

Identification and Phylogenetic Relationship of the Most Common Pathogenic *Candida* Species Inferred from Mitochondrial Cytochrome *b* Gene Sequences

KOJI YOKOYAMA,* SWARAJIT KUMAR BISWAS, MAKOTO MIYAJI,
AND KAZUKO NISHIMURA

Research Center for Pathogenic Fungi and Microbial Toxicoses, Chiba University,
1-8-1 Inohana, Chuo-ku, Chiba 260-8673, Japan

Received 29 June 2000/Returned for modification 9 August 2000/Accepted 16 September 2000

We sequenced a 396-bp region of the mitochondrial cytochrome *b* gene of the most common clinically important *Candida* species: *Candida albicans*, *C. glabrata*, *C. parapsilosis*, *C. tropicalis*, *C. krusei*, and *C. lusitaniae*. The recently described species of *Candida*, *C. dubliniensis*, associated with mucosal candidiasis in human immunodeficiency virus-infected individuals, was also included. Two to five strains of each species were examined. Some species represented intraspecies variation, which was not more than 1.8% (DNA). However, interspecies variations were more than 10 and 7%, respectively, for DNA and amino acid sequences. Multiple alignments of nucleotide and deduced amino acid sequences revealed species-specific nucleotides and amino acids. Nucleotide- and amino acid-based phylogenetic trees were constructed and are discussed. Using the database, it is possible to identify presumptive *Candida* species within a working day.

There has been a significant increase in the number of reports of systemic and mucosal infections caused by *Candida* species with the increase in the number of immunocompromised patients (3, 5, 9, 14, 21, 24, 25, 33). *Candida albicans* is

the most frequently isolated causative agent of candidal infection in humans (more than 50%) and is generally accepted as the most pathogenic species of the genus *Candida* (3). However, in recent years, non-*C. albicans* *Candida* species, e.g., *C.*

TABLE 1. Fungal strains, strain origins, and accession numbers of cytochrome *b* genes sequenced in this study^a

Species	IFM no.	Origin	Accession no.
<i>Candida albicans</i>	5728	Clinical isolate	AB044909
<i>Candida albicans</i>	46907	Clinical isolate	AB044910
<i>Candida albicans</i>	48311 ^T	CBS 562	AB044911
<i>Candida albicans</i>	48922	Clinical isolate	AB044918
<i>Candida albicans</i>	49689	Clinical isolate	AB044919
<i>Candida dubliniensis</i>	48184	Clinical isolate	AB044912
<i>Candida dubliniensis</i>	48313 ^T	CBS 7987	AB044913
<i>Candida dubliniensis</i>	48314	CBS 7988	AB044914
<i>Candida glabrata</i>	5489	Clinical isolate	AB044920
<i>Candida glabrata</i>	5768	Clinical isolate	AB044921
<i>Candida glabrata</i>	40065	Clinical isolate	AB044922
<i>Candida glabrata</i>	46843 ^T	CBS 138	AB044915
<i>Candida krusei</i> (“ <i>Issatchenkia orientalis</i> ”)	5798	Clinical isolate	AB044924
<i>Candida krusei</i> (“ <i>Issatchenkia orientalis</i> ”)	40019	Clinical isolate	AB044923
<i>Candida krusei</i> (“ <i>Issatchenkia orientalis</i> ”)	46834 ^T	CBS 573	AB044925
<i>Candida lusitaniae</i> (“ <i>Clavispora lusitaniae</i> ”)	49207 ^T	IFO 1019	AB044926
<i>Candida lusitaniae</i> (“ <i>Clavispora lusitaniae</i> ”)	49723	Clinical isolate	AB044927
<i>Candida parapsilosis</i>	5464	Clinical isolate	AB044928
<i>Candida parapsilosis</i>	40119	Clinical isolate	AB044929
<i>Candida parapsilosis</i>	46829 ^T	CBS 604	AB044916
<i>Candida tropicalis</i>	40085	Clinical isolate	AB044930
<i>Candida tropicalis</i>	46816 ^T	CBS 94	AB044917
<i>Candida tropicalis</i>	48776	Clinical isolate	AB044931
<i>Candida tropicalis</i>	49042	Clinical isolate	AB044932
<i>Candida tropicalis</i>	49101	Clinical isolate	AB044933
<i>Filobasidiella neoformans</i>	48637 ^T	CBS 132	AB040656

^a CBS, Centraalbureau voor Schimmelcultures, Baarn, The Netherlands; IFM, Institute for Food Microbiology (at present, the Research Center for Pathogenic Fungi and Microbial Toxicoses, Chiba University), Chiba, Japan; IFO, Institute for Fermentation, Osaka, Japan.

* Corresponding author. Mailing address: Research Center for Pathogenic Fungi and Microbial Toxicoses, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8673, Japan. Phone: 81-43-226-2789. Fax: 81-43-226-2486. E-mail: yoko@myco.pf.chiba-u.ac.jp.

TABLE 2. PCR primers used in this study

Primer	Sequence ^a	Description
E1M4	5'-TGRGGWGCWACWGTTACTACTA-3'	Forward
E1M5	5'-GTTACTACTAAAYTTATTMTTC-3'	Forward
E2M4	5'-GGWATAGMWSKTAAYAGCATA-3'	Reverse
E2mr5	5'-AGCACGTARWATTGCGTAGAATGG-3'	Reverse

^a R, A or G; W, A or T; Y, C or T; M, A or C; S, C or G; K, G or T.

glabrata (formerly *Torulopsis glabrata*) (10 to 30%), *C. parapsilosis* (10 to 20%), *C. tropicalis* (10 to 20%), *C. krusei*, and *C. lusitaniae*, have been recovered with increasing frequency from cases of candidiasis (9, 16, 24). *C. dubliniensis*, phenotypically very similar to *C. albicans* and associated primarily with recurrent oral infections in human immunodeficiency virus-infected individuals, was first described in 1995 (23). The clinical significance of this species is under investigation.

Identification of *Candida* isolates to the species level in the clinical laboratory has become more important. Some *Candida* species, such as *C. glabrata*, are known to rapidly acquire decreased susceptibility to fluconazole; *C. krusei* is considered to be inherently resistant to fluconazole (20, 32). However, amphotericin B resistance is uncommon but appears to be most common among isolates of *C. lusitaniae* (27). *C. parapsilosis* can survive in the hospital environment, a fact which increases the chance of nosocomial transmission (5, 15). Therefore, species-specific identification is necessary for timely, targeted, and effective antifungal therapy and to facilitate hospital infection control measures.

