JOURNAL OF CLINICAL MICROBIOLOGY, Oct. 2010, p. 3648–3653 Vol. 48, No. 10
0095-1137/10/$12.00 doi:10.1128/JCM.01071-10

Rapid Molecular Identification of Pathogenic Yeasts by Pyrosequencing Analysis of 35 Nucleotides of Internal Transcribed Spacer 2[▽]

Andrew M. Borman,* Christopher J. Linton, Debra Oliver, Michael D. Palmer, Adrien Szekely, and Elizabeth M. Johnson

Mycology Reference Laboratory, Health Protection Agency, Bristol, United Kingdom

Received 27 May 2010/Returned for modification 1 July 2010/Accepted 2 August 2010

Rapid identification of yeast species isolates from clinical samples is particularly important given their innately variable antifungal susceptibility profiles. Here, we have evaluated the utility of pyrosequencing analysis of a portion of the internal transcribed spacer 2 region (ITS2) for identification of pathogenic yeasts. A total of 477 clinical isolates encompassing 43 different fungal species were subjected to pyrosequencing analysis in a strictly blinded study. The molecular identifications produced by pyrosequencing were compared with those obtained using conventional biochemical tests (AUXACOLOR2) and following PCR amplification and sequencing of the D1-D2 portion of the nuclear 28S large rRNA gene. More than 98% (469/477) of isolates encompassing 40 of the 43 fungal species tested were correctly identified by pyrosequencing of only 35 bp of ITS2. Moreover, BLAST searches of the public synchronized databases with the ITS2 pyrosequencing signature sequences revealed that there was only minimal sequence redundancy in the ITS2 under analysis. In all cases, the pyrosequencing signature sequences were unique to the yeast species (or species complex) under investigation. Finally, when pyrosequencing was combined with the Whatman FTA paper technology for the rapid extraction of fungal genomic DNA, molecular identification could be accomplished within 6 h from the time of starting from pure cultures.

Invasive fungal infections caused by *Candida* spp. remain a significant cause of morbidity and mortality in immunocompromised patients and those undergoing invasive procedures (19, 22, 26). While most studies agree that *Candida albicans* is the principal agent of nosocomial yeast infections, more than 150 yeast species from *Candida* and other genera have now been reported from mammalian infections (9), with *C. glabrata*, *C. parapsilosis*, *C. tropicalis*, *C. lusitaniae*, and *C. krusei* emerging over recent years to be significant opportunistic pathogens (16, 18, 23, 25). It is now well established that the patterns of antifungal susceptibilities vary substantially between different *Candida* species (20, 21). In addition, the widespread use of antifungal agents has been postulated to have contributed to a shift in species distributions via the emergence of inherently resistant species as significant pathogens (12, 13). Thus, informed therapeutic decisions frequently require the correct and rapid identification of an ever increasing number of potential pathogens.

The commercial systems available for the conventional identification of pathogenic yeasts rely upon a combination of morphological and biochemical features, employ tests that are time-consuming, and are designed to identify only the more common pathogens. Indeed, previous studies have demonstrated that such methods frequently fail to identify the less common pathogens and those common organisms that are misbehaving (1, 15) or to discriminate between closely related species (15, 27). Molecular identification, employing PCR amplification and sequencing of genomic regions that evolve slowly and that show high degrees of conservation, represents a rapid and sensitive alternative to conventional identification for yeasts and is also useful for the establishment of phylogenetic relationships. Among the genomic regions targeted, the nuclear rRNA gene cassette (and, notably, the internal transcribed regions [ITSs] and the D1-D2 portion of the 28S large rRNA gene) have previously proved sufficient to discriminate between the majority of species of clinically important yeasts (7, 8, 15). However, such approaches are still costly and cannot easily be performed more quickly than *in vitro* antifungal susceptibility testing.

Pyrosequencing technology (originally created by Pyrosequencing AB, now Biotage) is an extremely rapid DNA sequencing approach that employs novel chemistry to robustly sequence relatively short (<70-bp) target regions. Several reports have suggested that pyrosequencing could allow the identification of medically important yeasts (5, 10, 17), although for the most part only a few isolates representing the more common species were examined. In addition, we have previously shown that pyrosequencing of a region within ITS2 could identify some rare and closely related yeast species and was able to distinguish *C. glabrata* from its close genetic relative, *C. nivariensis* (4), and also to discriminate between *C. parapsilosis*, *C. orthopsilosis*, and *C. metapsilosis* (3).

