Full Paper

Phylogenetic identification of *n*-alkane assimilating *Candida* yeasts based on nucleotide divergence in the 5' end of LSU rDNA gene

Mami Arie,* Hitoshi Matsuda, Keizo Furuhashi, and Masamichi Takagi¹

Pharmaceutical and Biotechnology Laboratory, Japan Energy Corporation, Toda 335-8502, Japan

¹Tokyo University of Agriculture, Setagaya-ku, Tokyo 156-8502, Japan, and

Niigata College of Pharmacy, Niigata City, Niigata 950-2081, Japan

(Received August 4, 2000; Accepted October 20, 2000)

Phylogenetic relationships of several species within the *n*-alkane assimilating *Candida* yeasts were investigated by using characters from the nucleotide sequence of the variable D1/D2 region at the 5' end of a large-subunit (26S) ribosomal DNA (rDNA) gene. First the nucleotide sequences of D1/D2 domain of *Candida* sp. 1098 (formerly identified as *C. tropicalis* 1098) and its dicarboxylic acid-producing-mutant strain M1210 were investigated. These two nucleotide sequences were identical and lacked only one base pair compared with that of *C. maltosa* CBS 5611 (type strain), and they were identified as *C. maltosa*. We then showed that *C. maltosa* IFO 1978 (formerly identified as *C. cloacae*) and *C. maltosa* IFO 1975 (formerly identified as *C. subtropicalis*) had the same nucleotide sequence and had only one base pair substitution compared with *C. maltosa* CBS 5611 (type strain), which is consistent with conventional classification. We also found that another widely studied *n*-alkane assimilating *Candida* yeast, *C. tropicalis* pk233, to be *C. viswanathii*.

Key Words—Ascomycetous yeasts; *Candida maltosa*; *Candida tropicalis*; *Candida viswanathii*; *n*-alkane; phylogeny; ribosomal DNA

Introduction

Dicarboxylic acids (DCAs) are major chemical intermediates in many industrial fields, such as perfume production and polymer synthesis. Brassylic acid (1,13-tridecanedioic acid), the most excellent and valuable base material for musk perfumes, is only commercially produced by fermentation (Uemura, 1985).

To obtain DCA-producing microorganisms, we isolated yeast strain 1098 from the soil of a petroleum refinery field in the 1960s. Strain 1098 converted *n*-tride-

cane into DCA at 1.4 g/L when cultured in a flask. Consecutive mutagenesis treatment gave us a mutant strain M1210, which produces DCA at 20 g/L, and this strain is now used for the commercial production of brassylic acid in our fermentation plant. Strain 1098 was identified as *Candida tropicalis* by Suzuki et al. (1975) from the results of a set of fermentation and assimilation tests and was deposited as *Candida tropicalis* FERM-P3291 (Patent Microorganism Depository, National Institute of Bioscience and Human Technology, Tsukuba, Ibaraki, Japan).

Recently, the *URA3* gene, which encodes orotidine-5'-phosphate decarboxylase (ODCase) of *Candida* sp. 1098, was cloned (Hara et al., 1999), and its deduced amino acid sequence (Acc. AB023638) was much

* Address reprint requests to: Ms. Mami Arie, Pharmaceutical and Biotechnology Laboratory, Japan Energy Corporation, 3-17-35 Niizo-Minami, Toda 335-8502, Japan.

E-mail: marie@j-energy.co.jp

more similar to that of *C. maltosa* (98.5%; Ohkuma et al., 1993) than that of *C. tropicalis* pk233 (82.1%; Kanayama et al., 1998). Moreover, we also found that the nucleotide sequence of the cytochrome P450 gene of *Candida* sp. 1098 is almost homologous to that of *C. maltosa* (unpublished data). From these results, we had reasonable doubt that our strain was *C. maltosa* instead of *C. tropicalis*.