The utility of a gene segment in accurate identification and phylogenetic analysis depends on the frequency of recombinational interchange of genetic material within the chosen gene segment. Mitochondria (mt) are self-replicating; in the budding yeast *Saccharomyces cerevisiae*, mt enter the bud immediately after the emergence of the bud and are equally distributed, although it is not clear how the mother cell maintains its

TABLE 3. Variation of cytochrome *b* sequences within different strains of a single *Candida* species

Species and strain	% Variation ^a	
	DNA	Amino acid
<i>C. albicans</i>		
IFM 48311 ^T		
IFM 5728	0.0	0.0
IFM 46907	0.0	0.0
IFM 48922	0.0	0.0
IFM 49689	0.3	0.0
<i>C. glabrata</i>		
IFM 46843 ^T		
IFM 5489	0.0	0.0
IFM 5768	0.0	0.0
IFM 40065	0.0	0.0
X53862	0.3	0.8
<i>C. tropicalis</i>		
IFM 46816 ^T		
IFM 40085	1.8	0.8
IFM 48776	0.0	0.0
IFM 49042	1.8	0.8
IFM 49101	1.8	0.8

^a Compared to the type strain.

TABLE 4. Levels of cytochrome *b* nucleotide and amino acid sequence similarities for the most common pathogenic *Candida* species

Species	% Cytochrome <i>b</i> sequence similarity ^a for:						
	<i>C. albicans</i>	<i>C. dubliniensis</i>	<i>C. parapsilosis</i>	<i>C. lusitaniae</i>	<i>C. krusei</i>	<i>C. tropicalis</i>	<i>C. glabrata</i>
<i>C. albicans</i>		92.4	87.9	86.4	84.1	84.1	80.3
<i>C. dubliniensis</i>	89.9		87.9	84.1	84.8	86.4	75.8
<i>C. parapsilosis</i>	84.1	84.1		79.5	81.8	87.1	81.8
<i>C. lusitaniae</i>	83.3	84.6	81.6		80.3	74.2	79.5
<i>C. krusei</i>	81.8	84.1	81.6	84.3		75.8	79.5
<i>C. tropicalis</i>	81.3	80.3	80.5	74.5	72.7		76.5
<i>C. glabrata</i>	81.1	79.0	79.3	83.8	83.8	73.0	

^a Data in the lower left portion of the table indicate nucleotide sequence similarity, and data in the upper right portion indicate amino acid sequence similarity.

own supply of mt (34). If this behavior is also true for the budding yeast *Candida*, the choice of a mitochondrial gene such as the cytochrome *b* gene is reasonable. Poulter et al. (17) reported a parasexual cycle for *C. albicans*; however, it is beyond our knowledge that parasexuality has any effect on genetic recombination of mitochondrial genes. Restriction fragment length polymorphisms of mitochondrial DNA have been shown to be useful genetic markers for estimating relationships among fungi (8, 11, 28). However, studies that investigate phylogenetic relationships based on the sequences of mitochondrial DNA (13, 29) are in their infancy. The utility of the sequences of the mitochondrial cytochrome *b* gene for identifying pathogenic *Aspergillus* species has been shown by Wang et al. (30, 31). However, similar techniques have not focused on yeasts.

In this study, we partially sequenced the mitochondrial cytochrome *b* genes of *C. albicans*, *C. glabrata*, *C. parapsilosis*, *C. tropicalis*, *C. krusei*, *C. lusitaniae*, and *C. dubliniensis*. Their identification and phylogenetic relationships are discussed.

MATERIALS AND METHODS

Yeast strains and isolation of DNA. Two to five strains of each species of the most common pathogenic *Candida* species, including *C. dubliniensis*, and *Filobasidiella neoformans* were used in this study (Table 1). Total cellular DNA from these strains was extracted using a Gen Toru Kun kit (Takara Shuzo Co. Ltd., Otsu, Shiga, Japan) as recommended by the manufacturer, except that one-third of a loop of yeast cells from the YPD (1% [wt/vol] yeast extract, 2% [wt/vol] Polypeptone, 2% [wt/vol] glucose) slant was used to isolate DNA.

PCR primers and amplification of the cytochrome *b* gene. PCR primers, E1M4, E1M5, E2M4, and E2mr5 (Table 2), were designed by comparing previously published amino acid sequences for the mitochondrial cytochrome *b* genes of several organisms as described by Wang et al. (30). One microliter of extracted DNA was used to amplify the mitochondrial cytochrome *b* gene with a *TaKaRa Ex Taq* PCR amplification kit (Takara Shuzo). The reactions were performed in a final reaction mixture (50 μ l) containing 10 pmol of each primer, 4 μ l of 2.5 mM each deoxynucleoside triphosphate (dATP, dCTP, dGTP, and dTTP), 2.0 U of *TaKaRa Ex Taq* polymerase, and 5 μ l of 10 \times reaction buffer (Takara Shuzo). The amplification reactions were performed with the following cycling parameters: 94°C for 2 min; 30 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 50°C, and extension for 1 min at 72°C; and a final extension at 72°C for 10 min.

Sequencing. PCR products were purified using a QIAquick PCR purification kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Then, both strands of the PCR products were sequenced directly on an ABI prism 377 or 310 DNA sequencer using a Big Dye terminator cycle sequencing ready reaction kit (Applied Biosystems Japan Co. Ltd., Tokyo, Japan) as recommended by the kit manufacturer. From the DNA sequences, cytochrome *b* amino acid sequences were deduced using the yeast (*Saccharomyces*) mitochondrial genetic code.

