JOURNAL OF CLINICAL MICROBIOLOGY, Feb. 1998, p. 367–374 Vol. 36, No. 2
0095-1137/98/$04.00+0

Rapid Differentiation of Closely Related *Candida* Species and Strains by Pyrolysis-Mass Spectrometry and Fourier Transform-Infrared Spectroscopy

ÉADAOIN M. TIMMINS,[1] SUSAN A. HOWELL,[2] BJØRN K. ALSBERG,[1] WILLIAM C. NOBLE,[2] AND ROYSTON GOODACRE[1]*

Institute of Biological Sciences, University of Wales, Aberystwyth, Ceredigion SY23 3DA,[1] and Department of Microbial Diseases, St. John's Institute of Dermatology, St. Thomas' Hospital, London SE1 7EH,[2] United Kingdom

Received 19 June 1997/Returned for modification 18 August 1997/Accepted 17 October 1997

Two rapid spectroscopic approaches for whole-organism fingerprinting of pyrolysis-mass spectrometry (PyMS) and Fourier transform-infrared spectroscopy (FT-IR) were used to analyze a group of 29 clinical and reference *Candida* isolates. These strains had been identified by conventional means as belonging to one of the three species *Candida albicans*, *C. dubliniensis* (previously reported as atypical *C. albicans*), and *C. stellatoidea* (which is also closely related to *C. albicans*). To observe the relationships of the 29 isolates as judged by PyMS and FT-IR, the spectral data were clustered by discriminant analysis. On visual inspection of the cluster analyses from both methods, three distinct clusters, which were discrete for each of the *Candida* species, could be seen. Moreover, these phenetic classifications were found to be very similar to those obtained by genotypic studies which examined the *Hin*fI restriction enzyme digestion patterns of genomic DNA and by use of the 27A *C. albicans*-specific probe. Both spectroscopic techniques are rapid (typically, 2 min for PyMS and 10 s for FT-IR) and were shown to be capable of successfully discriminating between closely related isolates of *C. albicans*, *C. dubliniensis*, and *C. stellatoidea*. We believe that these whole-organism fingerprinting methods could provide opportunities for automation in clinical microbial laboratories, improving turnaround times and the use of resources.

The differentiation of *Candida* species has classically been performed on the basis of biochemical reactions and morphological features. However, phenotypic variations within some species such as *Candida albicans* (1, 6, 26) has led to much confusion in the classification of this species. The high degree of relatedness between *C. albicans* and *C. stellatoidea* illustrates this, and several workers have suggested that *C. stellatoidea* is synonymous with or a variant of *C. albicans* (3, 34). Typing systems based on DNA analysis have been introduced (38, 42), but these approaches are slow and labor-intensive and require considerable technical expertise.

In immunocompromised patients the clinical appearance of the *C. albicans* infection is often very complex and identification of the organism is difficult. However, speedy diagnosis and management of candidosis is crucial for these patients. The ideal method for the rapid and accurate identification of microorganisms, particularly in the clinical laboratory, would have minimum sample preparation, would analyze samples directly, and would be rapid, automated, accurate, and inexpensive (at least relatively inexpensive) (10). With recent developments in analytical instrumentation, these requirements are being fulfilled by physicochemical spectroscopic methods, often referred to as "whole-organism fingerprinting." The most common such methods are pyrolysis-mass spectrometry (PyMR) (8), Fourier transform-infrared (IR) spectroscopy (FT-IR) (13, 18), and UV resonance Raman spectroscopy (33).

PyMS involves the thermal degradation of complex material (such as bacteria or fungi) in a vacuum by Curie-point pyrolysis; this causes molecules to cleave at their weakest points to produce smaller, volatile fragments called pyrolysate (20). A mass spectrometer can then be used to separate the components of the pyrolysate on the basis of their mass-to-charge (m/z) ratios to produce a pyrolysis mass spectrum, which can then be used as a chemical fingerprint of the complex material analyzed (28). PyMS is well established within microbiology for the characterization of bacterial systems. In particular, the technique has been successful for the interstrain comparison of a variety of medically important organisms (see references 10, 16, and 24 for reviews).

In contrast to measuring the bond strengths of molecules, FT-IR spectroscopy measures vibrations of functional groups and highly polar bonds such as O-H stretches. Thus, these "fingerprints" are made up of the vibrational features of all the cell components, i.e., DNA, RNA, proteins, and membrane and cell-wall components (31). FT-IR allows the chemically based discrimination of intact microbial cells, without their destruction, and produces complex biochemical fingerprints which are reproducible and distinct for different bacteria and fungi. Naumann and coworkers (18, 30) have shown that FT-IR absorbance spectroscopy (in the mid-IR range, usually defined as 4,000 to 400 cm^{-1}) provides a powerful tool with sufficient resolving power to distinguish microbes at the strain level.

The aim of this study was to compare the phenotypic differentiation of *Candida* isolates by PyMS and diffuse reflectance-absorbance FT-IR with the differentiation based on genotypic investigations of the same isolates. Strain similarity was examined by using the *Hin*fI endonuclease restriction enzyme digestion patterns of genomic DNA, while strain and species identifications were confirmed by hybridization of *Eco*RI digests with the 27A *C. albicans*-specific probe. The 29 isolates studied had previously been identified by conventional means as be-

* Corresponding author. Mailing address: Institute of Biological Sciences, University of Wales, Aberystwyth, Ceredigion SY23 3DA, United Kingdom. Phone: 44 (0)1970 621947. Fax: 44 (0)1970 622354. E-mail: rrg@aber.ac.uk.

TABLE 1. Strain numbers of the three *Candida* species analyzed by PyMS together with their corresponding identifiers and source.

Species	Strain no.[a]	Source	Identifier[b]
C. albicans	NCPF 3153	Mycology Reference Laboratory antigen strain, serotype A, 1965	A
C. albicans	NCPF 3156	Serotype B, 1965	B
C. albicans	NCPF 3157	Serotype B, 1965	C
C. albicans	NCPF 3116	Human chronic glossitis, United Kingdom, 1944	D
C. albicans	NCPF 3119	Human systemic candidosis, skin isolate, United Kingdom, 1944	E
C. albicans	ATCC 18804	Type culture	F
C. albicans	Rld	Laboratory strain[c]	G
C. albicans	R8a large	Laboratory strain[c]	H
C. albicans	R12a	Laboratory strain[c]	I
C. albicans	R18a	Laboratory strain[c]	J
C. albicans	27544A large	Laboratory strain[c]	K
C. stellatoidea	ATCC 11006	Type I	1
C. stellatoidea	Y2360	Gift from B. B. Magee, University of Minnesota	2
C. dubliniensis	NCPF 3108	Human bronchopneumonia, lung isolate, United Kingdom 1957	a
C. dubliniensis	NCPF 3949	Middorsum of the tongue of an Irish subject with human immunodeficiency virus infection and suffering from oral candidosis, 1988, type culture	b
C. dubliniensis	R1a buff	Laboratory strain[c]	c
C. dubliniensis	R2g cream[d]	Laboratory strain[c]	d
C. dubliniensis	R2g white[d]	Laboratory strain[c]	e
C. dubliniensis	R3b	Laboratory strain[c]	f
C. dubliniensis	R3i	Laboratory strain[c]	g
C. dubliniensis	R9a	Laboratory strain[c]	h
C. dubliniensis	R9g large	Laboratory strain[c]	i
C. dubliniensis	R11b large	Laboratory strain[c]	j
C. dubliniensis	R16a	Laboratory strain[c]	k
C. dubliniensis	R16b white[e]	Laboratory strain[c]	l
C. dubliniensis	R16b buff[e]	Laboratory strain[c]	m
C. dubliniensis	716	Laboratory strain[c]	n
C. dubliniensis	43194A	Laboratory strain[c]	o
C. dubliniensis	$63861A_o$	Laboratory strain[c]	p

