APPLIED AND ENVIRONMENTAL MICROBIOLOGY, June 1998, p. 2207–2214 Vol. 64, No. 6
0099-2240/98/$04.00+0

Rapid and Reliable Identification of Food-Borne Yeasts by Fourier-Transform Infrared Spectroscopy

MICHAEL KÜMMERLE, SIEGFRIED SCHERER, AND HERBERT SEILER*

Institut für Mikrobiologie, Forschungszentrum für Milch und Lebensmittel Weihenstephan, Technische Universität München, D-85354 Freising, Germany

Received 13 August 1997/Accepted 9 March 1998

Computer-based Fourier-transform infrared spectroscopy (FT-IR) was used to identify food-borne, predominantly fermentative yeasts. Dried yeast suspensions provided the films suitable for FT-IR measurement. Informative windows in the spectrum were selected and combined to achieve optimal results. A reference spectrum library was assembled, based on 332 defined yeast strains from international yeast collections and our own isolates. All strains were identified with conventional methods using physiological and morphological characteristics. In order to assess identification quality, another 722 unknown yeast isolates not included in the reference spectrum library were identified both by classical methods and by comparison of their FT-IR spectra with those of the reference spectrum library. Ninety-seven and one-half percent of these isolates were identified correctly by FT-IR. Easy handling, rapid identification within 24 h when starting from a single colony, and a high differentiation capacity thus render FT-IR technology clearly superior to other routine methods for the identification of yeasts.

Yeasts not only provided humans with the first biotechnologically produced food such as wine, bread, and fermented milk products but are also responsible for food spoilage (19), and some species are of medical importance. Therefore, a reliable method of yeast identification is economically significant (40). Furthermore, until now about 700 yeast species have been described. Since only a few habitats have been investigated in detail so far, a wide range of yeasts is likely to be discovered in the future (6). Exploration of new species includes the identification of a large number of isolates in order to eliminate duplicates and to discover unusual forms. For such tasks, a rapid, simple, low-cost identification method is needed. Conventional differentiation systems using morphological characters as well as patterns of the assimilation and fermentation of carbon sources (4, 22, 35) do not fulfil these requirements (9, 33, 38, 40). They are tedious and time-consuming, and, quite often, their capacity is limited since many species are distinguished from one another by a single physiological reaction which is often controlled by only one mutable marker (4, 20).

Alternative methods such as fatty acid analysis (1, 31), electrophoretic karyotyping (10), restriction fragment length polymorphism, and DNA fingerprinting (26, 37) have already been evaluated (8). Restriction enzyme analysis of PCR-amplified rDNA (2), randomly amplified polymorphic DNA (3, 27), and nucleic acid hybridization with oligonucleotide probes (21, 24) have also been used. While some of these techniques do provide satisfactory results, molecular methods in general are still difficult to perform on a routine basis in laboratories of the food industry.

Fourier-transform infrared (FT-IR) spectroscopy is used for the identification of substances in chemical analyses (14). The wavelength of infrared radiation ranges from 1 μm to 1 mm (32). In general, the wave number ν, the reciprocal of the wavelength, is used as a physical unit for FT-IR spectroscopy. Infrared radiation is divided into near (ν = 12,500 to 4,000 cm^{-1}), middle (ν = 4,000 to 200 cm^{-1}), and far (ν = 200 to 10 cm^{-1}) infrared. In this work, only the middle infrared section was used. FT-IR spectroscopy involves the observation of vibrations of molecules that are excited by an infrared beam. Molecules are able to absorb the energy of distinct light quanta and start a rocking or rotation movement. The FT-IR spectrum uses only vibrations that lead to a change in the dipole moment (14). An infrared spectrum represents a fingerprint which is characteristic for any chemical substance.

The composition of biological material and, thus, of its FT-IR spectrum, is exceedingly complex, representing a characteristic fingerprint. Some years ago, Naumann and coworkers suggested identifying microorganisms by FT-IR spectroscopy (28–30). In principle, a reference spectrum library is assembled based on well-characterized strains and species. The FT-IR spectrum of any unidentified isolate is then measured under the same conditions as those used for the reference spectra and is compared to spectra in the reference spectrum library. If the library contains an identical or a very similar spectrum, an identification is possible. The success of the method is, therefore, directly dependent on the complexity of the reference spectrum library. The application of FT-IR spectroscopy has been reported for some species of the genera *Lactobacillus* (7), *Actinomyces* (15), *Listeria* (18), *Streptococcus* (13), and *Clostridium* (11). There are two reports which present preliminary data indicating that eukaryotic microorganisms such as yeasts may also be identified by FT-IR (17, 36). However, all these studies are based on a very limited number of species and isolates. For verification of the method only a few strains, which often were part of the reference spectrum library as well, were used. It was, therefore, still unclear whether FT-IR spectroscopy indeed was a competitive identification method.

The aim of this study was to develop a standardized sample preparation procedure for yeasts (suitable for the normal laboratory), to select the most significant spectral windows for efficient identification, and to assemble a spectral reference library of sufficient complexity. Last, the identification of a great variety of unknown yeast isolates by FT-IR spectroscopy

* Corresponding author. Mailing address: Institut für Mikrobiologie, Forschungszentrum für Milch und Lebensmittel Weihenstephan, Technische Universität München, Vöttingerstrasse 45, D-85354 Freising, Germany. Phone: 08161/713519. Fax: 08161/714492. E-mail: seiler@lrz.tu-muenchen.de.