Molecular phylogenetic analysis. DNA and amino acid sequences were aligned using GENETYX-MAC genetic information processing software (Software Development Co., Ltd., Tokyo, Japan). Sequences were analyzed by the

<i>C. albicans</i>	1	AATTTATTATCAGCTATACCTTTCATAGGTAATGATATCGTACCATTATATGAGGAGGT	60
<i>C. dubliniensis</i>	1	...C.....T.....G.....A.....C.....	60
<i>C. glabrata</i>	1C.....T.....G.....C.....A.....T.....A..... <u>GAT</u>	60
<i>C. krusei</i>	1T.....T..... <u>GC</u>T.....	60
<i>C. lusitaniae</i>	1	..C.....A.....T.....A.....	60
<i>C. parapsilosis</i>	1	..C.....T.....A.....T.....A.....T.....T.....G.....T.....	60
<i>C. tropicalis</i>	1	..CC..C..G...C..T..A.....T.....C.....T.....C.....T.....T...	60
<i>F. neoformans</i>	1	...A..GC...T...T...GG..T...CA..CT..TAC..AG..G..T..G..T..A	60
<i>C. albicans</i>	61	TTCTCCGTTAGTAACCCCTACTATACAACGGTTCCTTGCATTGCATTTCTTATTACCTTTC	120
<i>C. dubliniensis</i>	61T..A....T....A....G..I.....A.....T	120
<i>C. glabrata</i>	61A..ATC...T....T...A..A.....A...AT...G...A..T	120
<i>C. krusei</i>	61	..T..A..A....T..A..A.....A..A..T.....A...AT.....	120
<i>C. lusitaniae</i>	61T..ATCA..T...A....A..A..T.....A...AT.....A..T	120
<i>C. parapsilosis</i>	61	..AGT..A.....A..T...A..A..... <u>IC.T</u> ..C.....A...	120
<i>C. tropicalis</i>	61A..A.....C.....C.....A..C...C..TC...A...	120
<i>F. neoformans</i>	61T...A...TG..C..AC..TA..C..A.....T..C..T..C..ATC..C...A...	120
<i>C. albicans</i>	121	ATACTTGCTGCATTAGTATGTATGCACTTGATGGCCTTACATGTACATGGTTTCATCTAAC	180
<i>C. dubliniensis</i>	121	..T..A..A.....A.....A.....A.....A...	180
<i>C. glabrata</i>	121	..T <u>A</u>T.....GTA.....T..T.....I.....T	180
<i>C. krusei</i>	121	..T..A.....TATAT..A..T..A..A.....A.....T	180
<i>C. lusitaniae</i>	121	..TT..A..A.....GTA.....T..T..A..A.....ACT....A.....T	180
<i>C. parapsilosis</i>	121	..TT..A.....T.....T..A.....A.....TA.....	180
<i>C. tropicalis</i>	121	..CT..A.....CC..T.....T..A.....A.....A.....A.....	180
<i>F. neoformans</i>	121	..T..A..A..TC...CTCTAG..A..A..C..AA..AC...CAC.....AGC..A...	180
<i>C. albicans</i>	181	CCTGTAGGTATTACTGGTAATATTGATAGATTGCCAATGCATCCTTACTTCATATTTAAA	240
<i>C. dubliniensis</i>	181A..A.....A.....A.....T.....	240
<i>C. glabrata</i>	181	..T.....A.....G.....A..GG.....GG...T.....T.....	240
<i>C. krusei</i>	181	..A.....A.....G.....AAGT.....T..T.....	240
<i>C. lusitaniae</i>	181A..AT..AA.....I..A.....A..A.....AG.....T.....	240
<i>C. parapsilosis</i>	181	..AT.....A.....CG.....A.....T..T.....	240
<i>C. tropicalis</i>	181T.....C..A.....C..C..CC...A.....C..C...	240
<i>F. neoformans</i>	181A.....GCTCA...GCA..A..AGGCA.....C..A.....T..C.....G	240
<i>C. albicans</i>	241	GACTTAATTACTGTCTTTGTATTCTTATTAATATTTAGTTTATTGATTCTATTACCT	300
<i>C. dubliniensis</i>	241	..TC..T..A.....A.....T.....I.....	300
<i>C. glabrata</i>	241	..T.....T.....A..TT..C..CTCA.....T.....TC.....	300
<i>C. krusei</i>	241	..TGCT..A...A..A.....A..T...CT...C..T.....T.....	300
<i>C. lusitaniae</i>	241	..T.....A.....TA.....G.....T.....TC.....	300
<i>C. parapsilosis</i>	241	..T...G..A...A.....G.....T.....I..A	300
<i>C. tropicalis</i>	241	..TC..G..A..A.....C.....A..CC..T...C..CC..G..T..G.....AGC...	300
<i>F. neoformans</i>	241	..TC..G..A..AA..TA..CC..T.....CA..TGCTC..A..CAGCTC..A..A..G...G.....	300
<i>C. albicans</i>	301	AATACATTAGGACATCCTGATAACTATATACCAGGTAACCCCTATGGTAACACCTCCTTCA	360
<i>C. dubliniensis</i>	301T.....T.....A.....A.....G.....A...	360
<i>C. glabrata</i>	301T.....T.....T.....T.....T.....A.....AG..A..T	360
<i>C. krusei</i>	301T.....T.....T.....T.....T.....T.....A..T..T...G...T	360
<i>C. lusitaniae</i>	301T.....T.....T.....T.....T.....G..A..T	360
<i>C. parapsilosis</i>	301T.....C.....T.....T.....T.....A..T..T.....T	360
<i>C. tropicalis</i>	301	..C..GC.....C..A.....C..C.....	360
<i>F. neoformans</i>	301	..C..IGC.....CAG...C.....C..T..C.....A...CA.....A.....G	360
<i>C. albicans</i>	361	ATTGTACCAGAATGATACTTATTACCATTCTACGCA	396
<i>C. dubliniensis</i>	361TC.....	396
<i>C. glabrata</i>	361T.....T.....I..T..T	396
<i>C. krusei</i>	361T..T.....T.....T.....T..T	396
<i>C. lusitaniae</i>	361	..A.....T.....T.....T..T	396
<i>C. parapsilosis</i>	361T.....G...TC.....T.....T..T	396
<i>C. tropicalis</i>	361	..C.....T..G.....C.....	396
<i>F. neoformans</i>	361G.....G..TC..TC.....	396

FIG. 1. Comparison of nucleotide sequences of cytochrome *b* genes of the most common pathogenic *Candida* species (only type cultures) and *F. neoformans*. Dots indicate nucleotides identical to those in *C. albicans*. Species-specific nucleotides are underlined.