[a] NCPF, National Collection of Pathogenic Fungi; ATCC, American Type Culture Collection.
[b] Identifier used in Fig. 3.
[c] Laboratory strains recovered from oral swabs of patients attending dental clinics.
[d] Strains isolated from the same patient.
[e] Strains isolated from the same patient.

longing to *C. albicans*, *C. stellatoidea*, or *C. dubliniensis*. *C. dubliniensis* is a newly proposed strain of *Candida* (40) and has been reported in the past as atypical *C. albicans* (27, 39). Therefore, an additional aim was to investigate the taxonomic position of *C. dubliniensis.*

MATERIALS AND METHODS

Strains and cultivation. Twenty-nine *Candida* isolates, which comprised a selection of *C. albicans*, *C. dubliniensis*, and *C. stellatoidea* strains (see Table 1 for strain numbers and sources), were aerobically cultivated on LabM Malthus blood agar base (37 mg · ml^{-1}) for 16 h at 37°C. After subculturing three times to ensure pure cultures, the biomass was carefully collected with sterile plastic loops and was suspended in 1-ml aliquots of sterile physiological saline (0.9% NaCl).

All isolates were identified to the species level by using the biochemical profiles from the API 20C AUX (BioMerieux) kit. In addition, all isolates were tested for their ability to produce germ tubes and chlamydospores and for the ability to grow at 42°C; the latter was the distinguishing feature for *C. dubliniensis* (40). The isolates listed as *C. albicans* were all identified by the API kit (generally with >99.5% certainty), were all germ tube and chlamydospore positive, and grew at 42°C. The two isolates of *C. stellatoidea* were identified by the API kit as *C. albicans* 2, were germ tube and chlamydospore positive, but failed to grow at 42°C. Isolates of *C. dubliniensis* produced a variety of API profiles for *C. albicans* (generally with levels of 95 to 99% certainty), 11 of 16 were germ tube positive, 15 of 16 were chlamydospore positive, and none grew at 42°C.

DNA extraction. Isolates were grown overnight in 5 ml of Sabouraud broth (1% mycopeptone and 4% dextrose; Oxoid, Unipath, Basingstoke, United Kingdom) at 30°C. Cells from 1.5 ml of broth culture were pelleted in a microcentrifuge tube by centrifugation and were then extracted as described previously (19). Briefly, the cell pellet was washed and resuspended in 1 ml of buffer (1 M sorbitol, 50 mM potassium dihydrogen phosphate [pH 7.5]) containing 1 mg of zymolyase 20T (ICN Biomedicals, High Wycombe, United Kingdom) and 3 μl of β-mercaptoethanol. After 90 min of incubation at 37°C the spheroplasts were pelleted by centrifugation and were lysed by resuspension in 0.5 ml of GES (60% guanidine thiocyanate, 0.1 M EDTA, 0.5% lauroyl sarcosine [Sigma, Poole, United Kingdom]) reagent (35). After 20 min at room temperature, 100 μl of 5 M potassium acetate was added, the solution was placed on ice for 30 min, and 0.5 ml of chloroform-pentanol (24:1; vol/vol) was then added. The mixture was centrifuged for 5 min at 10,000 × *g* after which the upper layer was saved and the DNA was precipitated by the addition of an equal volume of absolute ethanol. The DNA was next pelleted, washed with 70% ethanol, dried, and resuspended in 100 μl of TE (10 mM Trizma base, 1 mM EDTA [pH 8]) for 30 min at 37°C. The DNA was then reprecipitated, pelleted, and finally dissolved in 50 μl of TE and stored at −20°C until it was used.

Restriction endonuclease digestion. Aliquots of 20 μl of DNA solution were digested according to the manufacturer's instructions with either *Eco*RI (Gibco BRL, Uxbridge, United Kingdom) or *Hin*fI (Pharmacia, St. Albans, United Kingdom) for 4 h at 37°C. The DNA fragments were separated on 0.8% agarose gels immersed in TBE (89 mM Trizma, 32 mM boric acid, 2.5 mM EDTA) at 30 V for 16 h. The gels were stained with ethidium bromide, and the results were recorded photographically.

Southern blotting. The *Eco*RI restriction digests were capillary blotted onto a Stratagene (Cambridge, United Kingdom) Duralon UV nylon membrane by following a previously described procedure (36), and the DNA was fixed by UV cross-linking in a UV Stratalinker 1800 (Stratagene). Hybridization and detection were performed by following previously described procedures (2), with the exceptions that fat-free dried milk (0.2%; wt/vol) replaced sheared herring sperm in the hybridization solutions and the 27A probe was biotin labelled by nick translation with the BioNick Labelling System (Gibco BRL).

PyMS. Five-microliter aliquots of the yeast suspensions described above were evenly applied to clean iron-nickel foils which had been partially inserted into clean pyrolysis tubes. Samples were run in triplicate. Prior to pyrolysis the samples were oven dried at 50°C for 30 min, and the foils were then pushed into the tubes by using a stainless steel depth gauge so that the foils were 10 mm from the mouth of the tube. Viton O rings were next placed approximately 1 mm from the mouth of each tube.

PyMS was then performed on a Horizon Instrument PyMS-200X (Horizon Instruments Ltd., Heathfield, United Kingdom). Full operational procedures are described elsewhere (11, 12, 14, 41). The conditions used for each experiment involved heating the sample to 100°C for 5 s, followed by Curie-point pyrolysis at 530°C for 3 s with a temperature rise time of 0.5 s.

PyMS data may be displayed as quantitative pyrolysis-mass spectra (e.g., as in Fig. 1). The abscissa represents the 150 *m/z* ratios, while the ordinate contains information on ion count for any particular *m/z* value ranging from 51 to 200. Data were normalized as a percentage of the total ion count to remove the influence of sample size per se.

The normalized data for the 29 isolates (Table 1) were processed with the GENSTAT package (32) which runs under Microsoft DOS, version 6.2, on an International Business Machines (IBM)-compatible personal computer. This method has been described previously (17). The initial stage involved the reduction of the data by principal component analysis (PCA) (5, 21); this is a well-known technique for reducing the dimensionality of multivariate data while preserving most of the variance. Data were preserved by keeping only those principal components (PCs) whose eigenvalues accounted for more than 0.1% of the total variance. Discriminant function analysis (DFA) then discriminated between groups on the basis of the retained PCs and the a priori knowledge of which spectra were replicates (23, 43). The next stage involved the construction of a percent similarity matrix by transforming Mahalanobis's distance between a priori groups in DFA with the Gower similarity coefficient S_G (15).

