JOURNAL OF CLINICAL MICROBIOLOGY, Oct. 1992, p. 2674–2679
0095-1137/92/102674-06$02.00/0

Vol. 30, No. 10

Comparison of Molecular Typing Methods for *Candida albicans*

P. T. MAGEE,* LINDA BOWDIN, AND JEFF STAUDINGER†

Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55108

Received 20 April 1992/Accepted 20 July 1992

Four molecular approaches to determining the types of *Candida albicans* strains were compared. The strains used were those whose repeated DNA (ribosomal and mitochondrial) *Eco*RI restriction fragment length polymorphisms (RFLP) were determined by Stevens et al. (D. A. Stevens, F. C. Odds, and S. Scherer, Rev. Infect. Dis. 12:258–266, 1990). Scherer and Stevens (S. Scherer and D. A. Stevens, Proc. Natl. Acad. Sci. USA 85:1452–1456, 1988) used the same strains to examine the Southern blots of genomic *Eco*RI digests probed with the repeated sequence 27A. The results of these investigators were compared with determinations of RFLPs generated from repeated DNA by the enzyme *Hin*fI and examination of the karyotypes of strains under two sets of conditions, one for the smaller chromosomes and one for the larger ones. Analysis of RFLPs of repeated DNA is most convenient but shows the lowest degree of resolution. Use of the repeated sequence and use of karyotype have very high resolution, but the former method is more convenient than the latter. *Hin*fI digestion is more sensitive than *Eco*RI digestion but equally convenient. By using all four methods, separate types were identified for 18 of the 20 strains examined.

Candida albicans is a major fungal pathogen of immunosuppressed patients. Several studies of the epidemiology of *Candida* infections have been carried out (3, 6, 9–13), but definitive relationships between type and such properties as pathogenicity, commensality, and infectivity or between relapse and reinfection have not been established. In part, this lack is due to the difficulty of defining types specifically and reproducibly. As a result, biotyping of *Candida* species has been the subject of several recent reports (1, 2, 5, 8, 13–15).

These recent reports and the earlier literature have been reviewed by Merz (6), who compared the advantages and disadvantages of the various methods. He suggested that DNA typing, especially by determining restriction fragment length polymorphisms (RFLPs) and by electrophoretic karyotyping, is one of the methods with the greatest potential for discrimination.

These two methods theoretically seem very satisfactory in determining strain identity or difference, since they directly measure genotypic differences rather than phenotypic ones, and the definition of two identical organisms is that their genomes have the same DNA sequence. One of the simplest of the genomic approaches makes use of the highly repeated DNA sequences, including those of ribosomal DNA (rDNA) and mitochondrial DNA. Digestion with a single restriction enzyme like *Eco*RI yields several fragments which are present in concentrations high enough to reveal the fragments as distinct bands on a stained electrophoretic gel (4, 9). These bands are derived from the mitochondrial DNA and rDNA, and they vary enough in size to permit the assignment of various strains to biotyping groups (9). Other restriction enzymes, most notably *Hin*fI, which produces a characteristic fragment from the variable spacer region in the rDNA (4), have been proposed to replace *Eco*RI.

Analysis of RFLPs derived from the repeated DNA has the advantage of being easy and rapid, but the number of possible types distinguishable appears to be limited. Furthermore, some patterns are very common, and quite different strains may appear identical by this criterion. The two methods which seem to have the greatest potential resolving power are determining RFLPs by using a moderately repeated DNA sequence as a probe (10, 13) and examining the electrophoretic karyotype (1, 5, 6, 11, 15). One report has used both RFLPs of the repeated DNA and the karyotype and has argued that analysis of the karyotype is much more sensitive (14). In both cases, changes seem to occur in a stochastic manner with a relatively low frequency per cell division. Therefore, these properties are likely to remain stable over the length of time that one would want to study a particular strain, and they make the typing method reproducible. However, the properties vary often enough so that they have great resolving power in distinguishing among different strains. Of course, given the apparent lack of genetic exchange in *C. albicans*, divergence is irreversible, so any changes that occur are immediately fixed. This means that strains varying by only a single difference in either the restriction pattern or the karyotype are probably related.