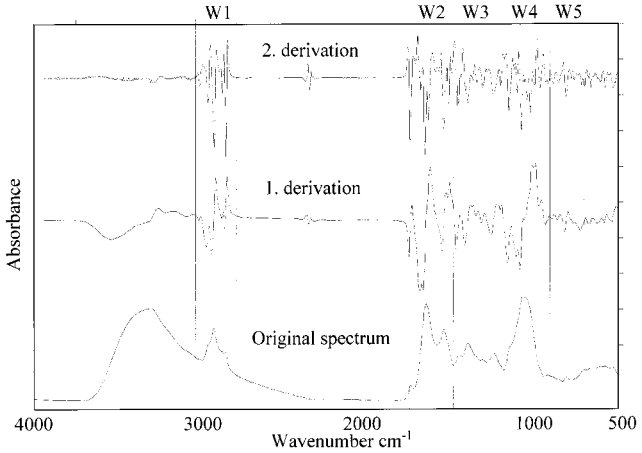

FIG. 1. Original spectrum as well as first and second derivations of an FT-IR measurement of an *S. cerevisiae* strain. Potentially informative spectral windows (W1 to W5) are indicated.

and conventional techniques had to be done in order to verify the method.

MATERIALS AND METHODS

Yeast strains. One hundred and seventy-four strains from international yeast culture collections provided the reference material. This collection was supplemented by 158 isolates from the Weihenstephan Yeast Collection (housed at our institute), representing a wide variety of habitats. All strains were identified by using miniature test samples in microtiter plates and conventional methods (35, 39), which test for about 100 physiological and morphological characters. The computer program of Barnett (5) was used to evaluate the results. The 332 reference yeasts of the resulting spectrum library represent 74 species of 18 genera.

Sample preparation. A single colony of yeast cells was transferred to an agar plate with a platinum loop, distributed with a Drigalski spatula, and incubated for 24 ± 1 h at 27 ± 2°C on YGCA (standard agar for yeasts in the food industry; Merck, Darmstadt, Germany) containing 5.0 g of yeast extract, 20.0 g of glucose, 0.1 g of chloramphenicol, and 14.9 g of agar per liter. For sample preparation, one loopful (1-mm-diameter platinum loop) of yeast cells scraped from this confluent lawn was suspended in 100 μl of distilled water. An aliquot of 35 μl was transferred to a ZnSe optical plate (sample holder) and dried at 42 ± 2°C for 1 h to yield transparent films, which were used directly for FT-IR spectroscopy. One sample holder accommodates 15 different samples. Measurement and comparison of the spectrum with the reference spectrum library containing spectra of defined strains take less than 2 min. In total, starting from a single yeast colony, identification is completed within 24 to 26 h.

FT-IR spectroscopy. All spectra between wave numbers 4,000 and 500 cm^{-1} were recorded with an IFS-28B FT-IR spectrometer (Bruker, Karlsruhe, Germany). For data processing, the software OPUS, version 2.2, for microbiological identification (Bruker) was used. The adjustment of instrument parameters was done according to the suggestions of the FT-IR workgroup of the Robert-Koch-Institut, Berlin, Germany (12). To diminish the difficulties arising from unavoidable baseline shifts and to improve the resolution of complex bands, the digitized original spectra were smoothed by the second derivation (16).

In principle, the five spectral windows W1 to W5 shown in Fig. 1 are potentially informative (16, 29, 30). Ranges of wave numbers can sometimes be associated with special chemical bonds. W1 is the so-called fatty acid region (3,050 to 2,800 cm^{-1}), where peaks mark the vibrations of the CH2 and CH3 groups of fatty acids. The W2 region is the amide section (1,750 to 1,500 cm^{-1}), where protein and peptide bands dominate. W3, which ranges from 1,500 to 1,200 cm^{-1}, is a mixed region containing vibrations of fatty acids, proteins, and polysaccharide. W4 (1,200 to 900 cm^{-1}) is dominated by polysaccharide peaks. Until now an exact correlation between peaks and molecules in this section was not possible. The so-called fingerprint region W5 ranges from 900 to 700 cm^{-1}. This window contains bands which are most characteristic at the species level. Again, just a few peaks can be assigned to the vibrations of special substances.

Selection of spectral windows. Cluster analysis was used to identify optimal windows (Fig. 1). For this purpose, windows were varied systematically until FT-IR identification was in accordance with the results of conventional microbiological identification. This variation included window size, the number of windows used, and the weighting factors imposed on each window. Best results were obtained with windows in the 3,030 to 2,830, 1,350 to 1,200, and 900 to 700 cm^{-1} ranges (all weighting factors were 1), and this configuration was therefore used as the standard. To cope with distances caused by unavoidable physical and biological variations such as slightly different growth in different batches due to medium preparation and variation in the dry microorganism film on the sample holder, each strain was measured at least three times in independent assays using different growth medium preparations of the standard agar. Then, an average spectrum was calculated and added to the reference spectrum library.

Cluster analysis and hit list identification. The spectral distance (also called the d value) is a measure of the similarity of the spectra of two isolates and reflects the size of nonoverlapping areas (29) of both spectra (for an example, see Fig. 2). d values between all spectra were calculated. The resulting distance matrix provided the basis for cluster analysis (construction of dendrograms by average linkage). By using identification analysis, the spectrum of an unknown isolate was compared to all spectra of the library. A hit list of 10 strains exhibiting the closest spectral distances was printed along with the d values. In Table 1, three identifications of different *Saccharomyces cerevisiae* strains are given.

RESULTS AND DISCUSSION

Reproducibility of measurements. FT-IR spectra are influenced by variation of plating methods, growth temperature, incubation time, and even the drying method of the microorganism suspension located on the sample holder. For a high level of reproducibility it was necessary to develop the standardized preparation procedure given in the sample preparation section above. The significance level of spectral distances is given by the d values observed for multiple, independent measurements of one strain. This level was always characterized by a d value less than 0.3 and turned out to be species dependent. In some cases, the d value was as low as 0.1.