Diffuse reflectance-absorbance FT-IR. Ten-microliter aliquots of the yeast suspensions were evenly applied onto a sand-blasted aluminum plate. Prior to analysis the samples were oven dried at 50°C for 30 min. Samples were run in triplicate. The FT-IR instrument used was the Bruker IFS28 FT-IR spectrometer (Bruker Spectrospin Ltd., Coventry, United Kingdom) equipped with a mercury-cadmium-telluride detector cooled with liquid N_2. The aluminum plate was then loaded onto the motorized stage of a reflectance thin-layer chromatography accessory (4, 7, 29).

The IBM-compatible personal computer used to control the IFS28 spectrometer was also programmed (using Opus, version 2.1, software running under IBM O/S2 Warp provided by the manufacturers) to collect spectra over the wave number range of 4,000 to 600 cm^{-1}. Spectra were acquired at a rate of 20 s^{-1}. A spectral resolution of 4 cm^{-1} was used. To improve the signal-to-noise ratio, 256 spectra were coadded and averaged. Each sample was thus represented by a spectrum containing 882 points, and spectra were displayed in terms of the absorbance calculated from the reflectance-absorbance spectra using the Opus software. Typical FT-IR spectra are shown in Fig. 2.

ASCII data were exported from the Opus software used to control the FT-IR instrument and imported into Matlab, version 4.2c.1 (The MathWorks, Inc., Natick, Mass.), which runs under Microsoft Windows NT on an IBM-compatible personal computer. To minimize problems arising from baseline shifts, the following procedure was implemented. (i) The spectra were first normalized so that the smallest absorbance was set to 0 and the highest absorbance was set to +1 for each spectrum; (ii) next, these normalized spectra were detrended by subtracting a linearly increasing baseline from 4,000 to 600 cm^{-1}; and (iii) finally, the smoothed first derivatives of these normalized and detrended spectra were calculated by using the Savitzky-Golay algorithm (37) with 5-point smoothing.

To reduce the dimensionality of the FT-IR data, Matlab was also used to perform PCA (according to the NIPALS algorithm [44]); of the original 882 spectral points, 97.5% of the total variance was retained in the first 15 PCs, illustrating the power of this technique. Next these 15 PCs were used as inputs to the DFA algorithm with the a priori knowledge of which spectra were replicates. DFA was programmed by Bjørn K. Alsberg according to the principles of Manly (25). Finally, the Euclidean distance between a priori group centers in DFA space (using the first eight discriminant functions) was used to construct a similarity measure, and these distance measures were then processed by an agglomerative clustering algorithm to construct a dendrogram (25).

RESULTS AND DISCUSSION

Examples of PyMS and FT-IR spectra for *C. albicans*, *C. stellatoidea*, and *C. dubliniensis* are shown in Fig. 1 and 2, respectively. For PyMS there was very little qualitative difference between these spectra, although on closer inspection quantitative differences may be observed. The FT-IR spectra all show broad and complex contours, and again, there was relatively little qualitative difference between the spectra. Such spectra readily illustrate the need to use multivariate statistical techniques in the analysis of both PyMS and FT-IR data.

The next stage was to use DFA to observe the relationships between these yeasts as judged from their PyMS and FT-IR spectra. The 87 spectra were coded so as to give 29 groups, 1 for each isolate (see Table 1), and the PyMS and FT-IR data were analyzed by DFA as detailed above. The resulting ordination plots for all 29 isolates (see Table 1 for identifiers) based on PyMS data and FT-IR data are shown in Fig. 3A and B, respectively. Both plots show the strains grouped into the same three main clusters: cluster 1 comprised all the *C. albicans* strains except *C. albicans* R8a large; cluster 2 was a rather loose cluster which included all the *C. dubliniensis* strains and *C. albicans* R8a large; finally, cluster 3 contained both *C. stellatoidea* strains. On closer inspection, cluster 2 can be seen to contain three subclusters: subcluster 2a comprised *C. dubliniensis* NCPF 3949, R1a buff, R3b, R2g cream, R2g white, R3i, R11b large, R16a, R16b buff, R16b white, 43194A, and 63861A_o and *C. albicans* R8a large; subcluster 2b contained *C. dubliniensis* R9a, R9g large, and 716; finally, subcluster 2c was a single-member cluster (SMC) of *C. dubliniensis* NCPF 3108.

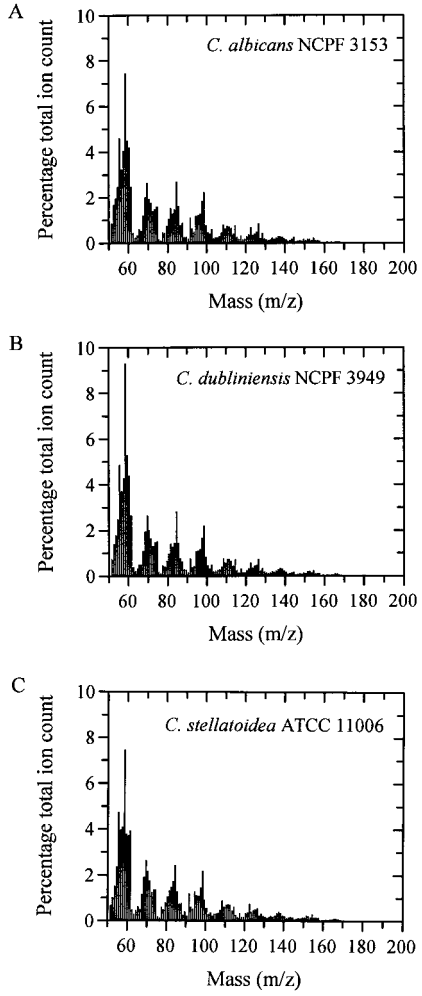

FIG. 1. Normalized pyrolysis-mass spectra of *C. albicans* NCPF 3153 (A), *C. dubliniensis* NCPF 3949 (B), and *C. stellatoidea* ATCC 11006 (C).

An alternative way of viewing the relationship between these isolates is by using hierarchical cluster analyses, and the resulting dendrograms are shown in Fig. 4 for both the PyMS and FT-IR data. This plot shows that there was indeed a great deal of congruence between both phenotypic analysis methods since the same three main clusters, as described above (Fig. 3), can be seen. This is very encouraging when one considers that although both methods fall within the framework of whole-organism fingerprinting and so give a phenotypic measure of the total biochemical makeup of the cells, they are doing so by measuring different physicochemical aspects. That is, PyMS gives a measure of the strengths of covalent bonds between molecules, while FT-IR measures the vibrations of functional groups and highly polar bonds.

