp. and *Cryptococcus albidus*, which were correctly identified by API 20C and AUXACOLOR2 testing, respectively, the majority of these more unusual yeast species were only satisfactorily identified by using molecular approaches or by further lengthy biochemical and morphological analyses (Table 1). Some of these isolates are likely to have been clinically significant, as judged by their culture from deep, usually sterile, body sites (Table 2), highlighting the importance of their correct identification. Interestingly, 11 of these 153 isolates corresponded to a new *Candida* species (*Candida* sp. nov.; Table 1), closely related to but distinguishable biochemically and genetically from *C. glabrata* (unpublished data).

Molecular identification of yeast isolates. Application of aqueous suspensions of yeast isolates to Whatman FTA cards (see Materials and Methods) generated PCR-grade genomic DNA for all 153 isolates tested, as judged by amplification success in PCRs with primers targeting the D1/D2 region of the LSU (20; data not shown). Even though the D1/D2 region is a relatively hypervariable part of the more conserved LSU element within the nuclear ribosomal repeat region, a surprising species-specific heterogeneity in the PCR amplicon lengths was observed for this region. As judged by the results of automatic sequencing of the amplicons in each case, PCR products (excluding PCR primer sites) ranged from 252 bp (for *Geotrichum* sp.) to 415 bp (for *Sporobolomyces* sp.; Table 3). Despite this heterogeneity, several species still shared amplicons of identical lengths, indicating that D1/D2 region PCR fragment lengths alone are not appropriate for identification of *Candida* species. It should also be noted that three isolates of *C. neoformans* were identified after PCR amplification and sequenc-

TABLE 3. Length of D1D2 amplicons sequenced for each of the yeast species analyzed, in order of ascending size

Organism	Amplicon length ^a (bp)
<i>Geotrichum</i> sp.	252
<i>Candida catenulata</i>	276
<i>Candida magnoliae</i>	276
<i>Candida lipolytica</i>	320
<i>Candida lusitaniae</i>	328
<i>Candida pseudointermedia</i>	329
<i>Metschnikowia pulcherrima</i>	331
<i>Candida haemulonii</i>	332
<i>Candida rugosa</i>	345
<i>Saccharomyces elongasporus</i>	357
<i>Candida boidinii</i>	361
<i>Candida kefyr</i>	361
<i>Candida blankii</i>	363
<i>Candida pararugosa</i>	364
<i>Candida lambica</i>	365
<i>Candida pseudoglebosa</i>	365
<i>Candida dubliniensis</i>	366
<i>Candida parapsilosis</i>	366
<i>Candida viswanathii</i>	366
<i>Candida tropicalis</i>	367
<i>Candida zeylanoides</i>	367
<i>Candida albicans</i>	368
<i>Candida guilliermondii</i>	368
<i>Candida krusei</i>	368
<i>Candida ciferrii</i>	369
<i>Candida fabianii</i>	369
<i>Candida pelliculosa</i>	369
<i>Saccharomyces cerevisiae</i>	369
<i>Kloekera</i> sp	370
<i>Candida utilis</i>	374
<i>Candida glabrata</i>	378
<i>Candida norvegensis</i>	378
<i>Candida inconspicua</i>	381
<i>Candida palmiophila</i>	383
<i>Issatchenkia terricola</i>	383
<i>Candida famata</i>	387
<i>Candida pintolopesii</i>	388
<i>Kazachstania bovina</i>	388
<i>Sporopachydermia cereana</i>	388
<i>Pichia onychis</i>	389
<i>Trichosporon</i> sp.	390
<i>Candida eremophila</i>	391
<i>Zygosaccharomyces lentus</i>	397
<i>Candida freysshussii</i>	401
<i>Candida</i> sp. nov.	403
<i>Malassezia pachydermatis</i>	412
<i>Sporobolomyces</i> sp.	415

^a The lengths given exclude PCR primer sequences.

ing of the internal transcribed spacer region 1 (ITS1) using the primers described previously (7), rather than the D1/D2 region of the LSU. We routinely find PCR amplification of the D1/D2 region for this organism to be extremely inefficient (data not shown).