Changes of the agar medium used had pronounced influences on the spectra. Any newly purchased charge of an identical medium produced slightly different spectra and had to be verified by recording the spectra of a standard set of seven yeast strains which had been shown to be especially sensitive to medium variations. A new charge of medium can only be considered suitable for FT-IR measurements if distances between the new spectra and the reference spectra, both taken from the standard species test set, are below 0.3. While this procedure can be performed easily in any research laboratory, "FT-IR grade" standard media have to be commercially available if the FT-IR technique is to be adopted for routine microbiological analysis.

FT-IR spectroscopy as a general taxonomic tool? A dendrogram of 332 well-characterized reference yeasts calculated from FT-IR spectra is shown in Fig. 3. It provides a graphic

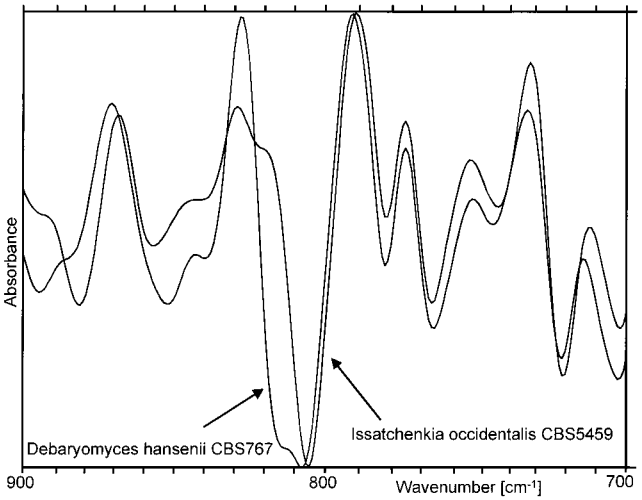

FIG. 2. Comparison of the fingerprint regions of two normalized FT-IR spectra. The d value is equivalent to the area which is covered by only one of the spectra.

TABLE 1. Examples of hit lists of different identification quality for FT-IR identification of three *Saccharomyces* isolates preidentified by conventional methods[a]

Sample hit no.	*S. cerevisiae* ze61204		*S. cerevisiae* mü70403		*S. cerevisiae* Kefir 438	
	d value	Reference strain	d value	Reference strain	d value	Reference strain
1	0.35	*S. cerevisiae* WSYC G1383	0.69	*S. cerevisiae* WSYC G1526	1.28	*S. cerevisiae* WSYC G1526
2	0.54	*S. cerevisiae* WSYC D19	1.75	*S. cerevisiae* WSYC G1383	1.28	*S. cerevisiae* DSM 70514
3	0.55	*S. cerevisiae* WSYC D9	2.01	*S. cerevisiae* WSYC D9	1.46	*S. cerevisiae* WSYC 117
4	0.67	*S. cerevisiae* WSYC G1526	2.22	*S. cerevisiae* WSYC D19	1.47	*S. cerevisiae* WSYC G1383
5	0.77	*S. cerevisiae* WSYC 117	2.34	*S. cerevisiae* WSYC 117	1.50	*S. cerevisiae* WSYC D9
6	1.03	*S. cerevisiae* WSYC G1305	2.47	*S. cerevisiae* DSM 70509	1.62	*S. cerevisiae* WSYC G1305
7	1.72	*S. cerevisiae* DSM 70509	2.53	*S. cerevisiae* DSM 70514	1.70	*S. cerevisiae* WSYC G1466
8	1.78	*S. pombe*[b] CBS 356	2.61	*S. cerevisiae* WSYC M240	1.71	*S. cerevisiae* DSM 70509
9	1.80	*K. marxianus* CBS 6432	2.64	*S. cerevisiae* WSYC G1305	1.91	*S. cerevisiae* WSYC G1461
10	1.82	*S. cerevisiae* WSYC M240	2.76	*S. unisporus* CBS 1575	1.99	*S. cerevisiae* DSM 1333

[a] CBS, Centraalbureau voor Schimmelcultures, Delft, The Netherlands; DSM, Deutsche Sammlung für Mikroorganismen, Braunschweig, Germany; WSYC, Weihenstephan Yeast Collection, Institut für Mikrobiologie, FML Freising-Weihenstephan, Germany.
[b] *S. pombe*, *Schizosaccharomyces pombe.*

impression of the distances dealt with in Table 2, where the isolates and cluster levels are listed in detail. Three arbitrary cluster levels were defined in order to divide the dendrogram into spectrally related groups. The 22 groups (A to V) of level 1 are separated by spectral distances of 2.0 to 2.5. The subclusters of those groups (level 2) have d values of 1.25 to 1.75, while level 3 is characterized by d values of 0.5 to 0.75. It is not possible to assign taxonomic interpretations to these levels.

The first clear result emerging from the data presented in Fig. 3 and Table 2 is that different species of the same genus generally did not cluster. This is most obvious for the genera *Pichia* and *Candida*, for which a variety of species were available. Other algorithms for cluster analysis also do not cluster all species of a single genus.

Since the taxonomy of yeasts is far from being finally settled and relies largely on phenotypic characters, one might suppose