In *C. albicans*, seven of the chromosomes are numbered 1 to 7 in order of decreasing size (1 is the largest). The rDNA-bearing chromosome is always one of the three slowest migrating chromosomes in a pulsed-field gel electrophoresis separation. However, its migration rate varies with respect to the other two large chromosomes. Hence, it is called R, rather than being given a number.

One highly variable aspect of the electrophoretic karyotype is the size of the two homologs of chromosome R. Apparently because of unequal crossing over, the migration rates of these homologs change with a very high frequency (16). One should therefore be cautious in using differences in the migration rate of these homologs to separate strains into different types.

Merz suggests that it would be useful to compare results of various typing methods on a large set of strains (6).

In order to compare the relative usefulness of typing by electrophoretic karyotype with typing by RFLPs (ranging

* Corresponding author.

† Present address: Department of Microbiology, University of Texas at Houston, Houston, TX 77030.

from 5 to 10 to 50 to 100 repeats per genome), we investigated a set of strains originally typed by Scherer et al. (10, 14). These were isolated during a hospital outbreak of candidosis, and their relatedness was determined both by the RFLPs of cellular DNA visible by ethidium bromide staining after digestion with the restriction enzyme *Eco*RI and by RFLPs revealed by probing the same digestions with the repeated element 27A. In this paper, we compare the types generated by analysis of karyotypes of the strains with the published results from the other two methods (9, 13). We also examine the usefulness of the *Hin*fI RFLPs in biotyping.

MATERIALS AND METHODS

Strains and culture conditions. The strains used for electrophoretic analysis were those described by Scherer and Stevens (10) and further examined by Stevens et al. (14). The cells were grown on YEPD medium (4) at 30°C.

DNA preparation and restriction enzyme digestion. The DNA was prepared according to the method of Scherer and Stevens (9), and digestion and Southern transfer were carried out as described by Magee et al. (4). Restriction enzymes were used according to the instructions of the suppliers. *Eco*RI and *Hin*fI were purchased from Gibco/BRL.

Determination of electrophoretic karyotype. For determination of the electrophoretic karyotype, cells were prepared as described elsewhere (7). Electrophoresis was carried out on a Bio-Rad CHEF-DRII. Conditions were varied according to the region of the gel where maximum resolution was desired. To separate the larger chromosomes (R to 3), 0.4% agarose was used, and the pulse regimen was as follows: 120 to 300 s for 30 h, 420 to 900 s for 66 h, and 900 to 1,200 s for 24 h, all at 65 V. To separate chromosomes 4 to 7 and smaller chromosomes 0.8% agarose was used, the voltage was 150 V, and the pulses were 120 s for 24 h followed by 240 s for 36 h. All runs were at 12°C in 0.45 M Tris–0.45 M boric acid–0.001 M EDTA buffer.

RESULTS

Comparison of three methods of biotyping. In 1990, Stevens et al. used 20 strains of *C. albicans* isolated from patients with *Candida* infections in the same hospital to demonstrate the use of *Eco*RI RFLPs in biotyping. Six identifiable groups were found among these strains (14). In an earlier publication, Scherer and Stevens had analyzed the same group of strains by using the middle repeated element 27A (10). Although they did not divide these strains into groups, inspection of their data yields 10 types based solely on identical band position. Table 1 summarizes the data on these strains from these two papers.