Automatic sequencing of the D1/D2 amplicons also revealed a dramatic degree of sequence variability among the isolates examined. The percent nucleotide identities among *Candida* species, calculated by using the CLUSTAL W multiple alignment program and ignoring gaps, varied from 96% identity (*C. palmiophila* and *C. pseudoglebosa*) to as little as 37% identity (*C. haemulonii* and *C. lipolytica*, *C. rugosa*, or *C. pararugosa*; data not shown). Thus, with the exception of 11

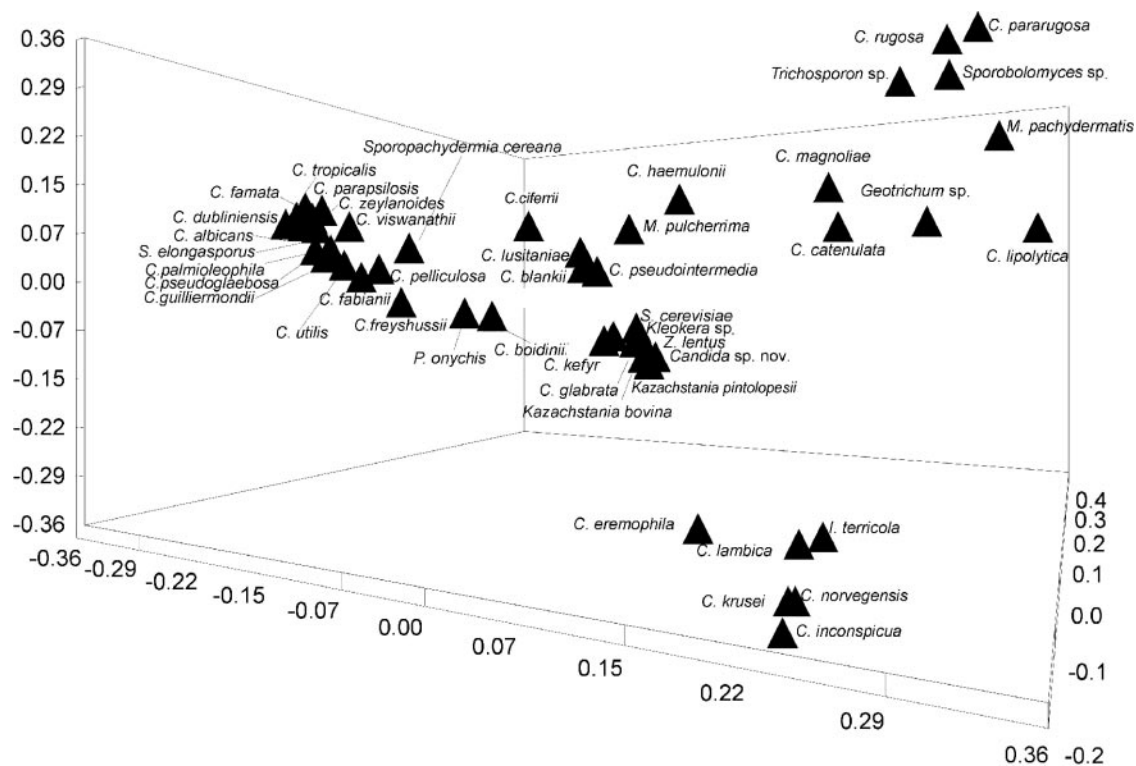


FIG. 1. Principle component ordination scores. The principle component ordination scores derived from analysis of 47 yeast sequences generated by PCR amplification and sequencing of the D1/D2 region are charted. The first three dimensions are shown.

isolates corresponding to a hitherto-undescribed species of *Candida* (*Candida* sp. nov; Table 1), BLASTN searches (2) of the resulting D1/D2 sequences against the existing synchronized EMBL and GenBank DNA databases yielded largely unequivocal hits against previously identified yeast reference isolates. To assess the reliability of molecular identification, the various sequences were subjected to PCA. The 47 DNA sequences were each recovered as distinct points in the PCA and could be represented as individual entries in the first three dimensions (Fig. 1). Total variance contained in the first three dimensions was rather low at 34%. Thus, it is extremely improbable that an unknown sequence would group incorrectly with a poorly related or unrelated reference sequence. In addition, for organisms not included in the databases of the AUXACOLOR2 and API 20C kits, the biochemical profiles obtained with AUXACOLOR2 and API 20C were compared to the known reactivities of the organisms from published sources (4), and a careful analysis of microscopic morphology was performed. In all cases, biochemical and phenotypic profiles obtained prior to molecular identification agreed extremely well with available published profiles for the given organism (data not shown).