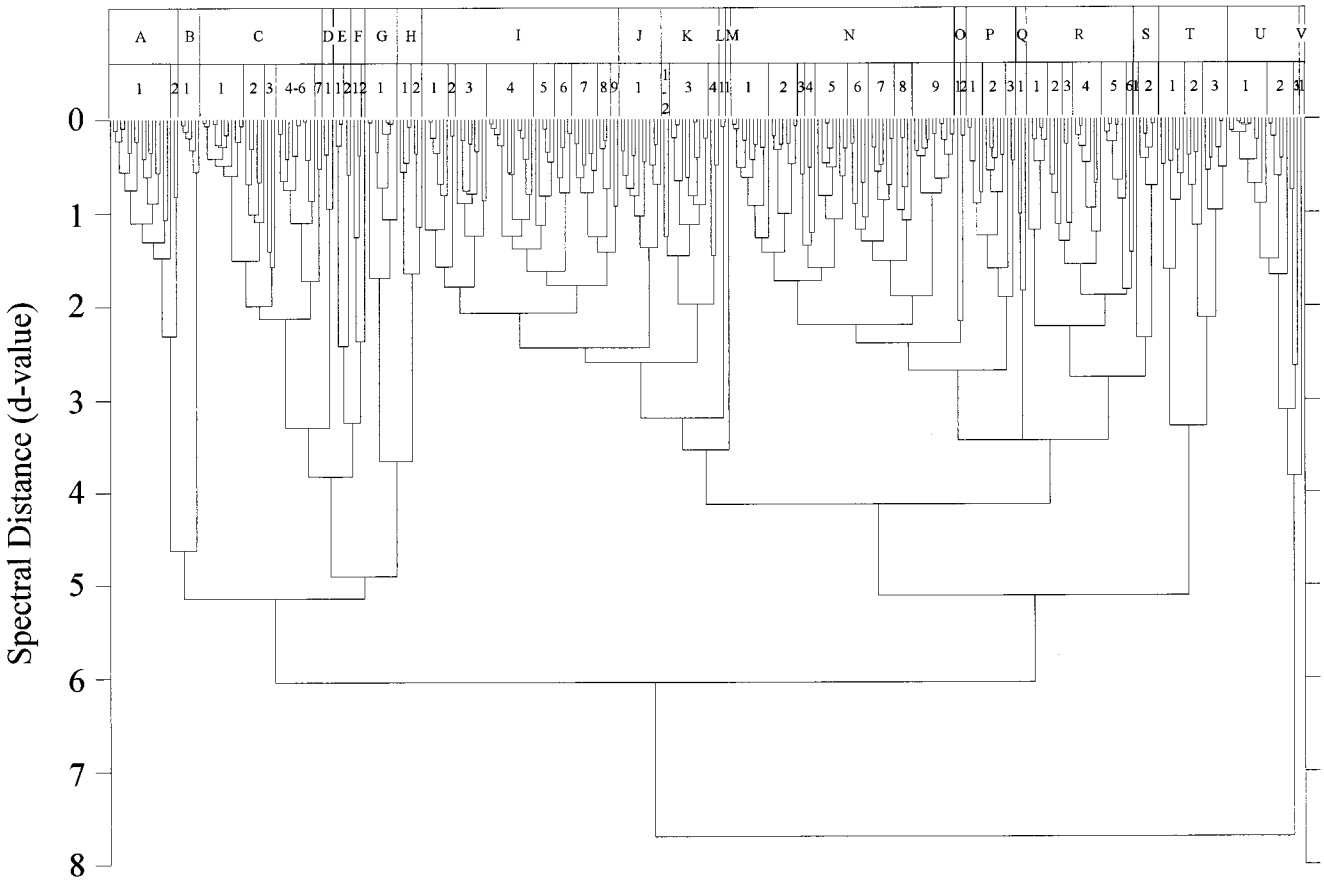

FIG. 3. Dendrogram of the mean spectra of the 332 yeast strains forming the reference spectrum library used for identification of unknown isolates. The dendrogram was calculated by an average-linkage algorithm and is divided into 22 major clusters (A to V). Each of those is further subdivided into second-order (1 to 9) and third-order (a to h; not listed) clusters. In Table 2, this nomenclature is also employed and can be used to identify individual strains.

TABLE 2. The 322 yeast strains used to create the FT-IR reference spectrum library[a]

Cluster	Species
A 1 a	*Issatchenkia occidentalis* WSYC D1
	Issatchenkia occidentalis CBS 5459T
	Issatchenkia occidentalis WSYC G348
	Issatchenkia occidentalis WSYC G323
b	*Issatchenkia orientalis* WSYC G1327
	Issatchenkia orientalis DSM 3433T
	Pichia norvegensis CBS 6564
c	*Pichia membranaefaciens* WSYC 247
	Pichia membranaefaciens CBS 1328
d	*Issatchenkia occidentalis* CBS 6887
	Pichia angusta CBS 4732
	Pichia angusta CBS 7031
	Pichia pini CBS 2613
e	*Pichia norvegensis* CBS 2126
	Pichia norvegensis CBS 6639
f	*Candida boidinii* DSM 70026T
g	*Candida ethanolica* CBS 8041T
h	*Pichia pini* DSM 70385T
2 a	*Issatchenkia occidentalis* WSYC M160
b	*Pichia cactophila* WSYC M111
B 1 a	*Candida sorbophila* WSYC M241
	Candida sorbophila WSYC M289
	Candida sorbophila WSYC G1332
	Candida sorbophila WSYC G1339
	Candida sorbophila WSYC M177
	Candida sorbophila WSYC G1324
C 1 a	*Clavispora lusitaniae* WSYC G1456
	Clavispora lusitaniae WSYC D6
	Clavispora lusitaniae WSYC D22
	Clavispora lusitaniae CBS 6936T
	Clavispora lusitaniae WSYC G1528
	Clavispora lusitaniae WSYC M267
	Clavispora lusitaniae WSYC M288
	Clavispora lusitaniae CBS 1944
	Clavispora lusitaniae CBS 4414
	Clavispora lusitaniae WSYC D14
	Clavispora lusitaniae WSYC G1568
	Clavispora lusitaniae WSYC D7
2 a	*Candida intermedia* CBS 572T
b	*Issatchenkia orientalis* DSM 70086
	Pichia norvegensis CBS 2125
c	*Pichia ohmeri* DSM 70813
	Pichia ohmeri DSM 70815T
d	*Candida kruisii* CBS 6451
3 a	*Candida rugosa* CBS 613
4 a	*Candida savonica* CBS 6563
5 a	*Candida sake* CBS 617
6 a	*Candida intermedia* WSYC M242
	Candida intermedia WSYC M273
b	*Candida tropicalis* CBS 643
	Candida tropicalis CBS 94T
c	*Candida tropicalis* WSYC M203
	Candida tropicalis WSYC M195
	Candida tropicalis DSM 1346
d	*Candida intermedia* WSYC M244
	Candida intermedia WSYC M252
	Candida intermedia WSYC M264
	Candida intermedia DSM 70753