Figure 1 shows the electrophoretic karyotypes of these strains under conditions designed to separate chromosomes R, 1, 2, and 3. Comparing the karyotypes shows several points immediately. In this figure, one can see that there are three bright bands in the upper part of the gel which do not vary significantly among the isolates. These are (from the top) chromosomes 1, 2, and 3. In contrast to the stability of these bands, there are numerous examples of lighter-staining bands between chromosomes 1 and 2 (e.g., strain 1 of Fig. 1A and strain 12, strain 14, and strain 15 in Fig. 1B). Some of these bands are the separated homologs of chromosome R, the rDNA-containing linkage group. Others are separated homologs of chromosome 1 or 2. Some strains (e.g., strain 11, Fig. 1B) have several new, large chromosomes between

TABLE 1. Results of three methods of biotyping *C. albicans* strains

Strain	RFLP analysis of type		Electrophoretic type		
	*Eco*RI[a]	27a[b]	Small chromosomes	Large chromosomes	All chromosomes
1	IB3	1	e1	E1	C1
2	IB3	2	e2	E2	C2
3	IA2	3	e2	E2	C2
4	IA2	4	e2	E2	C2
5	IA2	3	e2	E2	C2
6	IA8	2	e3	E2	C3
7	IA8	5	e4	E2	C4
8	IA11	6	e5	E2	C5
9	IA8	7	e4	E2	C4
10	IA8	8	e6	E3	C6
11	IA12	9	e7	E4	C7
12	IA11	6	e5	E2	C5
13	IA11	6	e8	E2	C8
14	IB3	10	e9	E5	C9
15	IA12	11	e2	E2	C10
16	IA8	8	e6	E2	C11
17	IB4	12	e2	E2	C2
18	IA12	11	e2	E2	C2
19	IA12	11	e2	E2	C2
20	IA11	10	e5	E2	C5

[a] From Fig. 2 in reference 14. Strains are numbered according to Scherer and Stevens (9) in order opposite to that in reference 14.
[b] Arbitrary type assignment based on the data in Fig. 5 of reference 9.

chromosomes 1 and 2. Chromosome R varies in size at a rapid rate; it is therefore somewhat unsatisfactory as a basis for biotyping. Separated homologs of the other chromosomes, on the other hand, are a more stable basis for strain differentiation. We can determine which of the light bands between chromosomes 1 and 2 are chromosome R homologs by blotting the gels and probing them with labeled rDNA. Figure 2 shows the results of such a blot for a separation of the chromosomes of strains 1 and 2 similar to that shown in Fig. 1A. The arrowheads indicate the positions of chromosomes 1 and 2 on this gel. Comparing the blot with Fig. 1A, one can see that for strain 1, one of the chromosome R homologs migrates at the same rate as chromosome 1, and the other is the slower, lighter band between chromosomes 1 and 2. For strain 2, on the other hand, both homologs of chromosome R migrate at the same rate, i.e., between chromosomes 1 and 2 under these conditions. Similar experiments (data not shown) suggest that the light bands in the chromosome 1-2 regions in strains 11, 16, 18, and 20 are variations in homologs of chromosomes other than chromosome R.

Using only the appearance of new bands which are not homologs of chromosome R, we can divide these strains into four types. Type E1 contains strain 1; type E2 contains strains 2 to 9, 12, 13, and 16 to 20; type E3 contains strain 10; type E4 contains strain 11; and type E5 is represented by strain 14.

Figure 3 shows electrophoretic karyotypes of the same strains run under conditions designed to separate the smaller chromosomes, 4 to 7. In this separation, variations in the number of bands appear in lanes 1, 10, 11, and 16 in Fig. 3A. Variations in migration rate appear in many of the remaining strains. Types e1 (strain 1), e2 (strains 2, 3, 4, 5, 15, 17, 18, and 19), e3 (strain 6), e4 (strains 7 and 9), e5 (strains 8, 12, and 20), e6 (strains 10 and 16), e7 (strain 11), e8 (strain 13),