In the current study, we have assessed the possibility of using pyrosequencing of a short fragment of ITS2 for identification of a wide spectrum of clinically important yeast species. In excess of 450 yeast isolates referred to the United Kingdom National Mycology Reference Laboratory (MRL),

* Corresponding author. Mailing address: Mycology Reference Laboratory, Health Protection Agency South-West Regional Laboratory, Myrtle Road, Kingsdown, Bristol BS2 8EL, United Kingdom. Phone: 0117 926 8683. Fax: 0114 922 6611. E-mail: Andy.Borman@uhBristol.nhs.uk.

▽ Published ahead of print on 11 August 2010.

TABLE 1. Selected NCPF reference and type strains included in the study

Organism	Collection designation(s) for representative reference isolate(s) employed in study
Candida albicans	NCPF 3179 (CBS6341, ATCC 10231)
Candida bracarensis	NCPF 8894^T
Candida blankii	NCPF 8830
Candida glabrata	NCPF 3309 (CBS138^T, ATCC 2001^T)
Candida pseudoglaebosa	NCPF 8832
Candida inconspicua	NCPF 3859 (CBS180^T, ATCC 16783^T)
Candida kefyr	NCPF 8678
Candida krusei	NCPF 3953 (CBS573^T, ATCC 6258^T)
Candida lambica	NCPF 8754, NCPF 8835
Candida lusitaniae	NCPF 2579, NCPF 2773, NCPF 3833, NCPF 3924, NCPF 3954, NCPF 3968, NCPF 8012, NCPF 8031, NCPF 8032, NCPF 8150, NCPF 8152, NCPF 8175, NCPF 8178, NCPF 8234, NCPF 8242, NCPF 8243
Candida lipolytica	NCPF 8630
Candida metapsilosis	NCPF 8768, NCPF 8789
Candida nivariensis	NCPF 8842, NCPF 8843, NCPF 8844, NCPF 8845, NCPF 8846, NCPF 8847, NCPF 8848, NCPF 8849, NCPF 8850, NCPF 8851, NCPF 8852, NCPF 8853
Candida norvegensis	NCPF 3861 (CBS1922, ATCC 22799)
Candida orthopsilosis	NCPF 8767, NCPF 8793, NCPF 8794, NCPF 8795, NCPF 8796, NCPF 8797, NCPF 8798, NCPF 8799, NCPF 8800, NCPF 8801
Candida palmioleophila	NCPF 8778
Candida parapsilosis	NCPF 8766 (CBS604^T, ATCC 22019^T)
Candida pelliculosa	NCPF 8623
Candida pseudointermedia	NCPF 8724
Candida rugosa	NCPF 8690
Candida zeylanoides	NCPF 8426
Cryptococcus gattii	NCPF 8446, NCPF 8600, NCPF 8868, NCPF 8869
Cryptococcus terreus	NCPF 8520
Filobasidiella uniguttulatus	NCPF 8427
Kazachstania sp.	NCPF 8895
Malassezia pachydermatis	NCPF 8494, NCPF 8562
Pichia membranifaciens	NCPF 3638
Pseudozyma aphidis	NCPF 848
Rhodotorula mucilaginosa	NCPF 8488, NCPF 8529, NCPF 8561
Sporobolomyces salmonicolor	NCPF 8488, NCPF 8735
Zygosaccharomyces bailii	NCPF 8416

Bristol, United Kingdom, for identification were submitted to conventional identification, molecular identification based on 28S rRNA gene sequencing, and also pyrosequencing. Pyrosequencing successfully allowed the rapid and unambiguous identification of >98% of isolates, which encompassed 40 different species of pathogenic yeasts.

MATERIALS AND METHODS

Test isolates. A total of 477 isolates were tested. These included 88 reference isolates and type strains that had been stored in the National Collection of Pathogenic Fungi (NCPF) (Table 1) and 389 additional clinical isolates that had been submitted to MRL for identification. 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 identities were revealed only after completion of conventional and molecular analyses.

Conventional identification methods. All the clinical isolates included in the present study were subjected to the conventional yeast identification strategy currently employed by the MRL: following initial germ tube testing, isolates were subjected to testing with the AUXACOLOR2 identification kit (Bio-Rad, Marnes-la-Coquette, France) exactly as described previously (6). Isolates were also subjected to PCR amplification-sequencing of the D1-D2 portion of the 28S rRNA gene exactly as described previously (15), using fungal genomic DNA that had been prepared using prepunched Whatman FTA papers as described previously (2).