Cluster	Species
C 7 a	*Candida santamariae* CBS 4515
	Candida santamariae WSYC G1532
D 1 a	*Hanseniaspora osmophila* CBS 313T
	Hanseniaspora vineae DSM 70283
	Hanseniaspora vineae CBS 2174
E 1 a	*Candida lactiscondensi* WSYC D10
	Candida lactiscondensi WSYC D11
	Candida lactiscondensi WSYC D21
2 a	*Hanseniaspora valbyensis* CBS 479
	Hanseniaspora valbyensis CBS 6618
F 1 a	*Dekkera anomala* DSM 70727
b	*Dekkera bruxellensis* DSM 3429
	Dekkera bruxellensis CBS 2797
2 a	*Dekkera custersiana* DSM 70736
G 1 a	*Candida etchelsii* WSYC M238
	Candida etchelsii WSYC M239
c	*Candida etchelsii* CBS 750
	Candida magnoliae DSM 70638T
	Candida etchelsii WSYC G1078
	Candida etchelsii WSYC G1081
	Candida etchelsii WSYC G1083
	Candida etchelsii WSYC G108
c	*Candida etchelsii* WSYC G1109
H 1 a	*Candida rugosa* WSYC G1349
	Candida rugosa DSM 70761
	Candida rugosa WSYC M153
	Candida rugosa WSYC M173
2 a	*Candida versatilis* CBS 1731
	Candida versatilis CBS 1752T
b	*Candida versatilis* CBS 1757
I 1 a	*Saccharomyces exiguus* WSYC D24
b	*Saccharomyces exiguus* WSYC G1510
	Saccharomyces exiguus WSYC G1502
	Saccharomyces exiguus CBS 6388
	Saccharomyces exiguus WSYC D26
c	*Saccharomyces dairensis* CBS 1579
	Saccharomyces dairensis CBS 421
2 a	*Candida sake* WSYC M249
	Candida sake WSYC M292
3 a	*Saccharomycodes ludwigii* DSM 3438
b	*Torulaspora delbrueckii* CBS 1146T
	Torulaspora delbrueckii DSM 70607
	Torulaspora delbrueckii DSM 70497
	Torulaspora delbrueckii CBS 818
	Torulaspora delbrueckii CBS 6104
	Torulaspora delbrueckii CBS 817
c	*Torulaspora pretoriensis* DSM 70525T
	Torulaspora pretoriensis CBS 5080
4 a	*Kluyveromyces marxianus* WSYC 239
	Kluyveromyces marxianus CBS 397
	Kluyveromyces marxianus CBS 1555
	Kluyveromyces marxianus CBS 712T
	Kluyveromyces marxianus DSM 70804
b	*Kluyveromyces marxianus* WSYC G1326
	Kluyveromyces marxianus DSM 70292
	Kluyveromyces marxianus CBS 5795

Cluster	Species
I 4 c	*Kluyveromyces marxianus* WSYC G1331
	Kluyveromyces marxianus WSYC G1336
	Kluyveromyces marxianus DSM 70343
	Kluyveromyces marxianus WSYC G1346
d	*Kluyveromyces polysporus* DSM 70294T
5 a	*Debaryomyces hansenii* WSYC M276
b	*Metschnikowia pulcherrima* CBS 2251
	Metschnikowia pulcherrima WSYC G1300
	Metschnikowia pulcherrima WSYC 238
c	*Pichia anomala* CBS 249
6 a	*Debaryomyces hansenii* WSYC M262
	Debaryomyces hansenii WSYC M218
	Debaryomyces hansenii WSYC M219
	Debaryomyces hansenii WSYC G1342
	Debaryomyces hansenii WSYC G1303
7 a	*Saccharomyces exiguus* NCYC 138
	Saccharomyces exiguus CBS 135
b	*Kluyveromyces lactis* CBS 141
	Kluyveromyces marxianus CBS 1558
c	*Kluyveromyces lactis* CBS 743
	Kluyveromyces marxianus DSM 70801
d	*Saccharomyces cerevisiae* DSM 70547
8 a	*Hanseniaspora guilliermondii* DSM 3432T
	Hanseniaspora uvarum CBS 312
	Hanseniaspora uvarum CBS 314T
	Hanseniaspora uvarum WSYC M148
9 a	*Candida sake* CBS 159T
b	*Hanseniaspora osmophila* DSM 70789
J 1 a	*Zygosaccharomyces bailii* DSM 70492
	Zygosaccharomyces bailii DSM 70834
	Zygosaccharomyces bisporus CBS 2449
b	*Zygosaccharomyces bisporus* DSM 70415T
	Zygosaccharomyces rouxii CBS 732T
c	*Zygosaccharomyces mellis* CBS 684
	Zygosaccharomyces mellis CBS 736
d	*Zygosaccharomyces mellis* WSYC G1449
	Zygosaccharomyces rouxii CBS 441
	Zygosaccharomyces rouxii CBS 731
	Zygosaccharomyces rouxii CBS 5717
K 1 a	*Saccharomyces exiguus* WSYC G1306
2 a	*Saccharomyces servazzii* CBS 4311T
3 a	*Saccharomyces cerevisiae* WSYC D8
	Saccharomyces cerevisiae WSYC G1465
	Saccharomyces cerevisiae WSYC G1467
	Saccharomyces cerevisiae CBS 2858
b	*Saccharomyces cerevisiae* CBS 6017
	Saccharomyces cerevisiae DSM 70449T
c	*Saccharomyces cerevisiae* DSM 70470
	Saccharomyces cerevisiae DSM 70511
	Saccharomyces cerevisiae DSM 70478
d	*Saccharomyces kluyveri* CBS 3082T
	Saccharomyces kluyveri DSM 70517
4 a	*Saccharomyces cerevisiae* WSYC 1305
b	*Saccharomyces cerevisiae* CBS 1489
	Saccharomyces cerevisiae DSM 70508

[a] The arrangement of strains in clusters is in accordance with Fig. 1. For abbreviations used in strain designations, see footnote *a* to Table 1. T, type strain. The cluster designations indicate spectral distances as follows: for A to V, d = 2.0 to 2.5; for 1 to 9, d = 1.25 to 1.75; for a to h, d = 0.5 to 0.75.