DISCUSSION

This report has presented the results of yeast identification at the United Kingdom MRL over a 24-month period. Although 50 different species of yeast were identified, the 12 most common species of yeast comprised 95% of isolates (Table 1). In accordance with previous reports (19, 26), *C. albicans*, *C.*

glabrata, *C. parapsilosis*, *C. tropicalis*, *C. lusitanae*, *C. krusei*, *C. guilliermondii*, *C. dubliniensis*, and *Cryptococcus neoformans* remain the most prevalent yeast species encountered in clinical isolates. However, it should be noted that the absolute numbers of isolates for a given species received at the MRL are likely to be biased. Due to the reference nature of our activity, isolates submitted for identification are by necessity those that have failed to be identified by the referring laboratories. Thus, this report is not an epidemiological survey of yeast species. Our data demonstrate that the vast majority of these common isolates can be readily identified by the commercially available AUXACOLOR2 kit, underscoring its utility as a first-line identification system (Table 1). Isolates that failed to be identified by this system could frequently be identified by using the API 20C system as a second-line test (Table 1). Again, it should be noted that we have not attempted to compare the absolute success rates for identifications using AUXACOLOR2 as opposed to API 20C but rather have adhered to the strategy used in the laboratory of sequential testing of difficult isolates with AUXACOLOR2, followed by API 20C.

Even 10 years ago, more than 150 species of yeast had been associated with human pathologies (17). It is thus not surprising that the remaining 5% of isolates received by the MRL included an additional 38 different yeast species. Some of these organisms (e.g., *C. inconspicua*, *C. kefyri*, *C. famata*, *C. pelliculosa*, *C. lipolytica*, *C. norvegensis*, and *Geotrichum* sp.) are relatively uncommon but are nonetheless recognized potential human pathogens and as such are included in the databases of the AUXACOLOR2 and API 20C kits. At least some isolates

of these species were correctly identified without recourse to molecular techniques, although the success rates were generally much lower than with the 12 most common species (Table 1). Others of these less common species (e.g., *C. blankii*, *C. fabianii*, *C. pararugosa*, *C. catenulata*, and *C. pintolopesii*) are very rarely encountered in the clinical setting and are not included in the databases supplied with commercial identification kits. Not surprisingly, all isolates of these rare yeast species, some of which had been cultured from clinically significant sites (Table 2), required molecular identification.

The molecular identification strategy used here involved PCR amplification of the D1/D2 region of the gene encoding the large ribosomal subunit (26S) RNA, followed by automated sequencing of the resulting amplicons. A similar approach but involving sequencing of the ITS regions of the nuclear rRNA gene cassette previously discriminated between some 40 species of clinically important yeasts (12, 13). The rationale for the choice of the D1/D2 region for the present study was twofold. The D1/D2 primers that we used are capable of binding virtually all yeast species (with the possible exception of *Cryptococcus neoformans*), whereas our own results with primers targeted to the ITS1 region have proved more variable (data not shown). Second, the number of sequences from different fungal species corresponding to the D1/D2 region in the public synchronized databases far exceeds those corresponding to other regions of fungal genomes. For example, basic string searches in EMBL using “*Candida*” yielded 1,805 sequences for “26S/28S” compared to only 1,167 using “internal transcribed spacer region 1/ITS1.” In addition, although the sequence of the D1/D2 region of the fungal genome is more conserved than that of the flanking spacer regions, we have shown here that sequence variation in the D1/D2 region is easily sufficient to allow unambiguous speciation of all yeast isolates encountered to date. Indeed, our unpublished results also indicate that this approach successfully identified isolates of several of the recently described *Candida* species that to date can only be identified by molecular approaches, including *C. orthopsilosis* and *C. metapsilosis* (29) and *C. nivariensis* (1; data not shown). Indeed, given the inability of standard phenotypic methods to distinguish some of these new *Candida* species, it is possible that molecular methods may ultimately become the primary means of identification of clinically important yeast isolates.

Species have been included in the present study due to their occurrence in clinical specimens. They have not been selected according to any systematic or phylogenetic criteria, and the full list includes both ascomycete and basidiomycete species. Some of these, such as *C. rugosa* and *C. pararugosa*, are very closely related, whereas others such as *Geotrichum* sp. and *Sporobolomyces* sp. are only very distantly related. Under these circumstances it would not be appropriate to use D1/D2 sequences to attempt to formally reconstruct phylogeny (see references 3, 10, and 25). The D1/D2 sequences have instead been used to produce identifications on the basis of the best match shown between unknown and reference sequences. In this process the final identification is made from a measure of agreement over all of the characters and so is a simultaneous method, as defined previously (28). Since PCA only places sequences in *n*-dimensional space and does not impose any grouping or phylogenetic criteria on results, it was used to

demonstrate the general pattern of variation between sequences (28). The PCA result here shows that there is sufficient variation between sequences for each to occupy a unique position within the ordination, and in this case that can be achieved with only the 34% of the total variation contained in the first three dimensions (Fig. 1).