Pyrosequencing of ITS2. Pyrosequencing analysis of a portion of ITS2 was performed essentially as described previously (3, 4), using genomic DNA prepared using prepunched Whatman FTA filters (2). Briefly, PCR amplification of a fragment of ITS2 was performed in 50-μl reaction volumes in the presence of 500 nM primers PyroF2B (5′-biotin-CTTTGAACGCACATTGCGCCCTCTGG TATTCC-3′) and PyroR2W (5′-TCCYCCGCTTATTGATATGCTTAAGTTC AGC-3′), 200 μM each deoxynucleoside triphosphate, 1.25 U of HotStar *Taq* polymerase (Qiagen, Valencia, CA), and a single FTA filter punch. Following enzyme activation at 94°C for 15 min, the reaction mixtures were subjected to 40 thermal cycles on a GeneAmp PCR systems 9700 thermocycler (Applied Biosystems, Foster City, CA) with the following parameters: 94°C (15 s), 50°C (15 s), and 72°C (90 s). 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. ITS2 amplification products were subjected to pyrosequencing analysis using the primer PyroYS (5′-GCGGGTAGTCCTACC TGATTTGAG-3′) and the reagents supplied with the Pyrogold SQA kit using a PyroMark ID pyrosequencing instrument (Biotage AB, Sweden). Sequencing was performed on 40 μl of each successful PCR product, using 15 cycles of introduction of each the 4 nucleotides (60 nucleotide introductions in total). Analysis of the resulting sequences was performed using IdentiFire software (Biotage) with an extended sequence database generated at MRL using reference and type species. IdentiFire default parameters were employed for sequence analysis, with the exception of "sequence to be analyzed," which was set to "user-defined, nucleotides 1 to 35." The extended sequence database is available from MRL upon request (enquiries to Andy.Borman@uhBristol.nhs.uk).

RESULTS

Table 2 shows the pyrosequencing profiles generated for the various yeast species analyzed. Unique pyrosequencing profiles were obtained for the majority of the species tested and ranged from 35 nucleotides to in excess of 60 nucleotides in length

TABLE 2. Manual alignments of the ITS2 pyrosequencing signature sequences[a]