TABLE 2—*Continued*

Cluster	Species
L 1 a	*Saccharomyces cerevisiae* WSYC D17
	Saccharomyces cerevisiae WSYC G1531
M 1 a	*Citeromyces matritensis* DSM 70187T
N 1 a	*Candida parapsilosis* WSYC D22
	Candida parapsilosis DSM 5784T
	Candida parapsilosis CBS 1954
	Candida parapsilosis WSYC M214
	Candida parapsilosis WSYC G1332
	Candida parapsilosis WSYC M213
	Candida parapsilosis WSYC G1546
	Candida parapsilosis WSYC M187
b	*Candida parapsilosis* DSM 70125
	Candida parapsilosis WSYC M51
c	*Candida parapsilosis* WSYC G1523
2 a	*Candida maltosa* WSYC D2
	Candida maltosa WSYC M245
	Candida maltosa WSYC M256
	Candida maltosa WSYC G1457
b	*Candida intermedia* WSYC M175
	Candida intermedia WSYC M151
	Candida intermedia WSYC D15
	Candida intermedia WSYC G1529
3 a	*Candida catenulata* CBS 565
	Candida catenulata DSM 70040
4 a	*Debaryomyces vanrijiae* CBS 5458
	Debaryomyces vanrijiae WSYC M79
b	*Wickerhamia fluorescens* DSM 70715T
5 a	*Pichia kluyveri* WSYC G1512
	Pichia kluyveri WSYC G1513
	Pichia kluyveri WSYC G1505
	Pichia kluyveri WSYC G1506
	Pichia kluyveri WSYC D23
	Pichia kluyveri WSYC G1504
b	*Pichia jadinii* WSYC 1514
	Pichia jadinii CBS 1600
	Pichia jadinii DSM 2361
6 a	*Pichia farinosa* CBS 185T
	Pichia farinosa CBS 5031
	Pichia farinosa WSYC D27
b	*Debaryomyces hansenii* CBS 2334
	Debaryomyces hansenii DSM 70244
c	*Wingea robertsiae* DSM 70870T
7 a	*Pichia guilliermondii* WSYC G1304
	Pichia guilliermondii WSYC D4
	Pichia guilliermondii DSM 6381
	Pichia guilliermondii CBS 2030T
b	*Candida diddensiae* CBS 2219T
	Candida diddensiae CBS 4514
	Candida diddensiae WSYC M184
8 a	*Metschnikowia reukauffii ex* CBS 5833
b	*Metschnikowia reukauffii* CBS 5834
	Metschnikowia reukauffii DSM 70880
c	*Pichia guilliermondii* DSM 70052
d	*Pachysolen tannophilus* DSM 70352T

Cluster	Species
N 9 a	*Pichia anomala* CBS 257
	Pichia anomala CBS 5759T
	Pichia anomala DSM 70130
	Pichia anomala WSYC 282
	Pichia anomala DSM 70260
	Pichia anomala WSYC G1438
b	*Pichia anomala* WSYC D3
	Pichia anomala WSYC D5
	Pichia anomala WSYC D20
	Pichia anomala WSYC D13
	Pichia anomala WSYC D16
c	*Pichia fabianii* CBS 5482
	Pichia fabianii CBS 5640
O 1 a	*Candida catenulata* WSYC 251
2 a	*Candida zeylanoides* CBS 619
	Candida zeylanoides CBS 641
P 1 a	*Kluyveromyces lactis* CBS 2359
	Kluyveromyces lactis WSYC 54
	Kluyveromyces lactis CBS 683T
b	*Kluyveromyces lactis* CBS 2360
c	*Kluyveromyces marxianus* var. *vanudeni* DSM 70807
2 a	*Saccharomyces cerevisiae* CBS 1464
	Saccharomyces cerevisiae CBS 5926
	Saccharomyces cerevisiae DSM 70471
	Saccharomyces cerevisiae DSM 70514
	Saccharomyces cerevisiae WSYC G1443
	Saccharomyces cerevisiae WSYC G1445
3 a	*Kluyveromyces marxianus* WSYC G1503
	Kluyveromyces marxianus WSYC G1515
Q 1 a	*Schizosaccharomyces pombe* WSYC G1308
b	*Schizosaccharomyces pombe* CBS 356
c	*Schizosaccharomyces pombe* DSM 1063
R 1 a	*Saccharomyces cerevisiae* WSYC G1526
b	*Saccharomyces cerevisiae* WSYC G1307
	Saccharomyces cerevisiae WSYC 117
	Saccharomyces cerevisiae WSYC D9
	Saccharomyces cerevisiae WSYC D18
	Saccharomyces cerevisiae WSYC D19
2 a	*Candida glabrata* DSM 6425
b	*Kluyveromyces marxianus* CBS 6432
c	*Endomyces fibuliger* WSYC M166
d	*Zygosaccharomyces bailii* WSYC G1530
3 a	*Candida zeylanoides* CBS 947
	Candida zeylanoides WSYC 27
	Candida zeylanoides WSYC M69
b	*Saccharomyces cerevisiae* DSM 70509
4 a	*Hanseniaspora guilliermondii* DSM 70285
	Hanseniaspora uvarum CBS 5074
	Hanseniaspora uvarum WSYC 221
	Hanseniaspora uvarum CBS 2585
b	*Kluyveromyces africanus* CBS 2665
c	*Kluyveromyces africanus* DSM 70290T