In contrast to giving a measure of the phenotypes of the

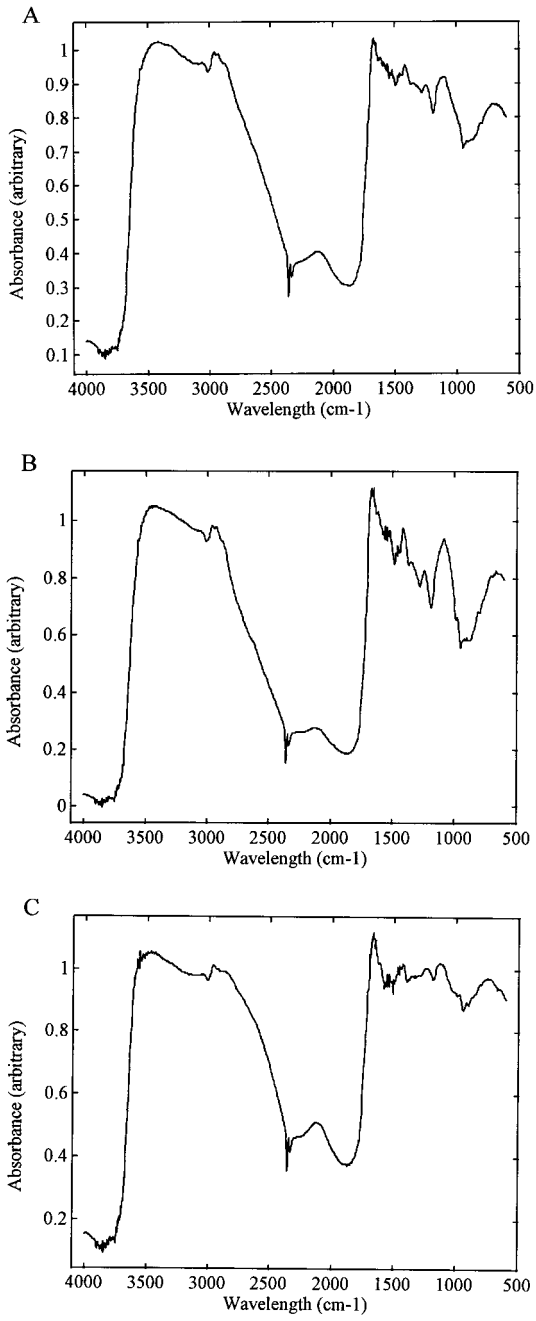

FIG. 2. FT-IR diffuse reflectance-absorbance spectra of *C. albicans* NCPF 3153 (A), *C. dubliniensis* NCPF 3949 (B), and *C. stellatoidea* ATCC 11006 (C).

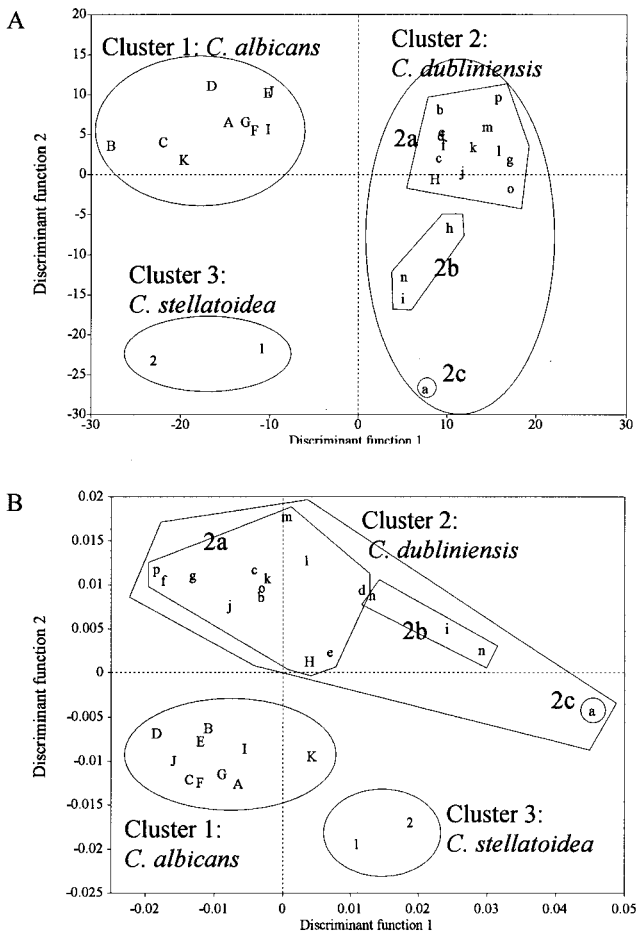

FIG. 3. Discriminant analysis plots based on PyMS data (A) and FT-IR data (B) showing the relationship between the 29 *Candida* strains. Analyses were conducted as described in Materials and Methods.

Candida cells, restriction enzyme digestion analysis of chromosomal DNA provides an indication of the genotypic relationship between the isolates. Therefore, genomic DNA from the 29 isolates was digested with the restriction endonuclease *Hin*fI, and the fragments were separated by agarose gel electrophoresis as detailed above.

Direct visual analysis of the band patterns (Fig. 5A and 5B) very easily allowed the *C. dubliniensis* strains to be differentiated from the *C. albicans* isolates. It can be seen (Fig. 5B) that the *C. dubliniensis* isolates produced quite a distinct pattern which was highly conserved, even though these isolates were recovered from different patients. On closer inspection, isolates R9a, R9g large, and 716 produced patterns which were similar to each other but which were slightly different from those for the other *C. dubliniensis* isolates. This was also seen in phenotypic analyses (Fig. 3A and B) in which these isolates were recovered together in cluster 2, subcluster 2b (see below). Figure 5 also shows that the pattern for *C. dubliniensis* NCPF 3108 more closely resembles the typical *C. dubliniensis* patterns, although some bands are missing; this isolate was originally reported to be *C. stellatoidea* (39); however, Sullivan and colleagues (40) have since assigned this isolate to the species *C. dubliniensis*. The PyMS and FT-IR analyses showed that this strain only loosely clustered with the other *C. dubliniensis* isolates and was more likely to be an intermediate between *C. dubliniensis* and *C. stellatoidea*. Therefore, it would appear that the exact taxonomic position of NCPF 3108 is still unclear.

The two *C. stellatoidea* banding patterns were very similar to one another (Fig. 5A), and when the lower-size bands are studied, these two patterns are more like those for *C. albicans* than those for *C. dubliniensis*; this result seems plausible since Kwon-Chung et al. (22) have referred to *C. stellatoidea* isolates as being sucrose-negative variants of *C. albicans*. However, the fact that the PyMS and FT-IR results indicated that *C. stellatoidea* is wholly distinct from both *C. albicans* and *C. dubliniensis* would suggest that visual inspection of banding patterns is often subjective.

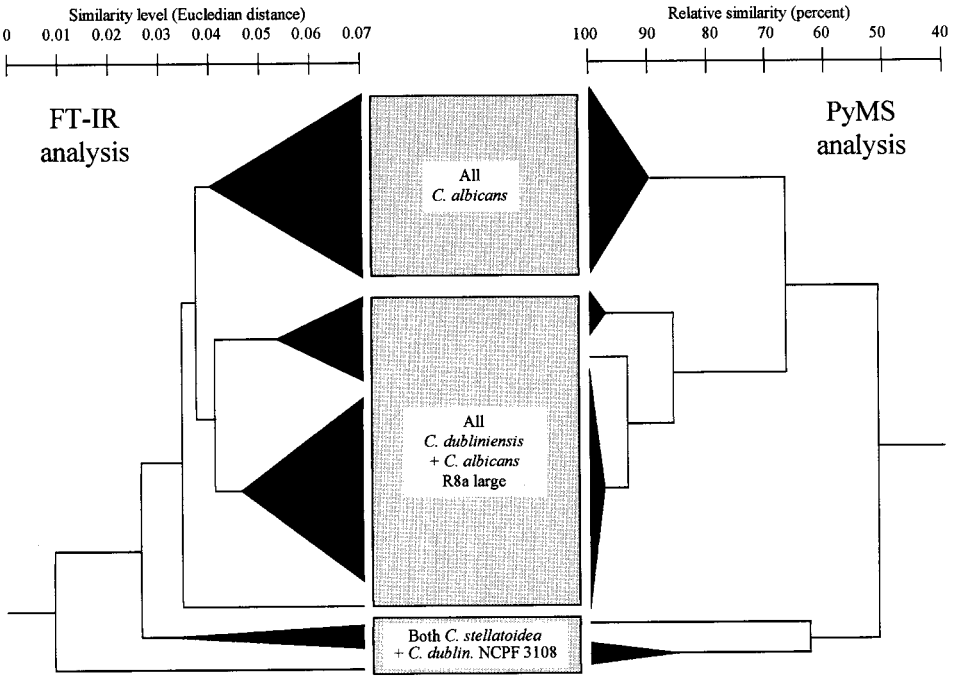

FIG. 4. Comparison of the 29 *Candida* strains clustered by using hierarchical cluster analyses; the dendrograms are from FT-IR data and PyMS data.