Organism	Pyrosequencing profile	Similarity (no. of nucleotides)	Total no. of isolates tested (no. of reference isolates/ no. of clinical isolates)	No. of isolates tested previously (reference)
Candida albicans	**GTCAAA----GTTTGAAGATATACG-TGG-TAGACGTTACC**	35	21 (1/20)	21 (3)
Lodderomyces elongisporus	**GTC**g**AA----GTTTGAA-ATATA**ga-**TT**t**G-**g**AG**ctt**TTA**tt	24	2 (0/2)	2 (3)
Candida dubliniensis	**GTCAAA----GTTTGAAGA**ata**A**aa-**TGG-**gc**GACG**ccaga	23	14 (0/14)	
Candida tropicalis	**GTCAAA----GTT**atg**A**a**ATA**a**A**ttg**TGG-TG**g**C**c**ac**t**TAG**C	23	34 (0/34)	
*Saccharomyces cerevisiae**	**GTCAAA----**c**TTTA**gaA**c**a**T**ttgttc**G**cc**TAGACG**c**T**ct**C**	22	16 (0/16)	
Candida metapsilosis	**GTC**g**AA-----TTTG**g**A**aga**AT**gtt-**TGG--AG**tttg**TACC**	21	4 (2/2)	3 (3)
Candida pelliculosa	**GTCAAA----**c**TTT**t**A**gtt**TAT**t-**G**t**TG**t-**TA**ag**C**cgag**CC**	21	10 (1/9)	
Candida parapsilosis	**GTC**g**AA-----TTTG**g**A**aga**A**gttt-**TGG--AG**tttg**TACC**	20	69 (1/68)	69 (3)
Candida orthopsilosis	**GTC**g**AA-----TTTG**g**A**aga**A**-ttt-**TGG--AG**tttg**TACC**	20	11 (10/1)	11 (3)
Filobasidiella uniguttulatus	**GTCA**g**A----**tg**T**ca**AAG-TATAC**aca**GG-**a**AG**caacac**C**a	20	1 (1/0)	
Pichia membranifaciens	**GTC**g**A-----G**c**T**ca**A**t**GATATA**ttt**T**c**G**c**T**c**GG**c**G**gc**A**aa	20	3 (1/2)	
Candida palmioleophila	**GTCAAA**ctt-**GTTTG**tt**G**t**T**t**T**ttaag**G**caa**AG**c**C**taaca**C**	19	2 (1/1)	
Candida glaebosa	**GTCAAA**ctt-**GTTTG**t-**G**t**T**t**TA**taag**G**caa**AG**c**C**taatga	19	1 (1/0)	
Zygosaccharomyces bailii	**GTCAAA----**c**TTTGA**gag**TAT**tgtt**TG**c-cca**A**g**G**cgcg**C**	19	1 (1/0)	
Candida kefyr	**GTCAAA----**c**TTTGA**gag**T**t**T**tg**G**t**T**aa-agcggta**T**g**CC**	19	16 (1/15)	1 (4)
Candida bracarensis	**GTCAAA----**c**TT**aa**A**g**G**t**T**tctgt-**TT**G-cc**G**tg**GT**c**AC**a	18	1 (1/0)	
Candida eremophila	**GTC**g**A**gctta**GTT**aa**AAG**t**T**cgg**CG--G**cca**A**ag**CGT**gcta	18	1 (1/0)	
Candida blankii	**GTCAAA----**t**TTTG**g**AG**cgg**T**tgttac**G**cct**G**t**C**tcg**AaC**	18	1 (1/0)	
Kazachstania sp.*	**GTCAAA----**c**TT**aa**AAGA**acact**G**t**T**c**G--CCACG**gctgt	18	3 (1/2)	
Candida nivariensis	**GTCAAA----**c**TT**aa**A**g**G**t**T**cctgt-**TT**g**G**cc**AG**ca**G**act**TC**	17	26 (12/14)	16 (4)
Candida utilis	**GTCAA**g----c**TT**a**GAA**aggtgtta-a**G**ccg**AG-C**tc**T**g**CC**	17	1 (0/1)	
*Cryptococcus neoformans**	**GTCAAA-----**caaa**AAGA**g**AT**ggt-**TG**ttatc**A**gcaag**CC**	17	17 (5/12)	
Sporopachydermia sp.	**GTCAAA---**ga**TTTA**g**A**tc**T**t**T**gtcaa**G**c-a**AGAC**aaacta	17	2 (0/2)	
Sporobolomyces salmonicolor	a**TC**t**AA**----tc**T**aa**AAG**g**TA**g**AC**tt**T**a**G**gg**A**tta**G**aag**C**t	17	2 (2/0)	
*Candida fabianii**	**GTCAAA----**c**TT**atg**A**aga**A**att**G**t**T**a**G**gcc**GA**gccc**A**aa	16	9 (0/9)	
Candida glabrata	**GTCAAA----**c**TT**aa**AG**acgtctgtc**TG**cc**T**cag**CG**acg**C**a	16	47 (1/46)	13 (4)
Candida inconspicua	**GTC**g**A**g----c**TT-GA**tt**A**a**A**ag**T**t-c**GG**cgg**G**ga**G**aacg**C**	16	11 (1/10)	1 (4)
Candida lusitaniae	**GgCgAA----**a**T**g**T**cgt**G**c**T**g**TA**acaa**G**ct**TA**act**GTT**tta	16	59 (35/24)	
Candida zeylanoides	**GTCAAA----**c**TTTG**ttt**g**T**tgtt**G**taa**G**gcc**GA**gcc**T**gtg	16	1 (1/0)	1 (4)
*Candida famata**	**GTCAAA----**c**TT**gttt**G**t**TATA**ttg**T**aaggccga**G**cctag	15	4 (0/4)	
*Candida guilliermondii**	**GTCAAA----**c**TT**gttt**GG**t**T**gtt**G-T**aaggccgg**G**cc**AaC**	15	27 (0/27)	
Candida norvegensis	**GTC**g**A**g----c**TT**a**GA**ttta**A**a**A**aaaaa**G**t**T**c**GG**c**G**ggcca	15	4 (1/3)	1 (4)
Pseudozyma sp.*	**G**c**C**g**A**t---gaa**TTGAA**at**TA**a**A**tcccttcctct**C**c**TT**c**C**g	15	1 (1/0)	
Rhodotorula mucilaginosa	a**TC**t**AA----**tc**TT**a**AA-AT**g**TA**gacattc**T**gattaga**AG**C**	15	7 (3/4)	
Cryptococcus saitoi	**G**c**CA**g**A----**tg**TT**atga**ATAT**taa-**T**ccga**AGA**tcaatgg	14	1 (0/1)	
*Cryptococcus terreus**	**G**c**CA**g**A----**taa**T**a**AA-A**a**A**gtca-**TG**tcc**A**cga**GgT**gga	14	1 (1/0)	
*Candida pseudointermedia**	**GgCgAA**--aa**G**aa**T**a**AAG**t**T**ga**A**gtaac**G-TA**ttgcaacaa	14	2 (2/0)	
Malassezia pachydermatis	ac**CA**g**A**----aa**T**ga**AA**a**A**gga**A**t**G**-**T**act**T**ttca**GTT**cgt	14	2 (2/0)	
*Candida krusei**	**GTC**g**A**g----c**TTT**ttgttgtct**CG**caacactcg**C**tc**T**cgg	11	28 (1/27)	1 (4)
Candida rugosa	aat**AA**c----**GT**caa**A**g**G**g**T**ccgtaacaagctt**A**ac**T**gttt	9	6 (3/3)	
Total			469 (40 species)	
Candida lambica	Successful PCR; no pyrosequence generated		2 (2/0)	
Candida lipolytica	Successful PCR; no pyrosequence generated		2 (1/1)	
Candida pararugosa	Successful PCR; no pyrosequence generated		4 (0/4)	

