Strain and Species Identification by Restriction Fragment Length Polymorphisms in the Ribosomal DNA Repeat of *Candida* Species†

BEATRICE B. MAGEE, TREVOR M. D'SOUZA, AND P. T. MAGEE*

Department of Microbiology, Michigan State University, East Lansing, Michigan 48824-1101

Received 25 July 1986/Accepted 5 January 1987

Restriction fragment length polymorphisms in the ribosomal DNA (rDNA) have been shown to be a useful criterion for distinguishing among various isolates of *Candida albicans*. In a sample of 12 clinical isolates, we found six different classes based on variations in the fragments produced from genomic DNA by *Eco*RI and visualized after Southern transfer by being probed with a plasmid containing *Saccharomyces cerevisiae* rDNA. Some of the classes appeared to be heterozygous at the rDNA locus. Similar digestion of other *Candida* species showed that each could be identified on the basis of its restriction patterns. Since these are highly reiterated genes, the differences were apparent on ethidium bromide-stained gels; Southern transfers were not necessary. *Eco*RI restriction maps of the rDNA of *C. albicans*, *C. stellatoidea*, *C. tropicalis*, and *C. guilliermondii* were determined.

Taxonomic classification of the clinically important yeasts is a difficult task. The absence of a sexual cycle in many of these organisms makes classification schemes based on the use of teleomorphic characteristics impossible. Nutritional characteristics, including fermentation and assimilation of various substrates, have been extremely useful, but it is clear from studies of perfect genera that nutritional variation can occur among members of a clearly defined species, since a single mutation can cause loss of the ability to ferment any one carbon source. Cell morphology, particularly for fungal species that undergo dimorphic transitions, is a useful criterion and is probably significantly less sensitive to mutation than the nutritional spectrum is. However, it is also somewhat less easy to score.

An ideal approach to classification would involve the use of variations in a property which is important if not crucial to cell survival, which is easy to score, and which shows wide variation among the various organisms to be classified. It occurred to us that ribosomal DNA (rDNA) restriction fragment length polymorphisms (RFLPs) might serve as a distinguishing characteristic to facilitate the classification of clinically important yeasts. In this paper, we show that the 17S and 25S rRNA genes are sufficiently conserved among a large number of genera that one probe can be used to study rDNA in species ranging from *Candida glabrata* to *Candida guilliermondii*, that in fact the general pattern of rDNA RFLPs is conserved within species as defined by classical criteria (with the exception of *Candida albicans* and *C. albicans* var. *stellatoidea*), and that the variation in this parameter is sufficient to make this approach an extremely rapid and accurate way to classify isolates.

MATERIALS AND METHODS

Preparation of DNA. The strains used in this study are listed in Table 1. The cells were grown on YEPD (6 g of yeast extract, 6 g of peptone, 12 g of glucose per 600 ml) at 30°C from single colonies in a volume of 50 ml. DNA was prepared by a modification of the method of Hereford et al. (2) as follows. Stationary-phase cells were suspended in 2.5

ml of 0.2 M Tris (unbuffered)–5% β -mercaptoethanol and incubated for 30 min at room temperature. A 5-ml amount of 1 M sorbitol–0.004 M potassium phosphate (pH 7.0) was added to the suspension. The cells were spun down and resuspended in sorbitol phosphate (1.25 ml), Zymolyase 5000 (Kirin Brewery, Tokyo, Japan) was added to a final concentration of 2.5 mg (wet weight) of cells per ml, and the cells were incubated at 37°C for 30 to 60 min until they contained about 80% spheroplasts. The spheroplasts were spun down and then lysed by incubation at 60°C for 30 min in lysing buffer (0.2 M NaCl, 0.1 M EDTA, 5% sodium dodecyl sulfate, 50 mM Tris [pH 8.5], 20 μ g of pronase per ml). Nucleic acids were extracted twice with TE (10 mM Tris, 1 mM EDTA [pH 8.3])–saturated phenol–chloroform–isoamyl alcohol (25:24:1). The aqueous phase was separated, DNA was precipitated with 2 volumes of 95% ethanol, and the DNA precipitate was washed in 70% ethanol and then in ether. The dried precipitate was dissolved in 1 ml of TE and incubated for 30 min at 37°C with 20 μ g of RNase per ml and then for 30 min at 37°C with 10 μ g of pronase per ml. The extraction and precipitation steps were repeated, and the DNA was dissolved and stored in TE.