<i>C. albicans</i>	1	NLLSAMPFMGNDIVPFMWGGFSVSNPTMQRFFALHFLLPFMTAALVCMHLMALHVHGSSN	60
<i>C. dubliniensis</i>	1	.T.....I..M.....L.....N....	60
<i>C. glabrata</i>	1	. <u>F</u> . <u>I</u> . <u>V</u> . <u>Q</u> .. <u>QWL</u>I.....Y.V.. <u>II</u> ...V..F.....	60
<i>C. krusei</i>	1I. <u>A</u>Y.....L.. <u>ML</u>	60
<i>C. lusitaniae</i>	1M.....Y... <u>IL</u> ...V..F...T....	60
<i>C. parapsilosis</i>	1I. <u>K</u>I.....I...T..... <u>IL</u>N....	60
<i>C. tropicalis</i>	1	.TT..I..I.....I.....I.....TT.. <u>IL</u> ..T.....N....	60
<i>F. neoformans</i>	1	. <u>MT</u> ..I.. <u>WI</u> ..T.. <u>FTQ</u> ..V.....N.A. <u>TN</u> .. <u>ST</u> ..YTT..I...TATV.. <u>MITT</u> ..T....	60
<i>C. albicans</i>	61	PVGITGNIDRLPMHPYFMFKDLITVFVFLMFSLFVFPSPNTLGHDPNYMGPMPVTPPS	120
<i>C. dubliniensis</i>	61	..M...M.....TM.....	120
<i>C. glabrata</i>	61	.L....M.. <u>IG</u> .. <u>G</u> ..I..... <u>IF</u>F.....I...L..A.	120
<i>C. krusei</i>	61	.M.M...M.. <u>S</u> <u>AM</u> ..M.. <u>ML</u> .. <u>CF</u> ...F.....I...L..A.	120
<i>C. lusitaniae</i>	61	..MSS.L..M.. <u>S</u>M.V.....F.....I...L..A.	120
<i>C. parapsilosis</i>	61	.L....V.....I...V.....V.....I...L....	120
<i>C. tropicalis</i>	61I...TV..... <u>II</u> ..I.....T.....I.....	120
<i>F. neoformans</i>	61	.E..SS. <u>AEKA</u>I...TV.. <u>III</u> .. <u>FIATTAT</u> ..M.A.. <u>MT</u> .. <u>S</u> ...I.A...Q....	120
<i>C. albicans</i>	121	IVPEWYLLPFYA	132
<i>C. dubliniensis</i>	121T.....	132
<i>C. glabrata</i>	121	132
<i>C. krusei</i>	121	132
<i>C. lusitaniae</i>	121	<u>M</u>	132
<i>C. parapsilosis</i>	121T.....	132
<i>C. tropicalis</i>	121T.....	132
<i>F. neoformans</i>	121 <u>TT</u>	132

FIG. 2. Comparison of deduced amino acid sequences for cytochrome *b* genes of the most common pathogenic *Candida* species (only type cultures) and *F. neoformans*. Dots indicate amino acids identical to those in *C. albicans*. Species-specific amino acids are underlined.

unweighted pair-group method with arithmetic mean (UPGMA) and neighbor joining (NJ) using GENETYX-MAC genetic information processing software or by maximum parsimony (MP) using the package Phylogenetic Analysis Using Parsimony, version 4.0b4a, for Macintosh (26). For the NJ analyses, the distance between the sequences was calculated using Kimura's two-parameter model (10).

Nucleotide sequence accession numbers. Mitochondrial cytochrome *b* genes of *Candida* species partially sequenced in this study have been deposited in the DDBJ/EMBL/GenBank data library under the accession numbers shown in Table 1.

RESULTS

The mitochondrial cytochrome *b* genes of major pathogenic *Candida* species (Table 1), including *C. dubliniensis*, were amplified by PCR, and a 396-bp segment corresponding to positions 445 to 840 in the *C. glabrata* cytochrome *b* coding sequence (GenBank accession no. X53862) (2) was sequenced. Strains of *C. albicans*, *C. glabrata*, and *C. tropicalis* showed intraspecies variation (Table 3). Intraspecies variation was the highest for *C. tropicalis* (1.8%) and was divided into two DNA types.

Although the intraspecies variations of these pathogenic yeasts were not higher than 1.8%, the interspecies variations were high. Table 4 (lower left portion) represents the pairwise nucleotide identities of these yeasts calculated from the nucleotide sequences of the cytochrome *b* genes. No pair of species represented more than 90% identical sequences. Even *C. albicans* and *C. dubliniensis*, the closest species, according to other DNA sequences, had differences in 40 nucleotides of the total of 396 nucleotides (89.9% identical).

The *Saccharomyces* mitochondrial genetic code system was used to deduce amino acid sequences from cytochrome *b* gene sequences. In the yeast mitochondrial gene, UGA encodes tryptophan instead of polypeptide chain termination, AUA encodes methionine instead of isoleucine, and CUN encodes threonine instead of the leucine found in the universal codon

system. *C. glabrata* has conserved the strong codon bias of *S. cerevisiae* (1); however, the CUN codon family codes for a leucine in *C. parapsilosis* mt (7). The reason for using the *Saccharomyces* mitochondrial genetic code system in the present study is that the system is accepted for yeasts by the DDBJ/EMBL/GenBank data library. Table 4 (upper right portion) represents the pairwise amino acid identities among different species. Based on amino acid sequences, except for *C. albicans* and *C. dubliniensis* (92.4% identical), no other pair had more than 90% identity. There were differences in 10 amino acids of the total of 132 amino acids between *C. albicans* and *C. dubliniensis*.

Multiple alignments of the nucleotide (Fig. 1) and amino acid (Fig. 2) sequences showed that individual species possessed characteristic sequences. The major pathogenic *Candida* species were positioned distinctly by any one of the phylogenetic trees constructed using UPGMA (Fig. 3), NJ (Fig. 4), or MP (Fig. 5) with basidiomycetous yeast *F. neoformans* as an outgroup. Although there was some difference in the degree of resolution when nucleotide and amino acid sequences were used, the topologies obtained by both sequence types and tree estimation algorithms used were consistent.

DISCUSSION

This study reveals that the cytochrome *b* gene sequences of the major pathogenic *Candida* species differ from each other by more than 10%. Their amino acid sequences also show considerable differences. Except for *C. albicans* and *C. dubliniensis* (92.4% identical), no other pair showed more than 90% identity based on amino acid sequences. These results show the higher divergence of mitochondrial cytochrome *b* genes than of nuclear genes and its usefulness for investigating relationships of closely related species.