[a] The sequences are aligned and nucleotide conservation (similarity column) with respect to the 35-nucleotide *C. albicans* sequence was calculated. Conserved nucleotides are depicted in uppercase bold letters, and variable positions are depicted in lowercase letters. Hyphens have been introduced to improve alignments. Asterisks denote those organisms for which the pyrosequencing profile was not unique to the species in question, as judged by BLAST searches performed against the public synchronized DNA databases (see text and Table 3).

(data not shown). In all cases there was 100% agreement between the pyrosequencing identification and the identification obtained by PCR amplification and conventional sequencing of some 300 to 400 bp of the D1-D2 region of each organism (Table 2 and data not shown). While correct identification of some species could be achieved by analysis of only 20 to 25 bp of sequence generated by pyrosequencing analysis, analysis of the first 35 bp allowed correct identification for 469 of the 477 isolates examined (Table 2). As reported previously, pyrosequencing analysis was able to discriminate between very closely related organisms and species complexes. Pyrosequencing separated *C. albicans* from *C. dubliniensis* and *C. glabrata* from *C. nivariensis* and *C. bracarensis* and also correctly identified the members of the *C. parapsilosis* species complex (*C. parapsilosis*, *C. orthopsilosis*, and *C. metapsilosis* [24]) (Table 2) (3, 4).

For the 469 isolates that were correctly identified by pyrosequencing, the similarities between the pyrosequencing signature sequence and the reference ITS2 sequences present in our database ranged from 80 to 100%. As previously reported, all mismatches resulted from machine errors in incorrectly discerning the numbers of nucleotides in homopolymeric stretches (17). These errors never resulted in the erroneous identification of one species as another and were easily corrected manually after close inspection of the pyrograms (data not shown). Eight isolates, corresponding to three species (*C. lambica*, $n = 2$; *C. lipolytica*, $n = 2$; *C. pararugosa*, $n = 4$) could not be identified by pyrosequencing analysis. Despite successful amplification of the appropriate portion of ITS2 with our PCR primers (Table 2; data not shown), no signature sequence could be obtained after pyrosequencing. A previous study (17) using pyrosequencing primers developed commercially by Biotage to target the same ITS2 reported similar failures with these species, which are likely due to sequence variations in the

TABLE 3. Organisms demonstrating potential sequence redundancy in the ITS2 target[a]

Organism	Synonym	No. of nucleotides aligned	Other key BLASTN EMBL match(es)	Synonym(s)
Candida fabianii	*Pichia fabianii* (Te)	37	*Pichia amylophila, Pichia mississippiensis*	
Candida famata	*Debaromyces hansenii* (Te), *Debaromyces fabryi*	37	*Debaromyces nepalensis, Debaromyces coudertii, Debaromyces marama, Candida psychrophila*	
Candida guilliermondii	*Pichia guilliermondii* (Te)	36	*Candida carpophila*	*Candida fukuyamaensis, Candida xestobii*
			Candida fermentati	*Pichia caribbica* (Te)
Candida krusei	*Issatchenkia orientalis* (Te)	37	*Candida xylopsoci, Issatchenkia terricola, Pichia cecembensis, Pichia sporocuriosa*	
Candida pseudointermedia	*Filobasidiella neoformans* (Te)	38	*Candida intermedia*	
Cryptococcus neoformans		35	*Cryptococcus neoformans* var. *neoformans*, *Cryptococcus neoformans* var. *grubii*, *Cryptococcus gattii*	
Cryptococcus terreus	*Cryptococcus elinovii*	35	*Cryptococcus albidus, Cryptococcus phenolicus*	*Cryptococcus bacillisporus*
Kazachstania spp.		35	*Kazachstania aquatica, Kazachstania bovina, Kazachstania telluris, Kazachstania unispora*	
Pseudozyma spp.		38	*Pseudozyma aphidis, Pseudozyma antarctica, Pseudozyma rugulosa*	
Saccharomyces cerevisiae	*Saccharomyces boulardii*	37	*Saccharomyces paradoxus*	*Saccharomyces uvarum*
			Saccharomyces bayanus	*Saccharomyces mikatii*
			Saccharomyces mikatae	

[a] The yeast species subjected to pyrosequencing in the current study (organism column) are given, together with known synonyms (teleomorphs [Te]) and the list of key species (identified in BLASTN searches of the EMBL database) that show 100% nucleotide similarity over the target organism ITS2 signature sequence.