Digestion of DNA and separation of fragments. DNA (10 μ g) was digested with restriction endonucleases *Eco*RI, *Hind*III, *Bam*HI, *Xba*I, *Kpn*I, *Xho*I, *Hinf*I, and *Eco*RV, (2 U of enzyme per μ g of DNA for 24 h at 37°C). To ensure complete digestion, an additional 1 U of enzyme was added (per microgram of DNA) to the reaction mixture and further incubated (24 h) at 37°C. Other conditions for digestion were as recommended by the supplier (New England BioLabs, Inc., Beverly, Mass., or Bethesda Research Laboratories, Inc., Gaithersburg, Md.). DNA fragments were separated electrophoretically on 1% agarose gels in Tris-acetate buffer (40 mM Tris, 20 mM sodium acetate, 2.0 mM EDTA [pH 7.5]) at 70 mA for approximately 3 h. The gels were stained with ethidium bromide and transferred to nitrocellulose filters by the method of Southern (7).

Cloned rDNA probes were prepared by using nick-translated DNAs (6). Labeled DNA was passed through a Sephadex G-50 column (Pharmacia Fine Chemicals, Piscataway, N.J.) to free it from low-molecular-weight material. A specific activity of 0.5×10^7 to 2×10^7 cpm/ μ g of labeled DNA was used for hybridization. The blots were

* Corresponding author.

† Journal article no. 12138 from the Michigan Agricultural Experiment Station.

TABLE 1. List of strains used in this study

Strain	Source
<i>C. albicans</i> FC18	A. Rogers
<i>C. albicans</i> C9	D. Ahern
<i>C. albicans</i> 1810	M. Gurwith
<i>C. albicans</i> 1840	M. Gurwith
<i>C. albicans</i> 1892	M. Gurwith
<i>C. albicans</i> CBS5736	T. Suzuki
<i>C. albicans</i> IAM 4996	T. Suzuki
<i>C. albicans</i> STN57	T. Suzuki
<i>C. albicans</i> AA	J. Bennett
<i>C. albicans</i> BB	J. Bennett
<i>C. albicans</i> CC	J. Bennett
<i>C. albicans</i> DD	J. Bennett
<i>C. albicans</i> EE	J. Bennett
<i>C. albicans</i> FF	J. Bennett
<i>C. albicans</i> var. <i>stellatoidea</i> Y-523	C. Kurtzman
<i>C. albicans</i> var. <i>stellatoidea</i> Y-2444	C. Kurtzman
<i>C. glabrata</i> 81/011	F. C. Odds
<i>C. glabrata</i> 85/038	F. C. Odds
<i>C. guilliermondii</i> Y-324	C. Kurtzman
<i>C. guilliermondii</i> Y-7572	C. Kurtzman
<i>C. tropicalis</i> 1856	M. Gurwith
<i>C. tropicalis</i> MDH-1	A. Rogers
<i>C. tropicalis</i> 584	A. Rogers
<i>C. tropicalis</i> 81/019	F. C. Odds
<i>C. parapsilosis</i> 81/042	F. C. Odds
<i>S. cerevisiae</i> AH-22	J. Hicks

hybridized for 48 h at 42°C on a rocking platform (Labquake; Labindustries, Berkeley, Calif.) (1). Hybridized blots were washed by being soaked in three changes of 0.010 M Tris–0.001 M EDTA–0.1 M NaCl–0.1% sodium dodecyl sulfate–0.1% sodium PP_i at room temperature and then in two changes of the same buffer at 60°C. The blots were air dried and exposed to Kodak X-Omat-ARS film at –70°C. Whenever needed, Cronex Lightning-Plus (E. I. du Pont de Nemours & Co., Inc., Wilmington, Del.) intensifier screens were used. Sizes of bands obtained by autoradiography were determined by reference to the seven *Eco*RI rDNA bands of *Saccharomyces cerevisiae*, to a *Hind*III digest of λ bacteriophage, and to the *S. cerevisiae* rDNA plasmid described by Petes et al. (5).

Isolation of rDNA plasmids from *C. albicans*. An incomplete library of *C. albicans* genomic DNA was prepared by ligating fragments from a *Sau*3A partial digest into the *Bam*HI site of plasmid pUC8 and transforming into *Escherichia coli* HB101. The library was screened by colony hybridization with partially hydrolyzed rRNA as a probe. Two positive plasmids were identified and purified, and the insert was shown to hybridize to 18S rRNA (for p5i) and 25S rRNA (for p1f). These plasmids were then used to probe genomic restriction digests to identify the fragments that encode the two species of rRNA.