Cluster	Species
R 4 d	*Saccharomyces servazzii* CBS 6865
e	*Hanseniaspora uvarum* WSYC M143
5 a	*Saccharomyces cerevisiae* WSYC G1469
	Saccharomyces cerevisiae WSYC G1461
	Saccharomyces cerevisiae WSYC G1466
	Saccharomyces cerevisiae ATCC 9763
	Saccharomyces cerevisiae DSM 1332
	Saccharomyces cerevisiae WSYC 118
	Saccharomyces cerevisiae WSYC G1468
6 a	*Saccharomyces cerevisiae* WSYC G1006
7 a	*Saccharomyces unisporus* CBS 1575
S 1 a	*Saccharomyces cerevisiae* WSYC M240
2 a	*Torulaspora delbrueckii* WSYC M150
	Torulaspora delbrueckii CBS 451
	Torulaspora delbrueckii CBS 6795
	Torulaspora delbrueckii WSYC G1540
	Torulaspora delbrueckii WSYC D12
	Torulaspora delbrueckii DSM 70526
T 1 a	*Candida famata* CBS 941
	Debaryomyces hansenii var. *fabyanii* CBS 789T
b	*Debaryomyces hansenii* CBS 6960
	Debaryomyces hansenii WSYC G1090
	Debaryomyces hansenii DSM 70238
	Debaryomyces hansenii CBS 767T
	Debaryomyces hansenii WSYC G1094
2 a	*Debaryomyces hansenii* CBS 117
	Debaryomyces hansenii CBS 772
	Debaryomyces hansenii WSYC 1100
	Debaryomyces hansenii WSYC G1096
b	*Pichia haplophila* DSM 70365
3 a	*Pichia triangularis* WSYC G1102
	Pichia triangularis WSYC G1097
	Pichia triangularis CBS 4094T
b	*Pichia triangularis* WSYC G1085
	Pichia triangularis WSYC G1087
	Pichia triangularis WSYC G1088
	Pichia triangularis WSYC G1092
U 1 a	*Debaryomyces hansenii* WSYC M269
	Debaryomyces hansenii WSYC G1334
	Debaryomyces hansenii WSYC G1319
	Debaryomyces hansenii WSYC G1321
	Debaryomyces hansenii WSYC G1345
	Debaryomyces hansenii WSYC G1340
	Debaryomyces hansenii WSYC G1343
	Debaryomyces hansenii WSYC G1325
	Debaryomyces hansenii WSYC G1341
	Debaryomyces hansenii CBS 1099
	Debaryomyces hansenii CBS 773
b	*Debaryomyces hansenii* WSYC G1105
	Debaryomyces hansenii WSYC G1080
	Debaryomyces hansenii WSYC G1084
	Debaryomyces hansenii WSYC G1134
	Debaryomyces hansenii WSYC G1131
c	*Debaryomyces hansenii* WSYC G1130
2 a	*Debaryomyces hansenii* WSYC G1347
	Debaryomyces hansenii WSYC G1079
3 a	*Debaryomyces vanrijiae* DSM 70252T
V 1 a	*Saccharomycopsis capsularis* DSM 70560T

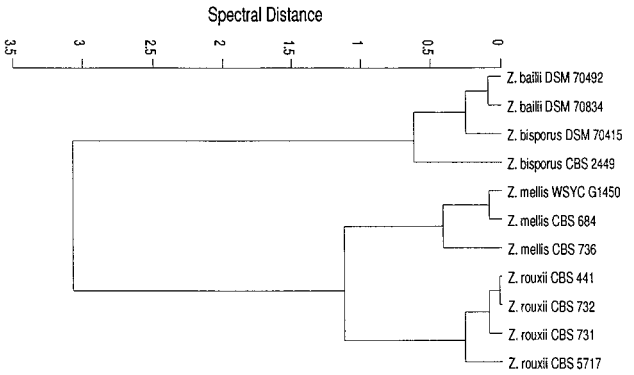

FIG. 4. Cluster analysis (average linkage) of four *Zygosaccharomyces* species. Spectral windows have been set at the following wave number ranges: 1,710 to 1,690, 1,213 to 1,202, and 777 to 767 cm^{-1}. The pattern shown in this figure corresponds closely to that derived from 18S rRNA sequences.

that molecular data may turn out to be in agreement with FT-IR spectroscopy. Some sequences do not confirm this idea. For instance, according to 18S rRNA, *Candida tropicalis*, *Candida parapsilosis*, and *Candida maltosa* are very closely related (see the phylogenetic tree in reference 21). However, *C. tropicalis* clusters far away from the other two species (cluster C6b versus clusters N1a and N2a in Fig. 3). At this stage, however, the relation between FT-IR and molecular taxonomy cannot be assessed conclusively, but we doubt that FT-IR can be used as a general taxonomic tool above the species level.

Strains of the same species may appear in different clusters. As is shown in Table 2, strains of many species cluster at level 3. However, there are a number of exceptions to this rule. For example, strains of *Issatchenkia orientalis* and *Issatchenkia occidentalis* fell into different clusters (clusters A1, A2, and C2b). They appear together with, e.g., *Pichia membranaefaciens* and *Pichia norvegensis*. It is interesting that these species are also difficult to separate with physiological markers (31). The same is true for *Kluyveromyces marxianus* and *Kluyveromyces lactis* (clusters I7b and -c), *Hanseniaspora uvarum* and *Hanseniaspora guilliermondii* (clusters I8a and R4a and -e) and *Hanseniaspora vineae* and *Hanseniaspora osmophila* (clusters D1a and I9b). A clear identification of these species was not always possible by physiological and morphological characteristics (unpublished data; see also reference 40).