In summary, we have presented data demonstrating that sequencing of the D1/D2 region of the 26S/28S rRNA gene is a robust means for identifying many clinically relevant yeast isolates. Although such an approach is not necessary for the more common pathogenic yeasts encountered (which are reliably identified by commercially available AUXACOLOR2 and API 20C yeast identification kits), it is an important if not essential adjunct for the correct identification of many of the rarer clinically relevant yeast species.

ACKNOWLEDGMENTS

We are grateful to Whatman International for supplying the FTA cards and wash reagents used in much of this work, and we thank the members of the MRL past and present for their contributions to yeast identification in the laboratory.

REFERENCES

- Alcoba-Florez, J., S. Mendez-Alvarez, J. Cano, J. Guarro, E. Perez-Roth, and M. del Pilar Arevalo. 2005. Phenotypic and molecular characterization of *Candida nivariensis* sp. nov., a possible new opportunistic fungus. *J. Clin. Microbiol.* **43**:4107–4111.
- Altschul S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Bainbridge, B. W. 1994. Modern approaches to the taxonomy of *Aspergillus*, p. 291–301. In K. A. Powell, A. Renwick, and J. F. Peberdy (ed.), *The genus Aspergillus*. Plenum Press, Inc., New York, NY.
- Barrett, J. A., R. W. Payne, and D. Yarrow. 2000. *Yeasts: characterisation and identification*, 3rd ed. Cambridge University Press, Cambridge, United Kingdom.
- Bautista-Munoz, C., X. M. Boldo, L. Villa-Tanaca, and C. Hernandez-Rodriguez. 2003. Identification of *Candida* spp. by randomly amplified polymorphic DNA analysis and differentiation between *Candida albicans* and *Candida dubliniensis* by direct PCR methods. *J. Clin. Microbiol.* **41**:414–420.
- Bhally, H. S., S. Jain, C. Shields, N. Halsey, E. Cristofalo, and W. G. Merz. 2006. Infection in a neonate caused by *Pichia fabianii*: importance of molecular identification. *Med. Mycol.* **44**:185–187.
- Borman, A. M., C. K. Campbell, C. J. Linton, P. D. Bridge, and E. M. Johnson. 2006. *Polycyrtella hominis* is a mutated form of *Scedosporium apiospermum*. *Med. Mycol.* **44**:33–39.
- Borman, A. M., C. J. Linton, S.-J. Miles, C. K. Campbell, and E. M. Johnson. 2006. Ultra-rapid preparation of total genomic DNA from isolates of yeast and mould using Whatman FTA filter paper technology: a re-usable DNA archiving system. *Med. Mycol.* in press.
- Boucher, H., S. Mercure, S. Montplaisir, and G. Lemay. 1996. A novel group I intron in *Candida dubliniensis* is homologous to a *Candida albicans* intron. *Gene* **180**:189–196.
- Bruns, T. D., T. J. White, and J. W. Taylor. 1991. Fungal molecular systematics. *Annu. Rev. Ecol. Systematics* **22**:525–564.
- Campbell, C. K., K. G. Davey, A. D. Holmes, A. Szekeley, and D. W. Warnock. 1999. Comparison of the API *Candida* system with the AUXACOLOR2 system for identification of common yeast pathogens. *J. Clin. Microbiol.* **37**:821–823.
- Chen, Y.-C., J. D. Eisner, M. M. Kattar, S. L. Rassoulian-Barrett, K. LaFe, S. L. Yarfitz, A. P. Limaye, and B. T. Cookson. 2000. Identification of medically important yeasts using PCR-based detection of DNA sequence polymorphisms in the internal transcribed spacer region 2 of the rRNA genes. *J. Clin. Microbiol.* **38**:2302–2310.
- Chen, Y.-C., J. D. Eisner, M. M. Kattar, S. L. Rassoulian-Barrett, K. LaFe, U. Bui, A. P. Limaye, and B. T. Cookson. 2001. Polymorphic internal transcribed spacer region 1 DNA sequences identify medically important yeasts. *J. Clin. Microbiol.* **39**:4042–4051.
- Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton. 2004. The Jalview Java alignment editor. *Bioinformatics* **20**:426–427.
- Davey, K. G., P. M. Chant, C. S. Downer, C. K. Campbell, and D. W. Warnock. 1995. Evaluation of the AUXACOLOR2 system, a new method of clinical yeast identification. *J. Clin. Pathol.* **48**:807–809.
- Felsenstein, J. 1989. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* **5**:164–166.