RESULTS

Some strains of *C. albicans* vary in the fragments generated by *Eco*RI from their rDNA repeats. We originally observed that the *S. cerevisiae* rDNA clone (5) could be used to probe a blot of *C. albicans* genomic DNA; thus there was a great deal of homology between the rDNAs of these two strains. In an attempt to find an enzyme which gave a reasonable number of well-separated bands, we examined the restriction patterns generated by *Eco*RI, *Hind*III, *Xba*I, *Kpn*I, *Xho*I, and *Eco*RV. The results with *Eco*RI were most

encouraging. With most *C. albicans* strains, this enzyme gave three hybridizing bands, one at 6.4 kilobases (kb) (A), one at 3.75 kb (B), and a third at 2.6 kb (C). We call this the standard or classic *C. albicans* pattern. These patterns were generally quite reproducible among strains of this species, but, as discussed below, there are some differences. Figure 1 shows a typical gel with separated digests of 12 strains of *C. albicans*. The classification of each of these strains was verified by Analytab Products tests (a classification approach based on the nutritional spectrum, marketed by Analytab Products, Plainview, N.Y.), and by the observation that they made germ tubes and chlamydospores. It is apparent from Fig. 1 that there are differences in several strains (1810, 1840, AA, FF, and IAM 4996). These include extra bands of about 6.0 kb in 1810 and 6.8 and 5.8 kb in 1840 and a decrease in the size of the 2.6-kb band to around 2.5 kb in AA, to 2.4 kb in 1840 and FF, and to 3.1 kb in IAM 4996 (which also has a band of 6.6 kb in place of the 6.4-kb band). The extra bands in CBS 5736 are most probably partial digests. Also shown for comparison in the figure is the pattern of *S. cerevisiae* AH-22. It is evident that the *S. cerevisiae* AH-22 bands are very different in size from those of *C. albicans*. Thus, for the 12 *C. albicans* strains, there are six classes distinguishable by variation of the *Eco*RI restriction fragment lengths. However, of the six classes, one contains seven members and the rest contain one member each. Thus, there seems to be a common pattern with relatively frequent variations.

***Eco*RI RFLPs and *Hinf*I are reliable ways of determining the species of medically important yeasts.** The observation that a single probe would hybridize with a large number of *C. albicans* strains encouraged us to examine a variety of

FIG. 1. Autoradiogram from a Southern blot of *Eco*RI-digested chromosomal DNA from various strains of *C. albicans* and *S. cerevisiae* after hybridization for 48 h with a ³²P-labeled rDNA probe. *Eco*RI-digested rDNA plasmid was also run. Fragment sizes are given in kilobases.

FIG. 2. (A) Agarose gel electrophoresis of *Eco*RI-digested chromosomal DNA obtained from various *Candida* species, *S. cerevisiae*, and *Schizosaccharomyces pombe*. *Eco*RI-digested rDNA was also run. (B) Autoradiogram from a Southern blot of *Eco*RI-digested chromosomal DNA from various *Candida* species and *S. cerevisiae*, after hybridization for 48 h with 32 P-labeled rDNA probe. rDNA digested with *Eco*RI was also run. Fragment sizes are given in kilobases.

Candida species to see whether we could reliably distinguish among medically important isolates. The bands corresponding to the rDNA and to other highly reiterated sequences make the patterns of the digests from the various strains distinguishable on ethidium bromide-stained gels; no blotting is needed (Fig. 2A). The bands from FC18 and STN57, typical *C. albicans* strains, are shown, along with the digests of two *C. stellatoidea* strains; although the small number of phenotypic differences between *C. albicans* and *C. stellatoidea* has led to the suggestion that these two organisms ought to be considered a single species, the RFLPs from reiterated DNA are distinct. The difference in the rDNA genes, however, appears to be due to a single restriction site change (see below). Two *C. tropicalis* strains are shown, whose patterns, although similar among themselves, diverge quite clearly from those of any of the other organisms. *C. glabrata*, *C. parapsilosis*, *C. guilliermondii*, *Schizosaccharomyces pombe*, and *S. cerevisiae* patterns are also shown. The patterns of each of these species differ from all others tested and are reproducible within the species for which more than one strain has been examined (which includes all except *Schizosaccharomyces pombe* and *C. parapsilosis*).

Figure 2B shows the rDNA blots of several strains. Each has its characteristic bands: in addition to the species shown in Fig. 1 for *C. albicans*, we can see bands of 6.4, 4.0, and 2.45 kb for *C. stellatoidea*, bands of 5.0, 2.9, 1.95, and 0.67 kb for *C. glabrata*, bands of 4.0, 3.0, and 0.95 kb for *C. tropicalis*, and bands of 3.8, 3.2, 2.1, and 1.7 kb for *C. parapsilosis*. *C. tropicalis* and *C. stellatoidea* each have a band of approximately the size of the B band (3.75 kb) of *C. albicans*.