Taxonomic classification of the genus *Candida* is usually based on cellular morphology and distinctive reactions in a standardized set of fermentation and assimilation tests (Meyer et al., 1998). These phenotypic characteristics, however, are often strain specific and inadequate for the recognition of species (Kurtzman and Robnett, 1995). For example, *C. maltosa* and *C. tropicalis* both belong to the *C. albicans* clade and are known to be physiologically very similar, but they can be distinguished by the soluble starch assimilation test (Meyer et al., 1975). *C. tropicalis* assimilates soluble starch and *C. maltosa* doesn't. A soluble starch assimilation test is sometimes ambiguous; even *C. maltosa* AJ4718 (CBS5611), a type strain of *C. maltosa*, was once reported to assimilate soluble starch slowly (Fukazawa et al., 1975).

Recently, molecular comparisons are leading to an understanding of phylogenetic relationships among yeasts and offer the opportunity for rapid and accurate identification of the species in the genus *Candida* (Kurtzman and Robnett, 1997, 1998). Some investigators have proposed identification of *Candida* spp. based on an analysis of restriction fragment length polymorphisms (RFLPs) of ribosomal DNA (rDNA) (Magee et al., 1987; Williams et al., 1995) and on oligonucleotide-based molecular probes that hybridize with species-specific RNA and DNA sequences (Fell, 1995). However, few species have been compared, and a successful application of molecular methods for identification requires prior characterization of all known species in the phylogenetic group of interest and an understanding of the extent of intraspecific nucleotide differences. Kurtzman and Robnett (1997, 1998) determined the sequences for the 600-nucleotide variable region (D1/D2) at the 5' end of the large subunit (LSU) rDNA gene for all recognized *Candida* spp. and concluded that this rDNA region is sufficiently variable to allow a reliable separation of all known *Candida* spp., including *C. maltosa* and *C. tropicalis*, and that within the species there are at most 2 to 3 nucleotide differences.

In this report, nucleotide sequences were determined and compared for the 600-nucleotide variable region (D1/D2) at the 5' end of the LSU rDNA gene for *Candida* sp. strains 1098 and M1210 and three other alkane-assimilating *Candida* spp. Furthermore, to determine the result of the molecular taxonomy, we carried out a set of fermentation and assimilation tests.

Materials and Methods

Strains and media. *Candida* spp. strains used in this study are listed in Table 1. The media used in this study were YPD (1% [w/v] yeast extract, 2% Bacto-peptone [Difco, Detroit, USA], 2% glucose and 2% Bacto-agar [Difco]) and YNB (0.67% Yeast Nitrogen Base [Difco] and 2% Bacto-agar [Difco]) supplied with appropriate carbon sources.

A set of fermentation and assimilation tests was performed according to the method described by Yarrow (1998).

rDNA sequence analysis. Genomic DNA of *Candida* spp. was prepared from cells grown for 16 h at 30°C in 10 ml of YPD medium at 300 rpm. Genomic DNA isolation for PCR was performed by the method described by Hoffman and Winston (1987).

The D1/D2 domain of the LSU rDNA gene was amplified with primers NL-1 (5'-GCATATCAATAAGCG-GAGGAAAAG-3') and NL-4 (5'-GGTCCGTGTTTCAA-GACGG-3') (O'Donnell and Gray, 1995). Amplification was performed for 36 PCR cycles with annealing at 52°C for 30 s, extension at 74°C for 1 min, and denaturation at 98°C for 30 s in TaKaRa PCR Thermal Cycler MP (TaKaRa, Shiga, Japan). The amplified DNA was electrophoresed in a 1% agarose gel, and the designated band was purified by using Qiagen Gel Extraction Kit (Qiagen, Hilden, Germany) as prescribed by the manufacturer.

Both strands of the rDNA regions from each strain were determined by using automated fluorescent DNA sequencer (ABI377, PE Applied Biosystems, Foster City, CA, USA). Four sequencing reactions were required for each DNA sample. Primers for these reactions were the external primers NL-1 and NL-4 and the internal primers NL-2A (5'-CTTGTTTCGCTACGGTCTC-3') and NL-3A (5'-GAGACCGATAGCGAACAAG-3').