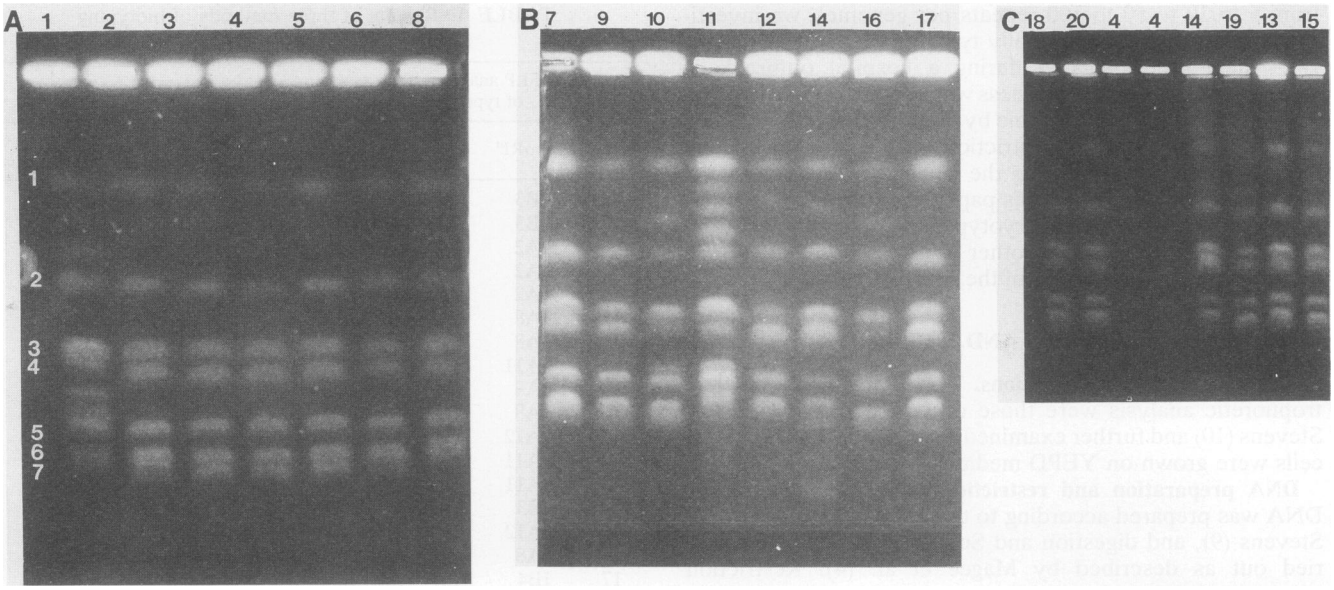

FIG. 1. Electrophoretic separation of chromosomes R, 1, 2, and 3. Samples were run as described in Materials and Methods for the large chromosomes. Numbers above the lanes indicate strains. Numbers on the side mark the regions of migration of the corresponding chromosomes. Chromosome R is not indicated because its migration rate varies highly from strain to strain, ranging from slower than chromosome 1 to as fast as chromosome 2.

and e9 (strain 14) can be identified. Many of these distinctions are subtle and depend on more than one gel to ascertain. Table 1 summarizes the typing data for the 20 strains examined here. The RFLPs with highly repeated DNA yield 6 types, the 27A RFLPs yield 10 types, and the electrophoretic karyotypes yield 9 types when the smaller chromosomes are examined and 5 types when the larger chromosomes are analyzed. Using the karyotyping data as a whole, 10 types can be distinguished (Table 1). By utilizing the data from all three analyses, 16 types can be distinguished for the 20 strains. Thus, the resolving powers of all three techniques come close to allowing the unequivocal identification of single strains, even closely related ones.

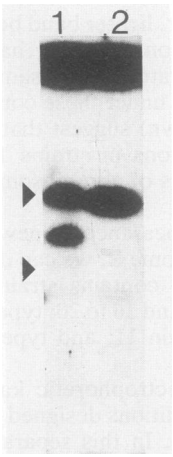

FIG. 2. rDNA hybridization of electrophoretic karyotypes of strains 1 and 2. The chromosomes were separated as in Fig. 1, transferred to Zetaprobe (Bio-Rad), and hybridized to an rDNA probe. Arrowheads indicate the positions of chromosomes 1 and 2 on the original separation.