Figure 3 shows an ethidium bromide-stained gel of *Hinf*I-

FIG. 3. Agarose gel electrophoresis of *Hinf*I digests of various yeast species. *Hind*III digests of phage λ were run as size standards.

FIG. 4. Restriction maps of the rDNA of several *Candida* species. Genomic DNA was isolated, digested, and subjected to gel electrophoresis as described in Materials and Methods. The Southern transfers were probed with the *HinfI* fragment and the plasmids shown at the bottom of the figure. It was assumed that fragments that hybridized with the *HinfI* probe contained the 5S RNA, that those complementary to p5i contained the 18S RNA, and that homology to p1f marked the 25S RNA. The unlabelled vertical lines mark *EcoRI* sites. There are a large number of *HinfI* sites in the *C. albicans* rDNA, and only the two that generate the large *HinfI* fragment are indicated. These vary from strain to strain (cf. Fig. 3).

digested DNA from the same *Candida* species shown in Fig. 2A. It is clear that this enzyme, which has a 4-base specificity, also yields fragments which are typical of particular species. However, polymorphisms are apparent in each of the pairs of strains shown. In the *C. albicans* strains, the largest band is 4.8 kb in FC18 and 6.1 kb in STN57, and in the *C. stellatoidea* pair, the variation is about 0.5 kb. For the *C. tropicalis* strains, the *HinfI* patterns are quite dissimilar, although there are several bands in common. For the *C. glabrata* strains, there is a band at about 2.7 kb in 81-011 which is absent in 85-038, but the other two observable bands are similar. Since all these pairs have identical *EcoRI* rDNA patterns, the *HinfI* patterns may be an excellent way of differentiating among isolates of a given yeast.

Organization of the rDNA genes in *Candida* species. Blots analogous to those in Fig. 2B were probed with the plasmids p5i and p1f, which contain portions of the *C. albicans* 18S and 25S RNA genes, respectively (data not shown). In addition, digestion of *C. albicans* DNA with *HinfI* gives a highly repeated fragment which gives a very weak signal when probed with the rDNA plasmid but which hybridizes moderately strongly with the *S. cerevisiae* rDNA plasmid when isolated and nick translated itself. We explain this by the argument that the fraction of the rDNA plasmid which is homologous to the *HinfI* band is a small fraction of the total

DNA. When the rDNA plasmid is nick translated, perhaps 2 to 5% of the counts are homologous to the *HinfI* band. On the other hand, perhaps 30% of the *HinfI* sequences are homologous to the rDNA plasmid; thus, when used as a probe, the band lights up the plasmid. The *HinfI* fragment also hybridizes with *S. cerevisiae* 5S rDNA and with DNA from the other *Candida* species. Using these probes, we have derived the *EcoRI* restriction maps of the rDNA region shown in Fig. 4 for a number of medically important yeasts and have completed a relatively detailed map for *C. albicans*. (We have assumed that transcription of the rRNA precursor in these organisms is similar to that in other eucaryotes.)

DISCUSSION

The first important point to be made about the findings we report here is that the yeast taxonomists have done a generally excellent job. All the species we examined by the RFLP method have proved to be reasonably homogeneous internally and quite distinct from other species. A striking finding was that *C. stellatoidea* and *C. albicans* are distinct by this criterion, although their biochemical and physiological phenotypes are quite similar. Their electrophoretic karyotypes, however, are quite different (B. B. Magee and P. T. Magee, *J. Gen. Microbiol.*, in press). Detailed restriction mapping of the rDNA genes of these two organisms suggests that the difference is due to a single *EcoRI* site change; i.e., both have three *EcoRI* sites in the rDNA repeat, but in *C. albicans* they are spaced at 0, 6.4, and 3.6 kb and in *C. stellatoidea* they are at 0, 6, and 4 kb (Fig. 4).

The large differences in the restriction patterns for the rDNA regions of the various *Candida* species may serve as a rapid means of differentiating among these organisms. The ethidium bromide-stained gels can be used to differentiate quite easily among these organisms (Fig. 2A). A *HinfI* digest is also definitive for distinguishing *C. albicans* and *C. stellatoidea* from other *Candida* species. This 4-base-pair-specific enzyme generates a very large number of polymorphisms and may be useful in conjunction with *EcoRI* in biotyping clinical isolates. This approach may therefore prove to be convenient to use for identification to species level and for biotyping at the level of clinical investigation.