Sequence data were aligned by using a GENETYX-MAC program (Software Development, Tokyo, Japan). The nucleotide sequences reported in this paper will

Table 1. *Candida* strains used in this study.

	Strain	Original designation	Isolation source	Acc. No. ^a
<i>Candida</i> sp.	1098 ^b	<i>Candida tropicalis</i> FERM-P3291	Petroleum refinery field	AB034685
<i>Candida</i> sp.	M1210 ^b		Mutant strain of 1098	AB034686
<i>Candida maltosa</i>	IFO1977	type strain (=ATCC28140=CBS5611)	Neutralizing tank in monosodium glutamate manufacture	U45845
	IFO1978	<i>C. cloacae</i> (=ATCC20184=CBS5612)	Mud	AB034687
	IFO1975	<i>C. subtropicalis</i> (=ATCC28241=CBS6465)		AB034688
<i>Candida tropicalis</i>	IFO1400 pk233	type strain (=ATCC750=CBS94) =ATCC20336	Bronchitic patient	U45749 AB034689

ATCC, American Type Culture Collection, Manassas, VA, USA; IFO, Institute for Fermentation, Osaka, Japan; CBS, Centraalbureau voor Schimmelcultures, Delft, The Netherlands.

^a GenBank accession No. of nucleotide sequence of D1/D2 region.

^b Our lab.; identified as *C. maltosa* in this study.

appear in the DDBJ/EMBL/GenBank nucleotide sequence databases with the accession numbers shown in Table 1.

Nucleotide alignment and phylogenetic analysis. The nucleotide substitution rate was determined by using Kimura's two-parameter method (Kimura, 1980), and phylogenetic trees were constructed by the distant matrix-neighbor-joining method on a CLUSTAL W package (Thompson et al., 1994). The topology of the phylogenetic tree was built by bootstrap analysis with 1,000 bootstrap trials.

Results and Discussion

Nucleotide sequences of D1/D2 region from the LSU rDNA gene from strains 1098 and M1210 were identical despite the consecutive mutagenesis treatments that made M1210 physiologically a different strain from 1098 because it produces DCA by 14 times. Furthermore, we could find several mutations in the nucleotide sequence of the P450alk gene between strains 1098 and M1210 (unpublished data). These facts show us that the D1/D2 region is stable among different strains of the same species. Only one base pair difference at bp. 538 was found in strains 1098 and M1210 compared with that of *C. maltosa* type strain CBS5611 (Acc. U45745). This means that *Candida* spp. strains 1098 and M1210 should be identified as *Candida maltosa*.

To determine that identification based on the molecular sequencing method is consistent with the conventional classification method, we carried out a set of fer-

mentation and assimilation tests for strain 1098. The results were consistent with the results of *C. maltosa* type strain (Table 2). Although strain 1098 and even the *C. maltosa* type strain showed a slight pseudohyphae-like growth on soluble starch, this can be easily identified as negative when *C. tropicalis* is grown on the same plate (positive) (Fig. 1). From this result, we could conclude that we first misunderstood strain 1098 as *C. tropicalis* because of the result that 1098 showed a pseudohyphae spot on soluble starch (Suzuki et al., 1975). Here we conclude strains 1098 and M1210 to be *Candida maltosa*.

At that time, we sequenced the D1/D2 domain from the LSD rDNA gene of *C. maltosa* IFO1978 and IFO1975, which had been classified as *C. cloacae* (Komagata et al., 1964a, b) and *C. subtropicalis* (Nakase et al., 1972), respectively, but because of its similarity they were reclassified as *C. maltosa* by Meyer et al. in 1975). The nucleotide sequence of IFO1978 and IFO1975 were identical and showed one base pair substitution compared with the type strain of *C. maltosa* (CBS5611). This result was consistent with the conventional classification.