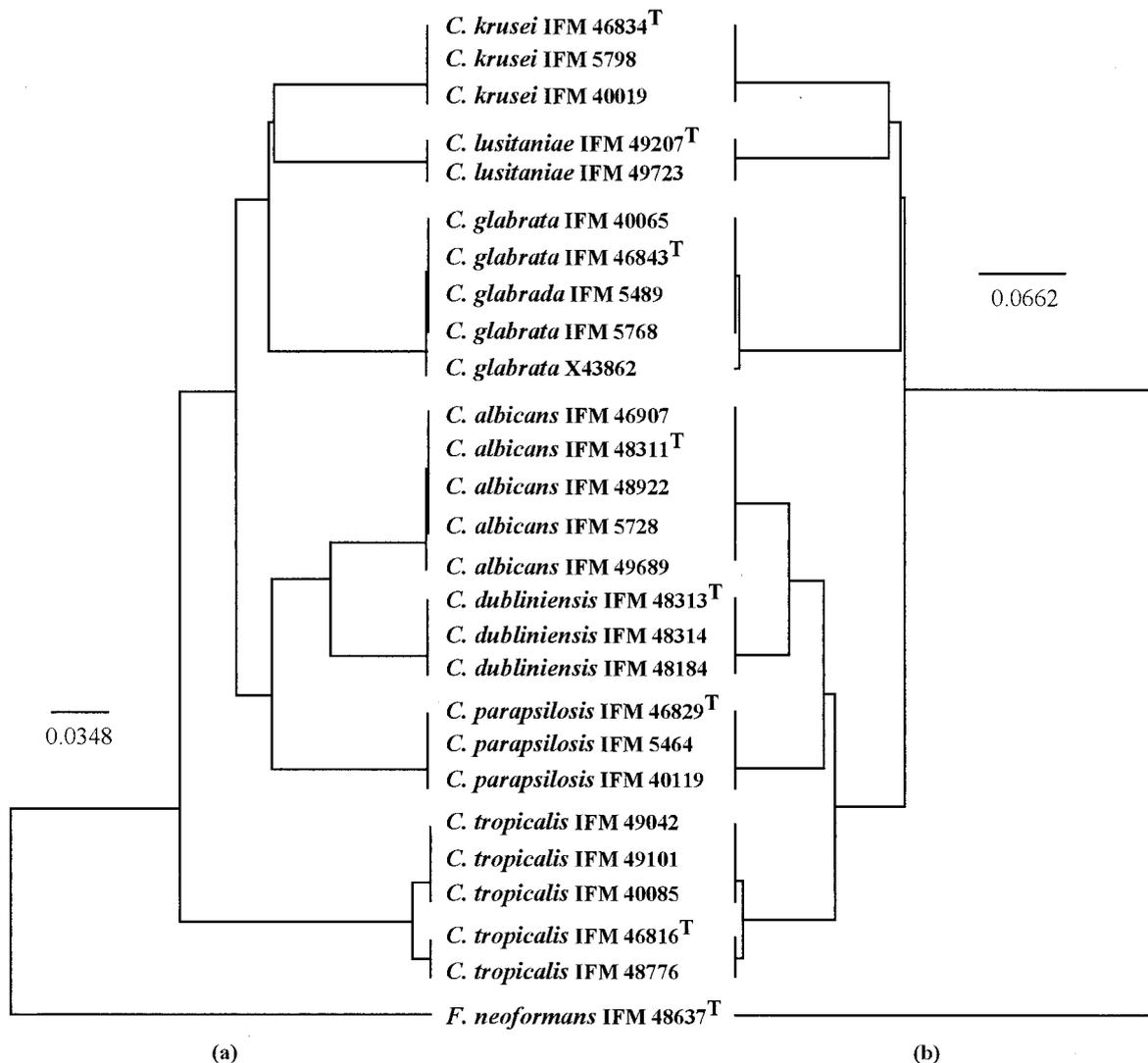


FIG. 3. UPGMA trees of the most common *Candida* species, based on nucleotide (a) and deduced amino acid (b) sequences for the cytochrome *b* genes. The basidiomycetous yeast *F. neoformans* represents an outgroup. Bars represent the numbers of nucleotide and amino acid substitutions per nucleotide and amino acid sites.

The sequence identity between *C. dubliniensis* and *C. albicans* is 89.9% based on the nucleotide sequences of the mitochondrial cytochrome *b* genes. However, the previously reported percentages of similarity between the organisms were 97.9% for the *ACT1* gene (exon) (4), 97.52 to 97.75% for the V3 variable region of the large-subunit rRNA gene (23, 25), and 98.6% for the small-subunit rRNA gene (6). The predicted *C. dubliniensis* *ACT1* protein sequence was identical to that of *C. albicans*, apart from a single conservative substitution, isoleucine to valine (4). On the other hand, the two organisms had 7.6% differences in the predicted protein sequences of the cytochrome *b* genes.

Wang et al. (30, 31) have shown that UPGMA is the best method for constructing phylogenetic trees based on cytochrome *b* sequences of fungi. In addition to UPGMA, sequences were analyzed using NJ and MP. There was very little difference among UPGMA (Fig. 3), NJ (Fig. 4), and MP (Fig. 5) phylogenetic trees based on the nucleotide or amino acid sequences for the cytochrome *b* genes of the most common

pathogenic *Candida* species. This difference seems to be due to bootstrap values and differences in the methods of analysis. Moreover, different nucleotide codons may give rise to the same amino acid, a result which may produce differences in nucleotide- and amino acid-based phylogenetic trees. Although there was some difference in the degree of resolution when nucleotide and amino acid sequences were used, the topologies obtained with both sequence types and the tree estimation algorithms used were mutually consistent. The phylogenetic relationships of these pathogenic yeasts based on cytochrome *b* sequences were not 100% congruent with those based on rRNA; in addition, in the latter case, the relationships varied depending on the part of the sequences used to draw the phylograms, large or small-subunit rRNA (6, 23, 25).