Use of *Hin*fI restriction enzyme polymorphisms. We pointed out in an earlier paper that digestion of genomic DNA with *Hin*fI yields a single band from the highly repeated rDNA repeat and that since this enzyme seems to cut in the nonconserved spacer region, the sizes of the fragments generated are highly variable (4). It seemed useful to test this molecular-typing approach on the strains in the present study. We therefore carried out a *Hin*fI digestion as described previously (4) and measured the separated bands. Figure 4 shows the stained gel, and Table 2 gives the sizes in base pairs of the major *Hin*fI bands. It is evident that this method, too, has excellent resolving power, yielding 9 types based on size of the major band among the 20 strains. Note that strain 1 has two prominent *Hin*fI bands. We attribute this to heterozygosity at the rDNA locus. The lighter bands provide yet another strain signature for typing. We do not know the source of these bands, but they appear reproducibly and must stem from some other repeated DNA, possibly mitochondrial. (Their relative brightness makes it clear that they are not derived from rDNA.) Using them allows one to generate 13 types by *Hin*fI digestion alone (Table 2). While the correlation of *Hin*fI polymorphisms with the *Eco*RI RFLP type is high (but not perfect), the correlation with other biotyping methods is rather low. For example, strains 8 and 12, 10 and 13, and 14 and 20, pairs indistinguishable on the basis of the RFLPs demonstrated by the 27A probe, have different-sized major *Hin*fI bands. The overall *Hin*fI types provide even more resolution. Combining the data in Tables 1 and 2, it can be shown that by the use of all the methods described in this paper, the 20 strains can be resolved into 18 types.

DISCUSSION

Methods of typing *C. albicans* isolated from infected patients are of use in answering such questions as whether a new episode of the disease is caused by reinfection or

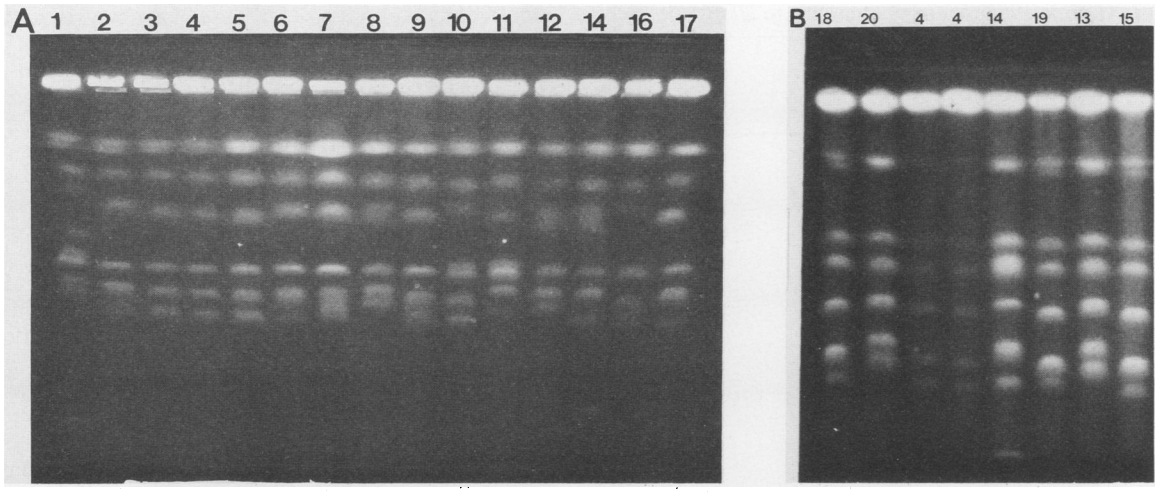

FIG. 3. Electrophoretic separation of chromosomes 4 through 7. The samples were run as described in Materials and Methods for these chromosomes. Numbers above the lanes indicate strains.

relapse, whether commensally carried strains are the source of infection, and whether there are particular strains adapted to particular body niches (e.g., the rectum compared with the vagina). Classic approaches to determining types, such as growth on a particular substrate or reactions to a particular antibiotic, are subject to the limitation that a specific phenotypic characteristic may be the result of one or several genotypes, and thus, strains that appear identical may in fact be different. Use of molecular approaches gets around this problem, since such approaches directly test genotypic character. No one approach (other than completely sequencing the genomes of two isolates) can ensure identity, but a combination of several such approaches can give a very high probability of detecting differences if any exist.