We also sequenced the D1/D2 domain from the LSU rDNA gene of *C. tropicalis* pk233, which has been often used for alkane assimilation studies (Kanayama et al., 1998). And the surprising result was that D1/D2 domain from the LSU rDNA gene of pk233 has the same sequence as that of *C. lodderae* CBS1924 (type strain), which is now classified as *C. viswanathii* CBS4024 (type strain) because those two species have almost the same sequence of D1/D2 domain ex-

Table 2. Physiological properties of tested strains.

		1098		pk233 ^a	Type strain ^b		
		Patent ^c	This study		<i>C. maltosa</i>	<i>C. tropicalis</i>	<i>C. viswanathii</i>
Fermentation	glucose	+	+	+	+	+	+
	galactose	+	+	+	v	+	+
	sucrose	+	+	—	+	v	v
	maltose	+	+	+	v	+	+
	lactose	—	—	—	—	—	—
	raffinose	—	—	—	—	—	—
	trehalose	+	+	+	+	+	+
	cellobiose	—	—	—			
	melibiose	—	—	—			
	melezitose	—	—	—			
Assimilation	glucose	+	+	+	+	+	+
	galactose	+	+	+	+	+	+
	L-sorbose	+	+	+	+	v	v
	sucrose	+	+	+	+	v	+
	maltose	+	+	+	+	+	+
	cellobiose	n.d.	+	+	+	+	+
	treharose	+	+	+	+	+	+
	lactose	—	—	—	—	—	—
	melibiose	—	—	—	—	—	—
	raffinose	—	—	—	—	—	—
	melezitose	+	+	—	+	v	+
	inulin	—	—	—	—	—	—
	soluble starch	±	—	+	—	+	+
	D-xylose	+	+	+	+	+	+
	L-arabinose	+	—	+	—	—	v
	D-arabinose	—	—	—	—	—	—
	D-ribose	—	—	—	—	v	v
	L-rhamnose	—	—	—	—	—	—
	ethanol	+	+	+	+	+	+
	glycerol	+	+	+	+	v	+
	erythritol	—	—	—	—	—	—
	ribitol	+	+	+	+	+	+
	galactitol	—	—	—	—	—	—
	D-mannitol	+	+	+	+	+	+
	D-glucitol	+	+	+	+	+	+
	α-methyl-D-glucoside	+	+	—	+	v	+
	salicin	+	+	+	+	v	+
	DL-lactic acid	+	—	—	v	v	—
	succinic acid	+	+	+	+	+	+
	citric acid	+	—	+	v	+	+
	inositol	—	—	—	—	—	—
	hexadecane	+	+	+	+	+	+
	Mol% G+C (ave.)				36.7	35.5	45.9

Symbols: +, positive; —, negative; v, variable.

^aThis study.

^bMeyer et al., 1998.

^cSuzuki et al., 1975.

cept for two nucleotide substitutions (Meyer et al., 1998). These results correlate to the GC contents of several sequences of *pk233* gene, such as *POX2*, *POX4*, and *POX5* (Okazaki et al., 1986, 1987) being 45.9–47.8%, higher than the 35.1% of *C. tropicalis* CBS94 (type strain), but almost the same as 46.3% of *C. viswanathii* CBS4024 (type strain) (Meyer et al., 1998).

We then carried out a phylogenetic analysis on the nucleotide sequence of D1/D2 domain from the LSU rDNA and constructed a phylogenetic tree (Fig. 2).

Fig. 1. Results of the soluble starch assimilation test. a, *C. maltosa* IFO1977 (type strain); b, strain 1098; c, *C. tropicalis* IFO1400 (type strain).

This figure clearly shows that strains 1098 and M1210 are *Candida maltosa*. And the overall structure is similar to evolutionary trees based on the 18S rRNA sequences showed by Ohkuma et al. (1993a, b). But as our result contains all known species in the *Candida albicans* clade, the resolution must be more precise.