This study represents the first phylogenetic investigation of the most common *Candida* species based on the sequences of the mitochondrial cytochrome *b* genes. The separation of the most common pathogenic *Candida* species, including the recently described species, *C. dubliniensis*, under investigation,

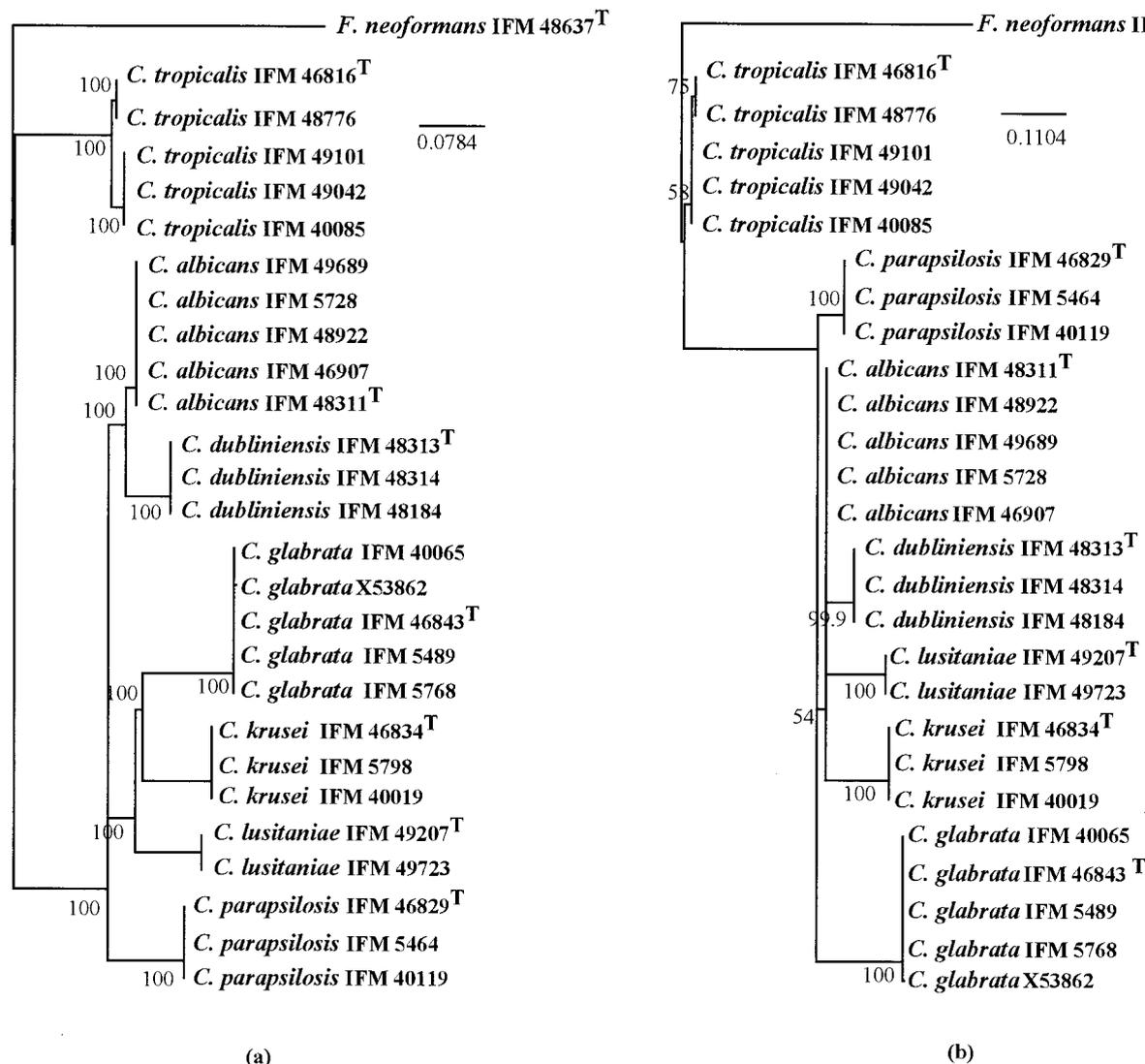


FIG. 4. NJ trees of the most common *Candida* species, based on nucleotide (a) and deduced amino acid (b) sequences for the cytochrome *b* genes. The basidiomycetous yeast *F. neoformans* represents an outgroup. The numerals given on the branches represent the confidence levels from 1,000 replicate bootstrap samplings (values lower than 50% are not shown). Bars represent the numbers of nucleotide and amino acid substitutions per nucleotide and amino acid sites, calculated using Kimura's two-parameter model (10).

was evident irrespective of the sequences, DNA or amino acid, and the methods of phylogenetic tree construction, UPGMA, NJ, or MP. The results supported the unique species designation of *C. dubliniensis* on the basis of its distinct phylogenetic position (supported by bootstrap analyses) and on the basis of the differences between its cytochrome *b* nucleotide and amino acid sequences and the corresponding sequences for *C. albicans*, 10.1 and 7.6%, respectively.

In a recent study, mitochondrial *COX2* sequences revealed 2.1% intraspecies variability of *C. glabrata* isolates representing two types (22). These types correlated with the geographical origins of the strains, 90% of U.S. isolates belonging to type 1 and 82% of Brazilian isolates belonging to type 2. However, mitochondrial cytochrome *b* genes are more conserved at the intraspecies level (although the numbers of strains examined are limited). In this study, strains of the same species had identical sequences, except for *C. albicans*, *C. glabrata*, and *C.*

tropicalis (Table 3). *C. tropicalis* had the highest variation at the intraspecies level—1.8%.

Rapid and accurate identification of pathogenic organisms is important clinically and epidemiologically (12, 18). Molecular methods provide accurate and rapid identification. Although a variety of methods, ranging from randomly amplified polymorphic DNA analysis to PCR-enzyme immunoassay, are used for DNA subtyping of *Candida* species, no "gold standard" exists (19). With an automated DNA sequencer, a DNA sequence can be obtained quite rapidly. Starting from DNA extraction from yeast cells, the sequence of a cytochrome *b* segment (396 bp) can be determined within a working day, and suspected *Candida* species can be identified correctly using a database.

This is the first report of the identification and phylogenetic analysis of the most common pathogenic *Candida* species by use of cytochrome *b* gene sequences. The 396-bp region of the cytochrome *b* gene examined in the present study has proved