ITS2 target that prevent annealing of the pyrosequencing sequencing primer.

To assess whether the individual signature sequences generated by pyrosequencing were unique to the species under study, each of the signature sequences depicted in Table 2 was subjected to BLAST searches against the fungal sequences present in the public synchronized databases (Table 3; data not shown). For 30 of the 40 signature sequences, searches returned only the species that had been identified. For the other 10 species (denoted by asterisks in Table 2), BLAST searches returned several alternative species identifications, all with 100% identity over the ITS2 analyzed by pyrosequencing. However, for the most part, the alternative species identifications suggested by these database searches corresponded to rare yeast species that have been proposed to comprise cryptic/closely related species in species complexes containing the principal organism of interest (Table 3). For all such database searches, sequence matches that were returned were scrutinized for their likely validity by assessing whether the database entries were for type strains of the respective organisms and whether multiple examples of each sequence had been deposited from several independent laboratories. However, since we did not attempt to pyrosequence the type strains of the organisms returned in such searches, it is possible that some of the database returns represent organisms that had been erroneously identified by the depositor of the sequence. In summary, to date potential pyrosequencing signature sequence redundancy has not been observed between yeast species from unrelated genera or species complexes.

DISCUSSION

The current study has evaluated the utility of pyrosequencing analysis in the rapid identification of yeast isolates cultured from clinical specimens. A total of 477 test isolates, encompassing 43 yeast species and including both reference and type strains and strains from clinical samples, were subjected to pyrosequencing analysis of a portion of ITS2. Species were included in the present study due to their occurrence in clinical specimens. Test organisms were selected to include as many disparate species as possible, while ensuring that multiple isolates of the most clinically relevant species were included. The organisms included therefore do not in any way reflect an epidemiological analysis of *Candida* infections. All isolates had previously been unambiguously identified using combinations of biochemical/phenotypic methods and conventional sequencing of the D1-D2 portion of the 28S rRNA gene but were submitted for pyrosequencing in a rigorously blinded manner.

Pyrosequencing of a 35-nucleotide portion of ITS2 allowed the successful identification of 469/477 isolates and 40/43 species. The 8 isolates that failed to be identified comprised three species (*C. lambica*, *C. pararugosa*, *C. lipolytica*) that are only rarely encountered in the clinical setting. Indeed, these three species account for less than 0.5% of all yeast isolates that are submitted to the MRL for identification (see, for example, reference 14). The ITS2 pyrosequencing signature sequences generated in the current study were unique to the species under study for 30 of the 40 species that were successfully identified. For the remaining 10 species, BLAST analyses of signature sequences against the public DNA databases re-

TABLE 4. Sequence variation detected by pyrosequencing of putative isolates of *C. rugosa*[a]

Candida rugosa isolate	Pyrosequencing signature sequence	EMBL match
Position	1 10 20 30	
Montero 1	**AATAAGATCAA-GAGTCTGTAACAAGCTTAA**	*C. rugosa*, 100% (2 strains)
Montero 2	**AATAAGATC**gt-**GAGTCTGTAACAAGCTTAA**	*C. rugosa*, 100% (9 strains; **T**)[b]
NCPF 8690	**AATAAGATCAA-GAGTCTGTAACAAGCTTAACTGTTTTA**	*C. rugosa*, 100% (2 strains)
MRL338	**AATAA**cg**TCAA**a**G**g**GTC**c**GTAACAAGCTTAACTGTTTT**	*C. rugosa*, 100% (2 strains)

[a] The pyrosequencing ITS2 signature sequences obtained in the current study (NCPF 8690 and MRL338) are manually aligned against those described by Montero et al. (17), together with the results of BLASTN searches conducted with each signature sequence against the EMBL database (identification returned, percent identity, number of strain matches). The signature sequences for Montero 1 and NCPF 8690 differ only by the additional 8 nucleotides sequenced for NCPF 8690. Hyphens have been introduced to improve alignments; conserved nucleotides are given in uppercase bold letters.

[b] **T** denotes the type strain sequence in this portion of ITS2 (included among the nine strains).

turned several potential alternative species identifications (Table 3). In all cases, the alternative species were close genetic and phenotypic relatives of the test organism and can reasonably be considered members of single species complexes (Table 3). Currently, the clinical need to distinguish between such very close genetic relatives is uncertain (see, for example, reference 11), although it is somewhat unfortunate that the 35-nucleotide sequence analyzed by pyrosequencing did not allow separation of members of the *Cryptococcus neoformans* complex. Further studies will be required to determine if pyrosequencing of a second independent locus might allow identification of the individual species in these complexes and also the three species that to date have failed to be identified by pyrosequencing altogether.