Here we revealed confusion in the previous identification among *n*-alkane assimilating *Candida* yeasts. Moreover, we found that the identification method based on the sequences of D1/D2 domain from the LSU rDNA gene is accurate and reliable, and because sequencing technology has become so common, this method can be routinely applicable to avoid mistakes.

Acknowledgments

We thank Dr. C. P. Kurtzman for useful suggestions. We also thank Dr. M. Ohkuma and Dr. T. Iida for help in carrying out a phylogenetic analysis on the D1/D2 region of LSU-rDNA.

References

- Fell, J. W. (1995) rDNA targeted oligonucleotide primers for the identification of pathogenic yeasts in a polymerase chain reaction. *J. Ind. Microbiol.*, **14**, 475–477.
- Fukazawa, Y., Nakase, T., Shinoda, T., Nishikawa, A., Kagaya,

Fig. 2. Phylogenetic tree based on the D1/D2 region sequences of *Candida albicans* clade.

Yeast strains and the sources of D1/D2 region of LSU-rDNA sequences, accession numbers in the DDBJ, EMBL and GenBank Nucleotide Sequence Database are *Saccharomyces cerevisiae* (U44806), *Pichia guilliermondii* (U45709), *Lodderomyces elongisporus* (U45763), *Candida parapsilosis* (U45754), *C. tropicalis* (U45749), *C. sojae* (U71070), *C. cloacae* (this study), *C. subtropicalis* (this study), *C. sp. 1098* (this study), *C. sp. M1210* (this study), *C. maltosa* (U45745), *C. viswanathii* (U45752), *C. lodderae* (U45755), *C. tropicalis* pk233 (this study), *C. albicans* (U45776), and *C. dubliniensis* (U57685). The bar under the tree indicates the distance corresponding to five changes per 100 nucleotide positions.