In these present studies, we utilized the information from two kinds of molecular studies previously described for these strains: the RFLPs generated from repeated DNA by digestion with *Eco*RI (14) and the RFLPs generated by the

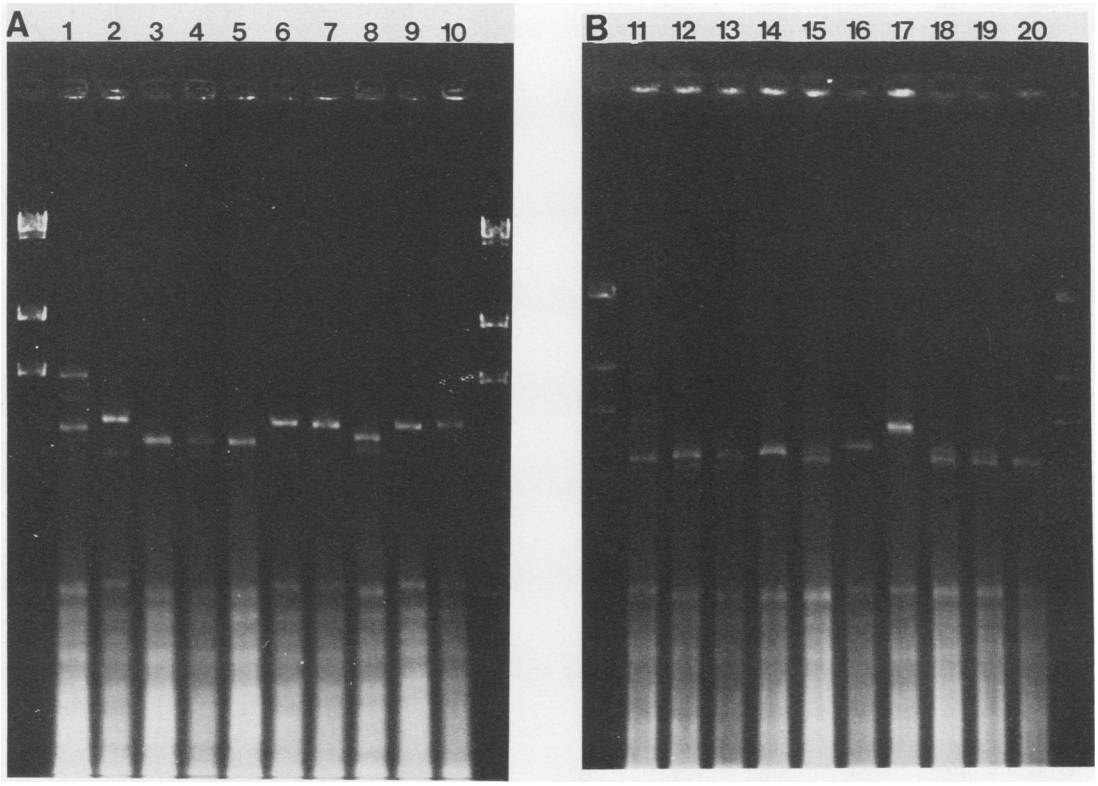

FIG. 4. *Hin*fI restriction fragments. DNA was isolated and digested as described in Materials and Methods. Digests were run on 0.7% agarose at 90 V and then stained with ethidium bromide. Numbers above the lanes refer to strains.