Although previous studies have demonstrated that conventional sequencing of the entire ITS2 allowed the identification of greater than 98% of clinical yeast isolates from 86 species examined (14), the discriminatory power of pyrosequencing analysis of the 35-nucleotide region analyzed here is impressive. Pyrosequencing, when coupled with the Whatman FTA paper technology for the rapid extraction of fungal genomic DNA, could be accomplished within 6 h from the time of starting pure cultures and robustly discriminated between closely related species in several species complexes, including the *C. parapsilosis* species complex (*C. parapsilosis*, *C. orthopsilosis*, *C. metapsilosis*) and distinguished *C. glabrata* from *C. nivariensis* and *C. bracarensis*. This high discriminatory power is likely to stem at least in part from the extremely variable nature of the portion of ITS2 examined (Table 2). No single nucleotide position in the 35-nucleotide *C. albicans* signature sequence was conserved among all 40 species that yielded sequences. Moreover, sequence variability increased steadily from 5′ to 3′ across the ITS2 target, such that 16 of the 35 nucleotides in the *C. albicans* sequence were conserved in less than 50% of the study species (Table 2). Thus, overall sequence identity with *C. albicans* ranged from 25.7% (*C. rugosa*) to 68.6% (*Lodderomyces elongisporus*) over the target region, and 21 of the 40 species shared less than 50% identity with *C. albicans* in this region (Table 2). The extremely variable nature of this portion of ITS2 is probably the reason why our pyrosequencing discriminated more yeast species than pyrosequencing targeting a portion of the more conserved 18S rRNA gene reported by Gharizadeh et al. (10) and also explains why occasional errors of resolution of homopolymer runs in the ITS2 target sequence do not affect the final identification obtained by pyrosequencing.

Once errors in the resolution of homopolymeric runs were corrected, no intraspecific sequence variations were detected in the relatively short region of ITS2 examined by pyrosequencing analysis for all but one of the species examined. This is perhaps unsurprising, given that the variability in highly conserved genes such as those encoding ribosomal DNA is more limited than that in many other genes. However, three related but distinct pyrosequencing signature sequences were generated from isolates of *Candida rugosa* tested in a previous study (17) and in the current study (see the aligned sequences in Table 4). When each signature sequence was submitted to BLAST searches, all returned matches with 100% identities with reference strains of *C. rugosa* in the public synchronized databases (Table 4), but only one of the sequences matched the type strain for this organism (strain Montero 2; Table 4). Although it remains plausible that *C. rugosa* exhibits exceptional sequence variability in this short region of ITS2, we believe that these data might indicate that *C. rugosa* is in fact a species complex of several closely related but distinct organisms. Future studies will address this possibility.

In conclusion, we have demonstrated that pyrosequencing analysis of a small portion of ITS2 is sufficient to reliably identify the vast majority of pathogenic yeast species. The pyrosequencing profiles produced for each species were reproducible and conserved across multiple isolates and are to date unique to the species (or species complex) of pathogenic yeast under investigation. Due to the extreme rapidity and facility of pyrosequencing analysis and the relatively low cost per sample analyzed, we believe that this approach is ideally suited to the accurate identification of yeast isolates. The high-throughput nature of the pyrosequencing technology also makes it ideally suited to evaluate the prevalence of cryptic minority species. Given the increasing evidence indicating species-specific variations in antifungal susceptibility profiles, we believe that rapid and accurate yeast identification will become increasingly important for informed therapeutic decisions.

ACKNOWLEDGMENTS

We are grateful to the other members of the Mycology Reference Laboratory for their interest and advice, to Frank Odds for sharing isolates of *Candida metapsilosis* and *C. orthopsilosis*, and to Whatman International and Biotage for the gift of reagents.

A.M.B. was funded in part by travel grants from Whatman International and Biotage AB.