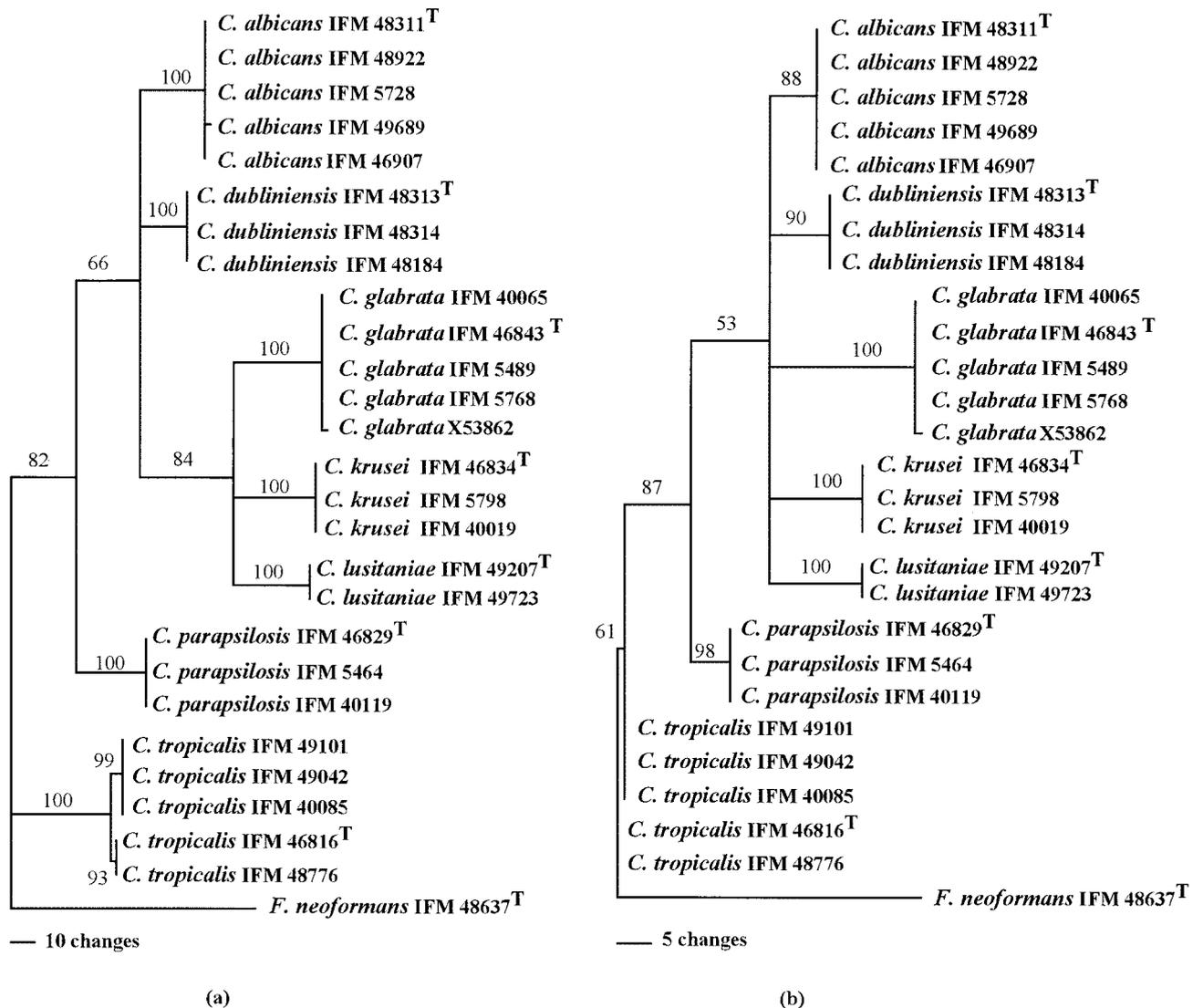


FIG. 5. Phylogenetic trees of the most common *Candida* species, based on nucleotide (a) and deduced amino acid (b) sequences for the cytochrome *b* genes. The trees were constructed by MP analysis (heuristic search, stepwise addition, tree-bisection-reconnection), where the basidiomycetous yeast *F. neoformans* represents an outgroup. Branch lengths are proportional to the numbers of nucleotide and amino acid changes, and the numerals given on the branches are the frequencies (>50%) with which a given branch appeared in 100 bootstrap replications.

effective for identification and for studying phylogenetic relationships. High interspecies and low intraspecies divergences of cytochrome *b* gene sequences are attractive and may facilitate the design of probes for specific and rapid identification of pathogenic *Candida* species.

ACKNOWLEDGMENT

We thank the Ministry of Education, Tokyo, Japan, for providing a scholarship to S. K. Biswas during the study.

REFERENCES

- Ainley, W. M., I. G. Macreadie, and R. A. Butow. 1985. *var1* gene on the mitochondrial genome of *Torulopsis glabrata*. *J. Mol. Biol.* **184**:565-576.
- Clark-Walker, G. D. 1991. Contrasting mutation rates in mitochondrial and nuclear genes of yeasts versus mammals. *Curr. Genet.* **20**:195-198.
- Coleman, D. C., M. G. Rinaldi, K. A. Haynes, J. H. Rex, R. C. Summerbell, E. J. Anaissie, A. Li, and D. J. Sullivan. 1998. Importance of *Candida* species other than *Candida albicans* as opportunistic pathogens. *Med. Mycol.* **36**(Suppl. 1):156-165.
- Donnelly, S. M., D. J. Sullivan, D. B. Shanley, and D. C. Coleman. 1999. Phylogenetic analysis and rapid identification of *Candida dubliniensis* based on analysis of ACT1 intron and exon sequences. *Microbiology* **145**:1871-1882.
- Fridkin, S. K., and W. R. Jarvis. 1996. Epidemiology of nosocomial fungal infections. *Clin. Microbiol. Rev.* **9**:499-511.
- Gillfillan, G. D., D. J. Sullivan, K. Haynes, T. Parkinson, D. C. Coleman, and N. A. Gow. 1998. *Candida dubliniensis*: phylogeny and putative virulence factors. *Microbiology* **144**:829-838.
- Guclin, E., M. Guerin, and J. Velours. 1991. Isolation of the ATP synthase subunit 6 and sequence of the mitochondrial *ATP6* gene of the yeast *Candida parapsilosis*. *Eur. J. Biochem.* **197**:105-111.
- Hamari, Z., F. Kevei, E. Kovacs, J. Varga, Z. Kozakiewicz, and J. H. Croft. 1997. Molecular and phenotypic characterization of *Aspergillus japonicus* and *Aspergillus aculeatus* strains with special regard to their mitochondrial DNA polymorphisms. *Antonie Leeuwenhoek* **72**:337-347.
- Hazen, K. C. 1995. New and emerging yeast pathogens. *Clin. Microbiol. Rev.* **8**:462-478.
- Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111-120.
- Kozlowski, M., and P. P. Stepień. 1982. Restriction enzyme analysis of mitochondrial DNA of members of the genus *Aspergillus* as an aid in taxonomy. *J. Gen. Microbiol.* **128**:471-476.