17. **Fromtling, R. A.** 1995. Mycology, p. 697–877. In P. R. Murray, E. J. Baron, M. A. Tenover, and R. H. Tenover (ed.), *Manual of clinical microbiology*, 6th ed. ASM Press, Washington, DC.
18. **Guiver, M., K. Levi, and B. A. Oppenheim.** 2001. Rapid identification of *Candida* species by TaqMan PCR. *J. Clin. Pathol.* **54**:362–366.
19. **Hazen, K. C.** 1995. New and emerging yeast pathogens. *Clin. Microbiol. Rev.* **8**:462–478.
20. **Issakainen, J., J. Jalava, J. Saari, and C. K. Campbell.** 1997. Relationship of *Scedosporium prolificans* with *Petriella* confirmed by partial LSU rDNA sequences. *Mycol. Res.* **103**:1179–1184.
21. **Johnson, E. M., and D. W. Warnock.** 1995. Azole drug resistance in yeasts. *J. Antimicrob. Chemother.* **36**:751–755.
22. **Law, D., C. B. Moore, H. M. Wardle, L. A. Ganguli, and D. W. Denning.** 1994. High prevalence of antifungal resistance in *Candida* spp. from patients with AIDS. *J. Antimicrob. Chemother.* **34**:659–668.
23. **Leinberger, D. M., U. Schumacher, I. B. Autenrieth, and T. T. Bachmann.** 2005. Development of a DNA microarray for detection and identification of fungal pathogens involved in invasive mycoses. *J. Clin. Microbiol.* **43**:4943–4953.
24. **Mannarelli, B. M., and C. P. Kurtzman.** 1998. Rapid identification of *Candida albicans* and other human pathogenic yeasts by using short oligonucleotides in a PCR. *J. Clin. Microbiol.* **36**:1634–1641.
25. **Moritz, C., and D. M. Hillis.** 1996. Molecular systematics: context and controversies, p. 1–13. In D. M. Hillis, C. Moritz, and B. K. Mable (ed.), *Molecular systematics*, 2nd ed. Sinauer Associates, Sunderland, MA.
26. **Pfaller, M. A., R. N. Jones, S. A. Messer, M. B. Edmond, and R. P. Wenzel.** 1998. National surveillance of nosocomial blood stream infection due to species of *Candida* other than *Candida albicans*: frequency of occurrence and antifungal susceptibility in the SCOPE Program. *Diagn. Microbiol. Infect. Dis.* **31**:327–332.
27. **Ruhnke, M.** 2006. Epidemiology of *Candida albicans* infections and role of non-*Candida albicans* yeasts. *Curr. Drug Targets* **7**:495–504.
28. **Sneath, P. H., and R. R. Sokal.** 1973. *Numerical taxonomy*. W. H. Freeman & Co., San Francisco, CA.
29. **Tavanti, A., A. D. Davidson, N. A. Gow, M. C. Maiden, and F. C. Odds.** 2005. *Candida orthopsilosis* and *Candida metapsilosis* spp. nov. to replace *Candida parapsilosis* groups II and III. *J. Clin. Microbiol.* **43**:284–292.
30. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
31. **Wright, W. L., and R. P. Wenzel.** 1997. Nosocomial *Candida*: epidemiology, transmission, and prevention. *Infect. Dis. Clin. N. Am.* **11**:411–425.
32. **Yang, R., D. Li, and J. Zhao.** 2001. 26S rDNA D1/D2 domain sequence analysis of the pathogenic strain of the first case of disseminated trichosporonosis in China. *Zhonghua Yi Xue Za Zhi* **81**:472–475.
33. **Yokoyama, K., S. K. Biswas, M. Miyaji, and K. Nishimura.** 2000. Identification and phylogenetic relationship of the most common pathogenic *Candida* species inferred from mitochondrial cytochrome *b* sequences. *J. Clin. Microbiol.* **38**:4503–4510.