TABLE 2. Classification of strains according to *Hin*fI RFLPS

Strain	Major band		*Hin*fI type[b]
	Size (kbp)	Type[a]	
1	8.8, 6.3	h1	H1
2	7.0	h2	H2
3	5.7	h3	H3
4	5.7	h3	H4
5	5.7	h3	H5
6	7.0	h2	H2
7	7.0	h2	H7
8	6.0	h4	H8
9	7.0	h2	H2
10	7.0	h2	H7
11	4.4	h6	H9
12	4.7	h6	H9
13	4.7	h6	H9
14	5.1	h7	H10
15	4.7	h6	H11
16	5.3	h8	H12
17	6.0	h4	H13
18	4.7	h6	H11
19	4.7	h6	H11
20	4.7	h6	H9

[a] The major-band type is based on the size(s) of the brightest one or two bands on the gel.
[b] The *Hin*fI type takes into account the less brightly staining bands.

same enzyme from the chromosomal repeated element 27A (10). We compared the types generated from 20 strains by these methods with those generated by examination of the electrophoretic karyotype and by the RFLPs obtained from the rDNA by *Hin*fI digestion. With the application of each method, the subdivision into types increased. That is to say, only three of the strains were indistinguishable by any of the four analyses. This is perhaps not surprising, since the lack of genetic exchange in *C. albicans* means that each daughter cell is permanently isolated from all others, so that genomic changes are fixed. It may be that relatedness is proportional to the number of differences between strains, but that possibility remains to be tested. Such an interpretation would suggest that strains 15, 18, and 19 are most closely related, since they differ in none of the parameters tested. Pair 3 and 5 and pair 8 and 12 differ only in their *Hin*fI patterns, while 8, 12, and 13 differ only in karyotype.

It is interesting that the *Hin*fI RFLPs resolve more strains than do the *Eco*RI patterns, although only one or two sequences are involved. This is most likely due to the fact that the major *Hin*fI piece comes from the nontranscribed spacer in rDNA (4) and thus is subject to minimal constraints on its genetic drift. *Hin*fI digestion seems to be preferable to *Eco*RI treatment as a method for looking at stained gels, since the enzyme generates only one or two major bands and one or two lighter ones, well separated from the rest of the digestion products, and the method shows great resolving power.

Vasquez et al. (15) have suggested that karyotyping is preferable to RFLP analysis as a method of biotyping strains. They argue that karyotyping is more sensitive, more reproducible, and simpler to perform and interpret, than RFLP analysis, since the number of bands of the karyotype is much smaller than the number of bands of the enzyme digestion. One disadvantage of karyotyping which they do not mention is the time required; their procedure takes 54 h, while ours, which resolves the larger chromosomes more effectively, takes 5 days (120 h). RFLP analysis, on the other hand, take less than a day from colony to result. Furthermore, the number of samples that can be put on a simple electrophoresis gel (>40) is much larger than can effectively be run on a contour-clamped homogeneous electric-field instrument. The use of *Hin*fI as the enzyme for digestion greatly simplifies interpretation, so this approach seems to be the method of choice for stained RFLPs.

The use of a middle repetitive sequence like 27A to analyze RFLPs is a very sensitive method. Since *Eco*RI is the enzyme of choice for 27A, if this pattern is to be analyzed, the gel may as well be stained before being blotted in order to take maximum advantage of the information available from analysis of the RFLPs from the highly repeated DNA. In the case of the 20 strains analyzed in this paper, use of such information generates 13 types. Adding the information from contour-clamped homogeneous electric-field karyotyping generates 16 types; including *Hin*fI digestion results gives 18 types. If only RFLPs (*Eco*RI stained with 27A as a probe and *Hin*fI stained) are used, 17 types are generated. Thus, the karyotyping data do resolve some otherwise identical strains.

Molecular approaches to biotyping thus promise eventually to allow every isolate to be distinguished from every other isolate. At that point, the relatedness of the various strains, that is, whether one isolate was recently derived from another, must be decided. Such information will depend on determining the rate at which the parameters we measured (restriction sites, chromosome length, and chromosomal configuration of 27A sequences) change as a function of cell growth. This is an area for which little or no information is available.