In contrast to the rather homogeneous nature of the *Hin*fI banding patterns for *C. dubliniensis*, *C. albicans* showed much more variation and no two patterns were the same (Fig. 5A). All *C. albicans* strains showed the similar feature of bands in the 4.4- to 6.2-kb region which were absent from the patterns for *C. stellatoidea* and *C. dubliniensis*, although *C. dubliniensis* R9a, R9g large, and 716 possessed a common band at approximately 4 kb.

Overall, the digestion patterns (Fig. 5) indicate that *C. dubliniensis* strains are very similar to each other, while PyMS and FT-IR detected differences and each produced a broad cluster (cluster 2; Fig. 3A and B). Conversely, the digests in Fig. 5A indicate that *C. albicans* isolates are diverse, while PyMS and FT-IR each produced a smaller cluster (cluster 1; Fig. 3A and B).

C. albicans R8a large, which possessed *Hin*fI banding patterns similar to those for *C. albicans*, was much more closely related to *C. dubliniensis* in both phenotypic analyses; Fig. 3 and 4 indicate this strain was recovered in the *C. dubliniensis* groups by both PyMS and FT-IR. Moreover, conventional testing also identified this isolate as *C. albicans*. To investigate this anomaly further, this isolate was subsequently reanalyzed by all techniques, and the results described above were found to be consistent. Identification was then confirmed by hybridizing *Eco*RI digests from some strains of each of the three *Candida* species with the *C. albicans*-specific 27A probe. Use of this probe was how atypical *C. albicans* isolates (later to be called *C. dubliniensis*) were first identified (27, 39). Figure 6A shows the *Eco*RI digestion patterns and Fig. 6B shows the corresponding hybridization patterns which clearly distinguish the three *Candida* species. Three separately stored subcultures of R8a large are shown on these gels, and they include subcultures from the St. Thomas Hospital stock (lane 5), the University of Wales stock (lane 6), and the sample after recovery from analysis by both PyMS and FT-IR (lane 7). From Fig. 6B it can be seen that the hybridization patterns for these subcultures were identical to one another and, more importantly, closely resemble the *C. albicans* profile. There are still two schools of thought as to how microorganisms should be classified: some believe that the way forward is to study the microbes's genotype; others pursue phenetic classifications. Few studies have exploited both approaches; often when this is done the results obtained are directly complementary (9) and both schools are happy, but sometimes, conflicting results are seen. For isolate R8a large the latter is unfortunately the case, and its identity remains unclear when both the phenotype and genotype are examined. The 27A repeat sequence probe for *C. albicans* shows that its genotype more closely resembles that of *C. albicans*; by contrast, however, the phenotypic experiments, which measure the expressed genotype, show unequivocally and reproducibly that R8a large clusters with *C. dubliniensis*. It is likely that as more taxonomic studies pursue both the phenetic and phylogenetic approaches other discrepancies will arise. Indeed, two organisms that are isogenic may be adapting (at the protein transcriptional level) to live in different ecological niches or host environments and will therefore display different phenotypes.

Finally, to investigate the finer relationships between the isolates recovered in clusters 1 and 2 from the PyMS analysis, multivariate statistical analysis was performed with these strains only, and the resulting dendrograms can be seen in Fig. 7 and 8. The dendrogram from cluster 1 (Fig. 7) showed that at 85% similarity three clusters could be seen: subcluster 1a contained *C. albicans* NCPF 3153, R12a, and Rld, which clustered together with >95% relative similarity, and these grouped with *C. albicans* NCPF 3156 at 87% similarity; subcluster 1b comprised *C. albicans* NCPF 3157, NCPF 3119, R18a, and ATCC 18804, which had >95% relative similarity, and these strains were 90.2% similar to *C. albicans* NCPF 3116; finally, subcluster 1c was an SMC consisting of *C. albicans* 27544A large. Similarly, the dendrogram from cluster 2 (Fig. 8) showed that at 90% similarity three clusters could be seen, and these mirrored those from DFA of all 29 *Candida* isolates (Fig. 3): subcluster 2a comprised two subgroups (which were formed at