- K., and Tsuchiya, T. (1975) Significance of cell wall structures on yeast classification : Proton magnetic resonance and serological and deoxyribonucleic acid characterization of *Candida sake* and related species. *Int. J. Syst. Bacteriol.*, **25**, 304–314.
- Hara, A., Ueda, M., Matsui, T., Furuhashi, K., Kanayama, N., and Tanaka, A. (1999) Construction of an autonomously replicating plasmid in *n*-alkane-assimilating yeast, *Candida tropicalis*. *J. Biosci. Bioeng.*, **87**, 717–720.
- Hoffman, C. S. and Winston, F. (1987) A ten-minutes DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene*, **57**, 267–272.
- Kanayama, N., Ueda, M., Atomi, H., and Tanaka, A. (1998) Genetic evaluation of physiological functions of thiolase isoenzymes in the *n*-alkane-assimilating yeast *Candida tropicalis*. *J. Bacteriol.*, **180**, 690–698.
- Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **16**, 111–120.
- Komagata, K., Nakase, T., and Katsuya, N. (1964a) Assimilation of hydrocarbons by yeasts I. Preliminary screening. *J. Gen. Appl. Microbiol.*, **10**, 313–321.
- Komagata, K., Nakase, T., and Katsuya, N. (1964b) Assimilation of hydrocarbons by yeasts II. Determination of hydrocarbon-assimilating yeasts. *J. Gen. Appl. Microbiol.*, **10**, 323–331.
- Kurtzman, C. P. and Robnett, C. J. (1995) Molecular relationships among hyphal ascomycetous yeasts and yeastlike taxa. *Can. J. Bot.*, **73**, S824–S830.
- Kurtzman, C. P. and Robnett, C. J. (1997) Identification of clinically important Ascomycetous yeasts based on nucleotide divergence in the 5' end of the large-subunit (26S) ribosomal DNA gene. *J. Clin. Microbiol.*, **35**, 1216–1223.
- Kurtzman, C. P. and Robnett, C. J. (1998) Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Leeuwenhoek*, **73**, 331–371.
- Magee, B. B., D'Souza, T. M., and Magee, P. T. (1987) Strain and species identification by restriction fragment length polymorphism in the ribosomal DNA repeat of *Candida* species. *J. Bacteriol.*, **169**, 1639–1643.
- Meyer, S. A., Anderson, K., Brown, R. E., Smith, M. TH., Yarrow, D., Mitchell, G., and Ahearn, D. G. (1975) Physiological and DNA characterization of *Candida maltosa*, a hydrocarbon-utilizing yeast. *Arch. Microbiol.*, **104**, 225–231.
- Meyer, S. A., Payne, R. W., and Yarrow, D. (1998) *Candida* Berkhout. In *The Yeasts: A Taxonomic Study*, Fourth ed., ed. by Kurtzman, C. P. and Fell, J. W., Elsevier Sci. Publ., Amsterdam, Netherlands, pp. 77–100.
- ed. by Kurtzman, C. P. and Fell, J. W., Elsevier Sci. Publ., Amsterdam, Netherlands, pp. 454–573.
- Nakase, T., Fukazawa, T., and Tsuchiya, T. (1972) A comparative study on two forms of *Candida tropicalis* (CAST.) Berkhout. *J. Gen. Appl. Microbiol.*, **18**, 349–363.
- O'Donnell, K. and Gray, L. E. (1995) Phylogenetic relationships of the soybean sudden death syndrome pathogen *Fusarium solani* f. sp. *phaseoli* inferred from rDNA sequence data and PCR primers for its identification. *MPMI*, **8**, 709–716.
- Ohkuma, M., Hwang, C. H., Masuda, Y., Nishida, H., Sugiyama, J., Ohta, A., and Takagi, M. (1993a) Evolutionary position of *n*-alkane-assimilating yeast *Candida maltosa* shown by nucleotide sequence of small-subunit ribosomal RNA gene. *Biosci. Biotechnol. Biochem.*, **57**, 1793–1794.
- Ohkuma, M., Muraoka, S., Hwang, C. W., Ohta, A., and Takagi, M. (1993b) Cloning of the *C-URA3* gene and construction of a triple auxotroph (*his5*, *ade1*, *ura3*) as a useful host for the genetic engineering of *Candida maltosa*. *Curr. Genet.*, **23**, 205–210.
- Okazaki, K., Takechi, T., Kambara, N., Fukui, S., Kubota, I., and Kamiryo, T. (1986) Two acyl-coenzyme A oxidases in peroxisomes of the yeast *Candida tropicalis*: Primary structures deduced from genomic DNA sequence. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 1232–1236.
- Okazaki, K., Tan, H., Fukui, S., Kubota, I., and Kamiryo, T. (1987) Peroxisomal acyl-coenzyme A oxidase multigene family of the yeast *Candida tropicalis*; nucleotide sequence of a third gene and its protein product. *Gene*, **58**, 37–44.
- Suzuki, S., Uemura, N., and Watanabe, A. (1975) JP75154993.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (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.
- Uemura, N. (1985) Biseibutsu ni yoru *n*-paraffins kara nienkisan seisan no kougyouka (Industrialization of production of dicarboxylic acid from *n*-paraffins) (in Japanese). *Hakko to Kogyo*, **43**, 436–441.
- Williams, D. W., Wilson, M. J., Lewis, M. A. O., and Potts, A. J. C. (1995) Identification of *Candida* species by PCR and restriction fragment length polymorphism analysis of intergenic spacer regions of ribosomal DNA. *J. Clin. Microbiol.*, **33**, 2476–2479.
- Yarrow, D. (1998) Methods for the isolation, maintenance and identification of yeasts. In *The Yeasts: A Taxonomic Study*, Fourth ed., ed. by Kurtzman, C. P. and Fell, J. W., Elsevier Sci. Publ., Amsterdam, Netherlands, pp. 77–100.