ACKNOWLEDGMENTS

This work was supported by U.S. Public Health Service grant AI16567.

We are grateful to D. Stevens and S. Scherer for supplying their strains and some unpublished data and to B. B. Magee for a critical reading of the manuscript.

REFERENCES

1. **Carruba, G., E. Pontieri, F. De Bernardis, P. Martino, and A. Cassone.** 1991. DNA fingerprinting and electrophoretic karyotype of environmental and clinical isolates of *Candida parapsilosis*. J. Clin. Microbiol. **29**:916–922.
2. **Childress, C. M., I. A. Holder, and A. N. Neely.** 1989. Modifications of a *Candida albicans* biotyping system. J. Clin. Microbiol. **27**:1392–1394.
3. **Fox, B. C., H. L. Mobley, and J. C. Wade.** 1989. The use of a DNA probe for epidemiological studies of candidiasis in immunocompromised hosts. J. Infect. Dis. **159**:488–494.
4. **Magee, B. B., T. M. D'Souza, and P. T. Magee.** 1987. Strain and species identification by restriction fragment length polymorphisms in the ribosomal DNA repeat of *Candida* species. J. Bacteriol. **169**:1639–1643.
5. **Magee, B. B., and P. T. Magee.** 1987. Electrophoretic karyotypes and chromosome numbers in *Candida* species. J. Gen. Microbiol. **133**:425–430.
6. **Merz, W. G.** 1990. *Candida albicans* strain delineation. Clin. Microbiol. Rev. **3**:321–334.
7. **Merz, W. G., C. Connelly, and P. Hieter.** 1988. Variation of electrophoretic karyotype among clinical isolates of *Candida albicans*. J. Clin. Microbiol. **26**:842–845.
8. **Odds, F. C., P. Auger, P. Krogh, A. N. Neely, and E. Segal.** 1989. Biotyping of *Candida albicans*: results of an international collaborative survey. J. Clin. Microbiol. **27**:1506–1509.
9. **Scherer, S., and D. A. Stevens.** 1987. Application of DNA typing methods to epidemiology and taxonomy of *Candida* species. J. Clin. Microbiol. **25**:675–679.

10. **Scherer, S., and D. A. Stevens.** 1988. A *Candida albicans* dispersed, repeated gene family and its epidemiologic applications. Proc. Natl. Acad. Sci. USA **85**:1452–1456.
11. **Snell, R. G., I. F. Hermans, R. J. Wilkins, and B. E. Corner.** 1987. Chromosomal variations in *Candida albicans*. Nucleic Acids Res. **15**:3625.
12. **Soll, D. R., R. Galask, S. Isley, T. V. Rao, D. Stone, J. Hicks, J. Schmid, K. Mae, and C. Hanna.** 1989. Switching of *Candida albicans* during successive episodes of recurrent vaginitis. J. Clin. Microbiol. **27**:681–690.
13. **Soll, D. R., M. Staebell, C. Langtimm, M. Pfaller, J. Hicks, and T. V. Gopala Rao.** 1988. Multiple *Candida* strains in the course of a single systemic infection. J. Clin. Microbiol. **26**:1448–1459.
14. **Stevens, D. A., F. C. Odds, and S. Scherer.** 1990. Application of DNA typing methods to *Candida albicans* epidemiology and correlations with phenotype. Rev. Infect. Dis. **12**:258–266.
15. **Vasquez, J. A., A. Beckley, J. D. Sobel, and M. J. Zervos.** 1991. Comparison of restriction enzyme analysis and pulsed-field gradient gel electrophoresis as typing systems for *Candida albicans*. J. Clin. Microbiol. **29**:962–967.
16. **Wickes, B., J. Staudinger, B. B. Magee, K.-J. Kwon-Chung, P. T. Magee, and S. Scherer.** 1991. Physical and genetic mapping of *Candida albicans*: several genes previously assigned to chromosome 1 map to chromosome R, the rDNA-containing linkage group. Infect. Immun. **59**:2480–2484.