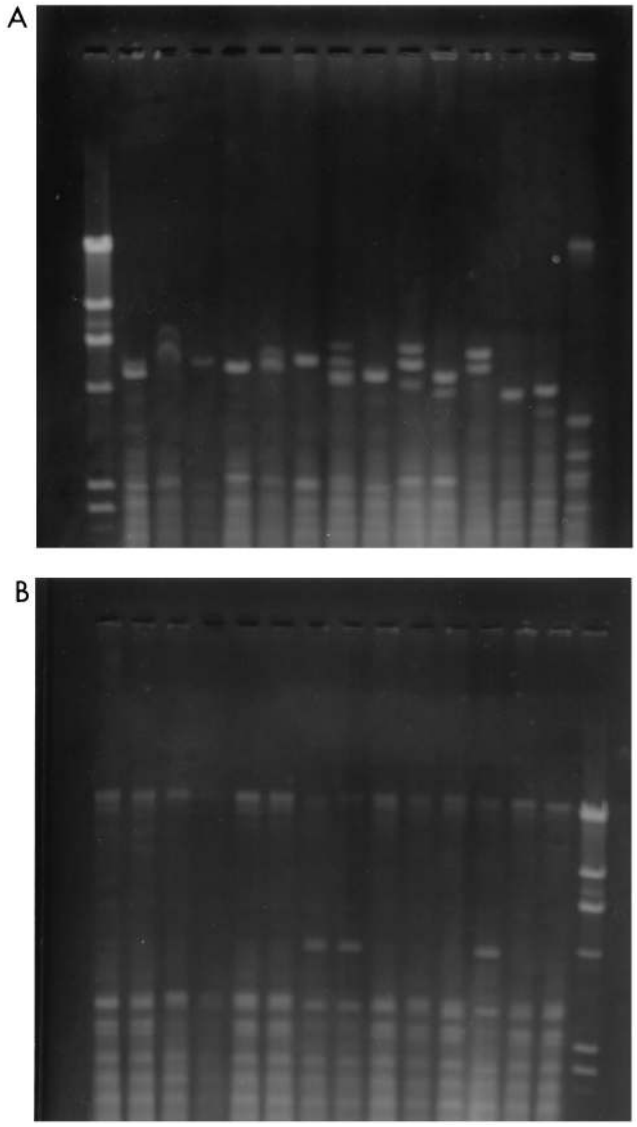

FIG. 5. *Hin*fI restriction patterns of genomic DNA. (A) Lane M (lanes are from left to right, respectively), bacteriophage λ *Hin*dIII marker; lanes 1 to 11, *C. albicans* isolates; lane 1, NCPF 3116; lane 2, NCPF 3119; lane 3, NCPF 3153; lane 4, NCPF 3156; lane 5, NCPF 3157; lane 6, ATCC 18804; lane 7, R1d; lane 8, R8a large; lane 9, R12a; lane 10, R18a; lane 11, 27544A; lanes 12 and 13, *C. stellatoidea* isolates; lane 12, ATCC 11006; lane 13, Y2360; lane 14, *C. dubliniensis* NCPF 3108. (B) Lanes 1 to 14 (lanes from left to right, respectively), *C. dubliniensis* isolates; lane 1, NCPF 3949; lane 2, R1a buff; lane 3, R2g cream; lane 4, R2g white; lane 5, R 3b; lane 6, R3i; lane 7, R9a; lane 8, R9g large; lane 9, R11b large; lane 10, R16a; lane 11, R16b white; lane 12, 716; lane 13, 43194A; lane 14, 63861A_o; lane M, λ *Hin*dIII marker. λ *Hin*dIII marker sizes of 23.1, 9.4, 6.6, 4.4, 2.3, and 2.0 kb are shown.

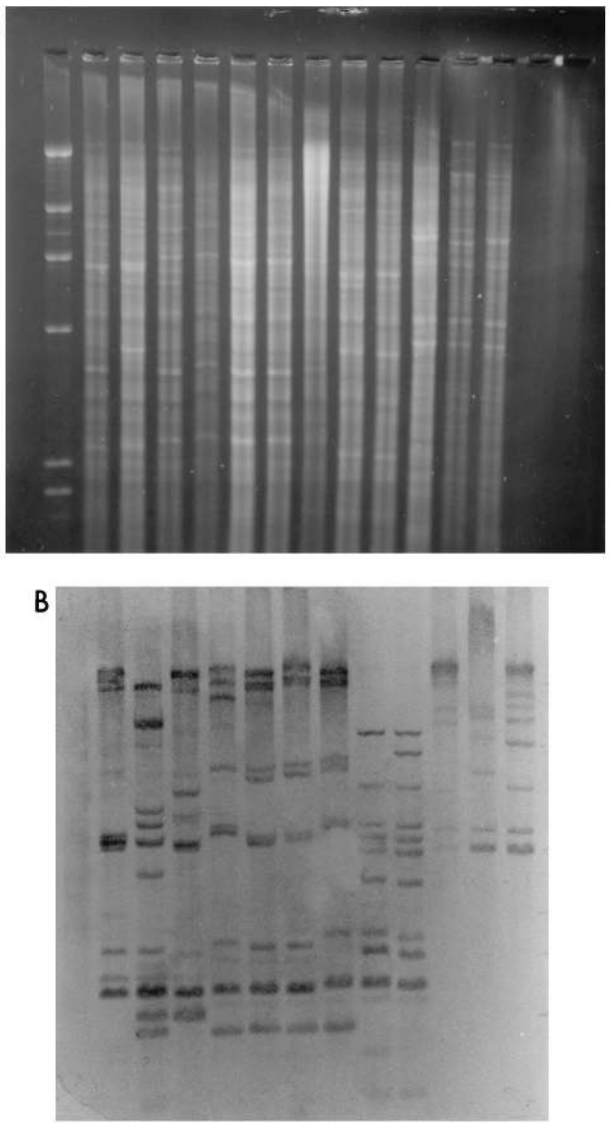

FIG. 6. *Eco*RI restriction patterns (A) and the corresponding hybridization patterns obtained with the *C. albicans* specific 27A probe (B). Lane M (lanes are from left to right, respectively), λ *Hin*dIII marker; lanes 1 to 7, *C. albicans* isolates; lane 1, NCPF 3153; lane 2, ATCC 18804; lane 3, R18a; lane 4, R8a small; lane 5, R8a large; lane 6, R8a large; lane 7, R8a large; lanes 8 and 9, *C. stellatoidea* isolates; lane 8, ATCC 11006; lane 9, Y2360; lanes 10 to 12, *C. dubliniensis* isolates; lane 10, NCPF 3108; lane 11, NCPF 3949; lane 12, R2g cream. λ *Hin*dIII marker sizes of 23.1, 9.4, 6.6, 4.4, 2.3, and 2.0 kb are shown in lane 1 of panel A.

92% similarity); the first comprised *C. albicans* R8a large and *C. dubliniensis* NCPF 3949, R1a buff, R3b, R2g cream, R2g white, while the second included *C. dubliniensis* R3i, R11b large, R16a, R16b buff, R16b white, 43194A, and 63861A_o; subcluster 2b contained *C. dubliniensis* R9a, R9g large, and 716; finally, subcluster 2c was an SMC consisting of *C. dubliniensis* NCPF 3108.

It was also very encouraging that duplicate cultures of *C. dubliniensis* from the same patients were recovered together, highlighting the reproducibility of these techniques. For example, it can be seen in Fig. 8 that the colonial color variants R2g cream and R2g white were recovered together in subcluster 2a′, while R16b buff and R16b white were recovered in subcluster 2a″.

Similar results were obtained by FT-IR (data not shown), demonstrating that both techniques are reproducible and have the ability to differentiate between *C. albicans* and *C. dubliniensis*. Moreover, these PyMS and FT-IR analyses were performed in triplicate on the day of assay and were repeated on two separate occasions and provided the same results, showing that these methods are indeed highly reproducible.

Concluding remarks. Twenty-nine isolates of *Candida* were classified by conventional means as being one of the following species: *C. albicans*, *C. dubliniensis*, or *C. stellatoidea*. These isolates were then analyzed by the two whole-organism finger-

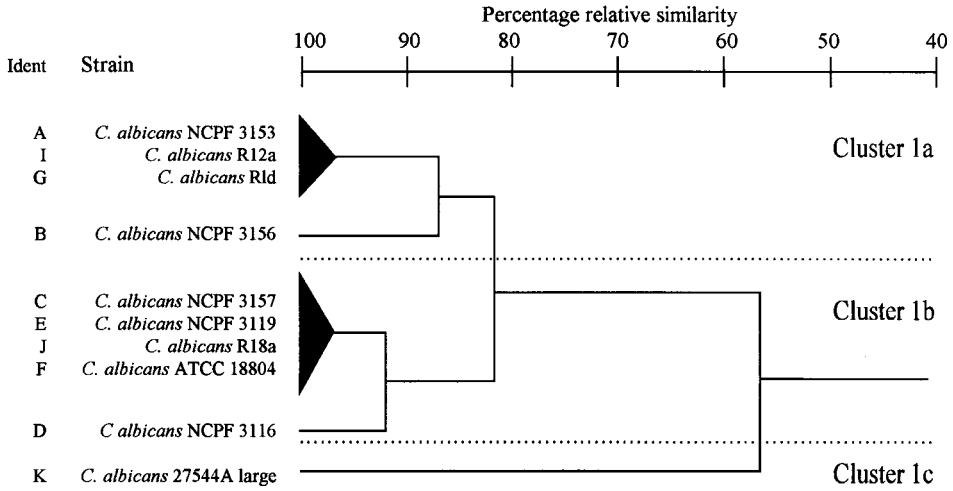

FIG. 7. Dendrogram representing the relationships between all the *C. albicans* strains (except R8a large) on the basis of PyMS data analyzed by GENSTAT.

printing techniques of PyMS and FT-IR, and their genotypes were examined by using restriction endonuclease digestion with the enzyme *Hin*fI followed by separation by gel electrophoresis. The identities of strains from each species were also confirmed by hybridization of *Eco*RI digests with the 27A *C. albicans*-specific probe.