12. Nguyen, M. H., J. E. Peacock, Jr., A. J. Morris, D. C. Tanner, M. L. Nguyen, D. R. Snyderman, M. M. Wagener, M. G. Rinaldi, and V. L. Yu. 1996. The changing face of candidemia: emergence of non-*Candida albicans* species and antifungal resistance. *Am. J. Med.* **100**:617–623.
13. Paquin, B., L. Forget, I. Roewer, and B. F. Lang. 1995. Molecular phylogeny of *Allomyces macrogynus*: congruency between nuclear ribosomal RNA- and mitochondrial protein-based trees. *J. Mol. Evol.* **41**:657–665.
14. Perfect, J. R., and W. A. Schell. 1996. The new fungal opportunists are coming. *Clin. Infect. Dis.* **22**(Suppl. 2):112–118.
15. Pfaller, M. A. 1995. Nosocomial fungal infections: epidemiology of candidiasis. *J. Hosp. Infect.* **30**(Suppl.):329–338.
16. 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.
17. Poulter, R., K. Jeffery, M. J. Hubbard, M. G. Shepherd, and P. A. Sullivan. 1981. Parasexual genetic analysis of *Candida albicans* by spheroplast fusion. *J. Bacteriol.* **146**:833–840.
18. Price, M. F., M. T. LaRocco, and L. O. Gentry. 1994. Fluconazole susceptibilities of *Candida* species and distribution of species recovered from blood cultures over a 5-year period. *Antimicrob. Agents Chemother.* **38**:1422–1427.
19. Reiss, E., K. Tanaka, G. Bruker, V. Chazalet, D. Coleman, J. P. Debeaupuis, R. Hanazawa, J. P. Latge, J. Lortholary, K. Makimura, C. J. Morrison, S. Y. Murayama, S. Naoe, S. Paris, J. Sarfati, K. Shibuya, D. Sullivan, K. Uchida, and H. Yamaguchi. 1998. Molecular diagnosis and epidemiology of fungal infections. *Med. Mycol.* **36**(Suppl. 1):249–257.
20. Rex, J. H., M. G. Rinaldi, and M. A. Pfaller. 1995. Resistance of *Candida* species to fluconazole. *Antimicrob. Agents Chemother.* **39**:1–8.
21. Rex, J. H., M. A. Pfaller, J. N. Galgiani, M. S. Bartlett, A. Espinel-Ingroff, M. A. Ghannoum, M. Lancaster, F. C. Odds, M. G. Rinaldi, T. J. Walsh, and A. L. Barry. 1997. Development of interpretive breakpoints for antifungal susceptibility testing: conceptual framework and analysis of in vitro-in vivo correlation data for fluconazole, itraconazole, and *Candida* infections. Subcommittee on Antifungal Susceptibility Testing of the National Committee for Clinical Laboratory Standards. *Clin. Infect. Dis.* **24**:235–247.
22. Sanson, G. F., and M. R. Briones. 2000. Typing of *Candida glabrata* in clinical isolates by comparative sequence analysis of the cytochrome *c* oxidase subunit 2 gene distinguishes two clusters of strains associated with geographical sequence polymorphisms. *J. Clin. Microbiol.* **38**:227–235.
23. Sullivan, D. J., T. J. Westerneng, K. A. Haynes, D. E. Bennet, and D. C. Coleman. 1995. *Candida dubliniensis* sp. nov.: phenotypic and molecular characterisation of a novel species associated with oral candidosis in HIV-infected individuals. *Microbiology* **141**:1507–1521.
24. Sullivan, D. J., M. C. Henman, G. P. Moran, L. C. O'Neill, D. E. Bennett, D. B. Shanley, and D. C. Coleman. 1996. Molecular genetic approaches to identification, epidemiology and taxonomy of non-*albicans* *Candida* species. *J. Med. Microbiol.* **44**:399–408.
25. Sullivan, D. J., K. Haynes, J. Bille, P. Boerlin, L. Rodero, S. Lloyd, M. Henman, and D. Coleman. 1997. Widespread geographic distribution of oral *Candida dubliniensis* strains in human immunodeficiency virus-infected individuals. *J. Clin. Microbiol.* **35**:960–964.
26. Swofford, D. L. 1998. PAUP*. Phylogenetic analysis using parsimony (and other methods). Version 4.0b4a for Macintosh. Sinauer Associates, Sunderland, Mass.
27. Vanden Bossche, H., F. Dromer, I. Improvisi, M. Lozano-Chiu, J. H. Rex, and D. Sanglard. 1998. Antifungal drug resistance in pathogenic fungi. *Med. Mycol.* **36**(Suppl. 1):119–128.
28. Varga, J., F. Kevei, A. Vriesema, F. Debets, Z. Kozakiewicz, and J. H. Croft. 1994. Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the *Aspergillus niger* aggregate. *Can. J. Microbiol.* **40**:612–621.
29. Walker, D. J., A. E. Wakefield, M. N. Dohn, R. F. Miller, R. P. Baughman, P. A. Hossler, M. S. Bartlett, J. W. Smith, P. Kazanjian, and S. R. Meshnick. 1998. Sequence polymorphisms in the *Pneumocystis carinii* cytochrome *b* gene and their association with atovaquone prophylaxis failure. *J. Infect. Dis.* **178**:1767–1775.
30. Wang, L., K. Yokoyama, M. Miyaji, and K. Nishimura. 1998. The identification and phylogenetic relationship of pathogenic species of *Aspergillus* based on the mitochondrial cytochrome *b* gene. *Med. Mycol.* **36**:153–164.
31. Wang, L., K. Yokoyama, M. Miyaji, and K. Nishimura. 2000. Mitochondrial cytochrome *b* gene analysis of *Aspergillus fumigatus* and related species. *J. Clin. Microbiol.* **38**:1352–1358.
32. Wingard, J. R. 1995. Importance of *Candida* species other than *C. albicans* as pathogens in oncology patients. *Clin. Infect. Dis.* **20**:115–125.
33. Witek-Janusek, L., C. Cusack, and H. L. Mathews. 1998. *Candida albicans*: an opportunistic threat to critically ill low birth weight infants. *Dimens. Crit. Care Nurs.* **17**:243–255.
34. Yang, H. C., A. Palazzo, T. C. Swayne, and L. A. Pon. 1999. A retention mechanism for distribution of mitochondria during cell division in budding yeast. *Curr. Biol.* **9**:1111–1114.