The variations in restriction patterns in the *C. albicans* strains seem to point in some cases to heterozygosity for the rDNA genes. For example, strains 1810 and 1840 both have the ABC pattern but show additional bands. The best explanation for this is that the organization on one chromosome differs from the organization on the other. Petes (4) and Szostak and Wu (9) have shown that in *S. cerevisiae*, gene conversion tends to correct such differences, so that the rDNA in diploids becomes homogeneous with respect to RFLPs, but this occurs during meiosis. This phenomenon may occur during vegetative growth in *C. albicans*. If so, it would account for the seeming homozygous variants from the majority case, such as FF and IAM 4996.

In *Histoplasma (Emmonsia) capsulatum*, both mitochondrial DNA and rDNA show variation in restriction enzyme sites (10). The variation in mitochondrial DNA has been more carefully examined than the one in rDNA, but in strains representing the three classes found by mitochondrial DNA differences, the rDNA patterns varied a great deal. This is somewhat surprising, given the fact that *H. capsulatum* is a perfect fungus, has a sexual cycle, and thus might be expected to correct polymorphisms during meiosis as does *S. cerevisiae*.

The recent observation by Suzuki et al. (8) that *C. albicans* and *C. guilliermondii* are to a limited extent interfertile, despite a large number of phenotypic differences, is interesting in the light of the very large differences in the rDNA restriction patterns we found in the two species. However, there are several varieties of fungi (e.g., *Issatchenkia sculata* var. *sculata* and *I. sculata* var. *oxigua*) in which conjugation and viable meiotic progeny are produced with 25% DNA homology (3). One would imagine that RFLPs in a conserved region are a less sensitive, although more convenient, measure of relatedness than DNA-DNA homology, so that *C. guilliermondii* and *C. albicans* may be more closely related than the two *Issatchenkia* species.

The observations reported here may have some implications for clinical mycology. It seems quite possible that RFLPs are very useful in biotyping various strains of *C. albicans*, thus both aiding in studies of the epidemiology of the infections and possibly allowing the identification of virulent strains (if such exist). At least they may aid in studies of reinfection versus relapse and autoinfection in susceptible patients. A broader survey is, of course, needed to determine whether the differences we found are common or rare in clinical isolates.

A second area of clinical interest is whether the differences among species could be exploited for rapid laboratory identification of these organisms. The fact that even on a stained gel the fragment pattern is so characteristic may mean that species identification can be made without hybridization. If this is so, simplified DNA preparation may be all that is required to place the technique at the disposal of clinical laboratories.

ACKNOWLEDGMENTS

We thank D. Primerano, E. Rikkerink, T. Suzuki, and A. Kennedy for helpful discussion.

This work was supported by Public Health Service grant AI16567 from the National Institutes of Health.

LITERATURE CITED

1. Engle, J. D., and J. B. Dodgson. 1981. Histone genes are clustered but not tandemly repeated in the chicken genome. *Proc. Natl. Acad. Sci. USA* 78:2856-2860.
2. Hereford, J., K. Fahrner, J. Woolford, Jr., M. Rosbash, and D. B. Kaback. 1979. Isolation of yeast histone genes H2A and H2B. *Cell* 18:1261-1271.
3. Kurtzman, C. P., M. J. Smily, and C. J. Johnson. 1980. Emendation of the genus *Issatchenkia kudriavzev* and comparison of species by deoxyribonucleic acid association, mating reaction, and ascospore ultrastructure. *Int. J. Syst. Bacteriol.* 30:503-513.
4. Petes, T. D. 1980. Unequal meiotic recombination within tandem arrays of yeast ribosomal genes. *Cell* 19:765-774.
5. Petes, T. D., L. M. Hereford, and K. Skryabin. 1978. Characterization of two types of yeast ribosomal DNA genes. *J. Bacteriol.* 134:295-305.
6. Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase. *J. Mol. Biol.* 113:237-251.
7. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
8. Suzuki, T., A. L. Rogers, and P. T. Magee. 1986. Inter- and intraspecies crosses between *Candida albicans* and *Candida guilliermondii*. *Yeast* 2:53-58.
9. Szostak, J. W., and R. Wu. 1980. Unequal meiotic crossing-over in the ribosomal DNA of *Saccharomyces cerevisiae*. *Nature (London)* 284:426-429.
10. Vincent, R. D., R. Goewart, W. E. Goldman, G. S. Kobayashi, A. M. Lambowitz, and G. Medoff. 1986. Classification of *Histoplasma capsulatum* isolates by restriction fragment polymorphisms. *J. Bacteriol.* 165:813-818.