Cluster analysis of the PyMS and FT-IR spectral data showed that three distinct groups were seen to be discrete for each of the *Candida* species. Moreover, these phenetic classifications were found to be very similar to those obtained by using *Hin*fI digestion patterns. The important conclusion from this study is that *C. dubliniensis* and *C. albicans* are indeed separate species, as judged by their genotypes and phenotypes. Finer discriminatory analyses showed that both spectroscopic methods could be used for the differentiation of *C. albicans* and *C. dubliniensis* isolates to below the subspecies level. However, subspecies analysis of *C. albicans* was found to be more successful when genotypic methods were used, while for *C. dubliniensis*, PyMS and FT-IR offer a more interpretable means of subspecies identification. Finally, duplicate cultures

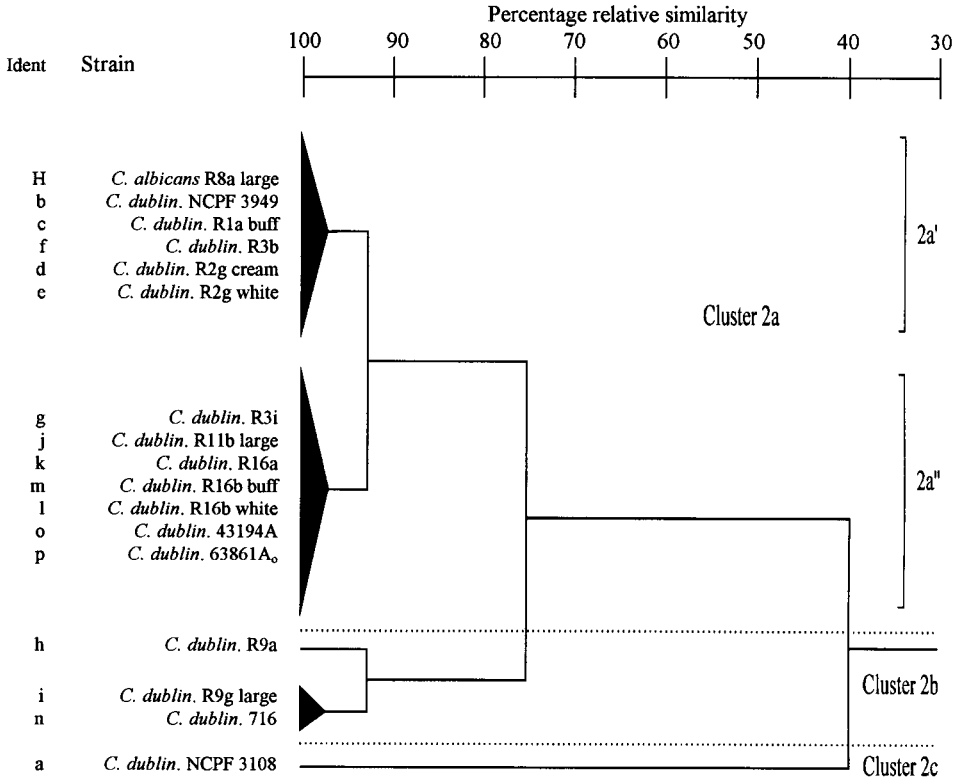

FIG. 8. Dendrogram representing the relationships between all the *C. dubliniensis* (*C. dublin.*) strains tested and *C. albicans* R8a large based on PyMS data analyzed by GENSTAT.

were recovered together, showing that all methods are highly reproducible.

The application of PyMS and FT-IR to microbiology is undoubtedly useful for the discrimination between these *Candida* species at the species and subspecies levels. Both techniques have the major advantages of speed, sensitivity, and the ability to analyze many hundreds of samples per day. We therefore conclude that these whole-organism fingerprinting methods could provide in the future opportunities for automation in the clinical microbiology laboratory, depending on the location of the laboratory and access to suitable support facilities. These new techniques are rapid and accurate and present the potential for investigating outbreaks of infection almost as they occur.

ACKNOWLEDGMENTS

We thank Douglas B. Kell for use of PyMS and FT-IR, M. J. Cunningham for the API data, and B. B. Magee for the 27A probe and isolate Y2360.

R.G. and E.M.T. are indebted to the Wellcome Trust for financial support (grant 042615/Z/94/Z). B.K.A. thanks the Chemicals and Pharmaceuticals Directorate of the UK BBSRC for financial support.