REFERENCES

1. **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.
2. **Borman, A. M., M. Fraser, C. J. Linton, M. D. Palmer, and E. M. Johnson.** 2010. An improved method for the preparation of total genomic DNA from isolates of yeast and mould using Whatman FTA filter papers. Mycopathologia **169:**445–449.
3. **Borman, A. M., C. J. Linton, D. Oliver, M. D. Palmer, A. Szekely, F. C. Odds, and E. M. Johnson.** 2009. Pyrosequencing analysis of 20 nucleotides of the internal transcribed spacer 2 discriminates *Candida parapsilosis*, *Candida metapsilosis*, and *Candida orthopsilosis*. J. Clin. Microbiol. **47:**2307–2310.
4. **Borman, A. M., R. Petch, C. J. Linton, M. D. Palmer, P. D. Bridge, and E. M. Johnson.** 2008. *Candida nivariensis*, an emerging pathogenic fungus with multidrug resistance to antifungal agents. J. Clin. Microbiol. **46:**933–938.
5. **Boyanton, B. L., R. A. Luna, L. R. Fasciano, K. G. Menne, and J. Versalovic.** 2008. DNA pyrosequencing-based identification of pathogenic *Candida* species by using the internal transcribed spacer 2 region. Arch. Pathol. Lab. Med. **132:**667–674.
6. **Campbell, C. K., K. G. Davey, A. D. Holmes, A. Szekely, A., 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.
7. **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.
8. **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.
9. **Fromtling, R. A.** 2007. Mycology. *In* P. R. Murray, E. J. Baron, J. Jorgensen, M. A. Pfaller, and M. L. Landry (ed.), Manual of clinical microbiology, 9th ed. ASM Press, Washington, DC.
10. **Gharizadeh, B., E. Norberg, J. Loffler, S. Jalal, J. Tollemar, H. Einsele, L. Klingspor, and P. Nyren.** 2004. Identification of medically important fungi by the Pyrosequencing technology. Mycoses **47:**29–33.
11. **Gomez-Lopez, A., A. Alatruey-Izquierdo, D. Rodriguez, B. Almirante, A. Pahissa, J. L. Rodriguez-Tudela, M. Cuenca-Estrella, and the Barcelona Candidemia Project Study Group.** 2008. Prevalence and susceptibility profile of *Candida metapsilosis* and *Candida orthopsilosis:* results from population-based surveillance of candidemia in Spain. Antimicrob. Agents Chemother. **52:**1506–1509.
12. **Johnson, E. M., and D. W. Warnock.** 1995. Azole drug resistance in yeasts. J. Antimicrob. Chemother. **36:**751–755.
13. **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.
14. **Leaw, S. N., H. C. Chang, H. F. Sun, R. Barton, J.-P. Bouchara, and T. C. Chang.** 2006. Identification of medically important yeast species by sequence analysis of the internal transcribed spacer regions. J. Clin. Microbiol. **44:** 693–699.
15. **Linton, C. J., A. M. Borman, G. Cheung, A. D. Holmes, A. Szekeley, M. D. Palmer, P. D. Bridge, C. K. Campbell, and E. M. Johnson.** 2007. Molecular identification of unusual pathogenic yeast isolates by large ribosomal subunit gene sequencing: 2 years experience at the United Kingdom Mycology Reference Laboratory. J. Clin. Microbiol. **45:**1152–1158.
16. **Marr, K. A.** 2004. Invasive *Candida* infections: the changing epidemiology. Oncology **14:**9–14.
17. **Montero, C. I., Y. R. Shea, P. A. Jones, S. M. Harrington, N. E. Tooke, F. G. Witebsky, and P. R. Murray.** 2008. Evaluation of Pyrosequencing® technology for the identification of clinically relevant non-dematiaceous yeasts and related species. Eur. J. Clin. Microbiol. Infect. Dis. **27:**821–830.
18. **Nucci, M., and K. A. Marr.** 2005. Emerging fungal diseases. Clin. Infect. Dis. **41:**521–526.
19. **Pfaller, M. A., and D. J. Diekema.** 2007. Epidemiology of invasive candidiasis: a persistent public health problem. Clin. Microbiol. Rev. **20:**133–163.
20. **Pfaller, M. A., R. N. Jones, G. V. Doern, A. C. Fluit, J. Verhoef, H. S. Sadler, S. A. Messer, A. Houston, S. Coffman, and R. J. Hollis.** 1999. International surveillance of blood stream infections due to *Candida* species in the European SENTRY program: species distribution and antifungal susceptibility including the investigational triazole and echinocandin agents. SENTRY Participant Group (Europe). Diagn. Microbiol. Infect. Dis. **35:**19–25.
21. **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.
22. **Ruhnke, M.** 2006. Epidemiology of *Candida albicans* infections and role of non-*Candida albicans* yeasts. Curr. Drug Targets **7:**495–504.
23. **Snydman, D. R.** 2003. Shifting patterns in the epidemiology of nosocomial *Candida* infections. Chest **123:**500S–503S.
24. **Tavanti, A., A. D. Davidson, N. A. R Gow, M. C. J. 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.
25. **Wingard, J. R.** 1994. Infections due to resistant *Candida* species in patients with cancer who are receiving chemotherapy. Clin. Infect. Dis. **19**(Suppl. 1)**:**49–53.
26. **Wright, W. L., and R. P. Wenzel.** 1997. Nosocomial *Candida*. Epidemiology, transmission and prevention. Infect. Dis. Clin. North Am. **11:**411–425.
27. **Yang, R., D. Li, and J. Zhao.** 2001. 26SrDNA 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.