REFERENCES

1. **Allen, C. M., and F. M. Beck.** 1983. Strain related differences in pathogenicity of *Candida albicans* for oral mucosa. J. Infect. Dis. **147:**1036–1040.
2. **Anthony, R. M., J. Midgley, S. P. Sweet, and S. A. Howell.** 1995. Multiple strains of *Candida albicans* in the oral cavity of HIV positive and HIV negative patients. Microb. Ecol. Health Dis. **8:**23–30.
3. **Barnett, J. A., R. W. Payne, and D. Yarrow.** 1990. Descriptions of the species arranged alphabetically, p. 79–695. Yeasts: characteristics and identification, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.
4. **Bouffard, S. P., J. E. Katon, A. J. Sommer, and N. D. Danielson.** 1994. Development of microchannel thin layer chromatography with infrared microspectroscopic detection. Anal. Chem. **66:**1937–1940.
5. **Causton, D. R.** 1987. A biologist's advanced mathematics. Allen and Unwin, London, United Kingdom.
6. **Ghannoum, M. A., I. Swairjo, and D. R. Soll.** 1990. Variation in lipid and sterol contents in *Candida albicans* white and opaque phenotypes. J. Med. Vet. Mycol. **28:**103–115.
7. **Glauninger, G., K. A. Kovar, and V. Hoffmann.** 1990. Possibilities and limits of an online coupling of thin-layer chromatography and FTIR spectroscopy. Fresenius' J. Anal. Chem. **338:**710–716.
8. **Goodacre, R.** 1994. Characterisation and quantification of microbial systems using pyrolysis mass spectrometry: introducing neural networks to analytical pyrolysis. Microbiol. Eur. **2**(2)**:**16–22.
9. **Goodacre, R., A. Hartmann, J. E. Beringer, and R. C. W. Berkeley.** 1991. The use of pyrolysis mass spectrometry in the characterization of *Rhizobium meliloti*. Lett. Appl. Microbiol. **13:**157–160.
10. **Goodacre, R., and D. B. Kell.** 1996. Pyrolysis mass spectrometry and its applications in biotechnology. Curr. Opin. Biotechnol. **7:**20–28.
11. **Goodacre, R., M. J. Neal, and D. B. Kell.** 1994. Rapid and quantitative analysis of the pyrolysis mass spectra of complex binary and tertiary mixtures using multivariate calibration and artificial neural networks. Anal. Chem. **66:**1070–1085.
12. **Goodacre, R., M. J. Neal, D. B. Kell, L. W. Greenham, W. C. Noble, and R. G. Harvey.** 1994. Rapid identification using pyrolysis mass spectrometry and artificial neural networks of *Propionibacterium acnes* isolated from dogs. J. Appl. Bacteriol. **76:**124–134.
13. **Goodacre, R., É. M. Timmins, P. J. Rooney, J. J. Rowland, and D. B. Kell.** 1996. Rapid identification of *Streptococcus* and *Enterococcus* species using diffuse reflectance-absorbance Fourier transform infrared spectroscopy and artificial neural networks. FEMS Microbiol. Lett. **140:**233–239.
14. **Goodacre, R., S. Trew, C. Wrigley-Jones, M. J. Neal, J. Maddock, T. W. Ottley, N. Porter, and D. B. Kell.** 1994. Rapid screening for metabolite overproduction in fermentor broths using pyrolysis mass spectrometry with multivariate calibration and artificial neural networks. Biotechnol. Bioeng. **44:**1205–1216.
15. **Gower, J. C.** 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika **53:**325–338.
16. **Gutteridge, C. S.** 1987. Characterization of microorganisms by pyrolysis mass spectrometry. Methods Microbiol. **19:**227–272.
17. **Gutteridge, C. S., L. Vallis, and H. J. H. MacFie.** 1985. Numerical methods in the classification of microorganisms by pyrolysis mass spectrometry, p. 369–401. *In* M. Goodfellow, D. Jones, and F. Priest (ed.), Computer-assisted bacterial systematics. Academic Press, London, United Kingdom.
18. **Helm, D., H. Labischinski, G. Schallehn, and D. Naumann.** 1991. Classification and identification of bacteria by Fourier transform infrared spectroscopy. J. Gen. Microbiol. **137:**69–79.
19. **Howell, S. A., R. M. Anthony, and E. Power.** 1996. Application of RAPD and restriction enzyme analysis to the study of oral carriage of *Candida albicans*. Lett. Appl. Microbiol. **22:**125–128.
20. **Irwin, W. J.** 1982. Analytical pyrolysis: a comprehensive guide. Marcel Dekker, Inc., New York, N.Y.
21. **Jolliffe, I. T.** 1986. Principal component analysis. Springer-Verlag, New York, N.Y.
22. **Kwon-Chung, K. J., B. L. Wickes, and W. G. Merz.** 1988. Association of electrophoretic karyotype of *Candida stellatoidea* with virulence for mice. Infect. Immun. **56:**1814–1819.
23. **MacFie, H. J. H., C. S. Gutteridge, and J. R. Norris.** 1978. Use of canonical variates in differentiation of bacteria by pyrolysis gas-liquid chromatography. J. Gen. Microbiol. **104:**67–74.
24. **Magee, J. T., J. S. Brazier, I. K. Hosein, C. D. Ribeiro, D. W. Hill, A. Griffiths, C. Dacosta, A. J. Sinclair, and B. I. Duerden.** 1993. An investigation of a nosocomial outbreak of *Clostridium difficile* by pyrolysis mass spectrometry. J. Med. Microbiol. **39:**345–351.
25. **Manly, B. F. J.** 1994. Multivariate statistical methods: a primer. Chapman & Hall, London, United Kingdom.
26. **Martinez, J. P., L. Gill, M. Casanova, J. Ribot-Lopez, J. G. De Lomas, and R. Sentandreu.** 1990. Wall mannoproteins in cells from colonial phenotypic variants of *Candida albicans*. J. Gen. Microbiol. **136:**2421–2432.
27. **McCullough, M., B. Ross, and P. Reade.** 1995. Characterization of genetically distinct subgroup of *Candida albicans* strains isolated from oral cavities of patients infected with human immunodeficiency virus. J. Clin. Microbiol. **33:**696–700.
28. **Meuzelaar, H. L. C., J. Haverkamp, and F. D. Hileman.** 1982. Pyrolysis mass spectrometry of recent and fossil biomaterials. Elsevier, Amsterdam, The Netherlands.
29. **Mitchell, M. B.** 1993. Fundamentals and applications of diffuse reflectance infrared fourier transform (DRIFT) spectroscopy. Adv. Chem. Ser. **236:**351–375.
30. **Naumann, D., D. Helm, H. Labischinski, and P. Giesbrecht.** 1991. The characterization of microorganisms by Fourier-transform infrared spectroscopy (FT-IR), p. 43–96. *In* W. H. Nelson (ed.), Modern techniques for rapid microbiological analysis. VCH Publishers, New York, N.Y.
31. **Naumann, D., D. Helm, and C. Schultz.** 1994. Characterization and identification of micro-organisms by FT-IR spectroscopy and FT-IR microscopy. *In* F. G. Priest, A. Ramos-Cormenzana, and B. J. Tindall (ed.), Bacterial diversity and systematics (FEMS Symposium No. 75). Plenum Press, New York, N.Y.
32. **Nelder, J. A.** 1979. Genstat reference manual. Scientific and Social Service Program Library, University of Edinburgh, Edinburgh, United Kingdom.
33. **Nelson, W. H., R. Manoharan, and J. F. Sperry.** 1992. UV resonance Raman studies of bacteria. Appl. Spectrosc. Rev. **27:**67–124.
34. **Odds, F. C.** 1988. Biological aspects of pathogenic *Candida* species, p. 7–15. *In* F. C. Odds (ed.), Candida and candidosis. Baillière Tindall, London, United Kingdom.
35. **Pitcher, D. G., S. N. A., and R. J. Owen.** 1989. Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. Lett. Appl. Microbiol. **8:**151–156.
36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
37. **Savitzky, A., and M. J. E. Golay.** 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36:**1627–1633.
38. **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.
39. **Sullivan, D., D. Bennet, M. Henman, P. Harwood, S. Flint, F. Mulcahy, D. Shanley, and D. Coleman.** 1993. Oligonucleotide fingerprinting of isolates of *Candida* species other than *C. albicans* and of atypical *Candida* species from human immunodeficiency virus-positive and AIDS patients. J. Clin. Microbiol. **31:**2124–2133.
40. **Sullivan, D. J., T. J. Westerneng, K. A. Haynes, D. E. Bennett, and D. C. Coleman.** 1995. *Candida dubliniensis* sp. nov.: phenotypic and molecular characterization of a novel species associated with oral candidosis in HIV-infected individuals. Microbiology **141:**1507–1521.
41. **Timmins, É. M., and R. Goodacre.** 1997. Rapid quantitative analysis of binary mixtures of *Escherichia coli* strains using pyrolysis mass spectrometry with multivariate calibration and artificial neural networks. J. Appl. Microbiol. **83:**208–218.
42. **Vazquez, J. A., A. Beckley, J. D. Sobel, and M. 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.
43. **Windig, W., J. Haverkamp, and P. G. Kistemaker.** 1983. Interpretation of sets of pyrolysis mass spectra by discriminant analysis and graphical rotation. Anal. Chem. **55:**81–88.
44. **Wold, H.** 1966. Estimation of principal components and related models by iterative least squares, p. 391–420. *In* K. R. Krishnaiah (ed.), Multivariate analysis. Academic Press, Inc., New York, N.Y.