During the creation of the reference spectrum library, it often happened that a new strain of a species clustered far away from the other strains already investigated. However, when more strains were included, it became clear that such an "aberrant" strain just represented the first example of a new cluster including several representatives. Therefore, species with many strains often formed more than one independent cluster. Typical examples are *S. cerevisiae*, *K. marxianus*, and *Debaryomyces hansenii*. The taxonomic significance of this finding has not been studied in detail so far and must be assessed in the future by using molecular taxonomic markers (compare references 3 and 20).

Optimization of spectral windows used for closely related yeast groups. There are several possible causes which may account for the formation of different clusters by strains of the same species. First, the window combination used for FT-IR spectroscopy may have been suboptimal in some cases. For instance, strains of *Zygosaccharomyces bisporus*, *Zygosaccharomyces bailii*, *Zygosaccharomyces rouxii*, and *Zygosaccharomyces mellis* (cluster J1) do not form species clusters. By conventional methods it is often difficult to identify these species (33, 40). Growth rates in the presence of 50 or 60% glucose and 1% acetic acid are the only criteria used to distinguish between these *Zygosaccharomyces* species (4). Analysis of 18S rRNA reflected a very close relationship between *Z. bailii* and *Z. bisporus* and a slightly greater distance between *Z. rouxii* and *Z. mellis* (20). Specific problems such as the identification of *Zygosaccharomyces* strains by FT-IR may be solved by optimizing the spectral window combination. For example, the windows characterized by wave numbers of 1,710 to 1,690, 1,213 to 1,202, and 777 to 767 cm^{-1} yielded the dendrogram shown in Fig. 4, which corresponds exactly to the results of the 18S rRNA analysis. This example demonstrates that a separation of yeasts which are difficult to identify by the general yeast identification window setting is possible by a stepwise optimization of spectral window selection. However, to do so, the "real" relationship of the isolates according to genomic DNA sequences, in this case 18S rRNA, must be known beforehand.

Detection of novel species, subspecies, or mutants by FT-IR? Another reason why strains of the same species do not cluster may be the existence of new subspecies or species. For instance, separate clusters for one species may be due to "hab-

TABLE 3. Field isolates of yeast strains which are not included in the FT-IR reference spectrum library[a]

Species[b]	No. of isolates
C. boidinii	1
C. etchellsii	6
C. glabrata	6
C. intermedia	28
C. oleophila	7
C. parapsilosis	16
C. rugosa	2
C. sorbophila	6
C. tropicalis	17
C. versatilis	2
C. zeylanoides	2
Clavispora lusitaniae	49
Debaryomyces hansenii	47
H. uvarum	15
H. vineae	4
H. valbyensis	1
Issatchenkia occidentalis	39
K. lactis	14
K. marxianus	40
Metschnikowia pulcherrima	5
P. anomala	100
P. cactophila	3
P. guilliermondii	10
P. jadinii	1
P. membranaefaciens	2
P. kluyveri	9
P. pseudocactophila	2
P. triangularis	12
S. cerevisiae	185
S. servazzii	3
S. unisporus	14
Sterigmatomyces haplophilus	2
Torulaspora delbrueckii	56
Yarrowia lipolytica	3
Zygosaccharomyces sp.	4
Z. rouxii	2
No identification possible	5

[a] These 722 isolates had been identified by conventional methods and were used subsequently for validation of the FT-IR identification. Five of 722 isolates could not be identified by conventional methods (last line of the table).

[b] *C.*, *Candida*; *H.*, *Hanseniaspora*; *K.*, *Kluyveromyces*; *P.*, *Pichia*; *S.*, *Saccharomyces.*

itat variants" (34). Successful adaptation to completely different habitats may result in the evolution of subspecies which may have different FT-IR spectra. One example is *D. hansenii*: strains from clusters U1a to -c, U2a, and T2a and -b were isolated from cheese and brine, while strains within cluster T1b come from beer, cattle, and yogurt. A second example is provided by strains of *S. cerevisiae*: those contained in clusters K3a to -c and L1a (Table 2) are from beer, wine, and must, while the strains combined in clusters P2a and R1a and -b were isolated from yogurt, diseased nails, and other sources. It is well known that strains of *S. cerevisiae* fall into different groups (3, 20, 25), but it is not clear whether these groups represent isolates from certain habitats.

Another indication for the existence of different subspecies or even species is extremely different degrees of homogeneity in the spectral distances between strains. This is, for instance, the case when clusters of *D. hansenii*, *Pichia anomala*, and *Torulaspora delbrueckii* are compared. While the *P. anomala* cluster and the two *T. delbrueckii* clusters exhibit an internal spectral distance of approximately 0.7, the distances between *D. hansenii* strains within the three species clusters were more than 1.7 (clusters T1 and T2 and U1 and U2; Table 2). Again, Kurtzman (23) noted that *D. hansenii* is a heterogeneous species according to a taxonomy based on physiological markers.

While such data are in accordance with the hypothesis of different taxonomic forms, molecular data are clearly needed to clarify the situation. In particular, there might be other reasons for single strains clustering independently, e.g., slow-growing mutants, strains with mutations of a biochemical character such as slime production, and strains with fast ascospore formation. Such mutants most probably would have significantly different FT-IR spectra due to major differences in cellular composition (11).

Validation of FT-IR identification. It appears that the use of FT-IR spectroscopy for taxonomic purposes is limited. This fact, however, does not prevent it from being a powerful identification system. To evaluate this potential, the method has been tested with 722 independent yeast isolates, which were obtained from different habitats, mostly from the food industry. They were not included in the reference spectrum library and constituted 36 yeast species belonging to 11 genera (Table 3). All isolates were identified in parallel by physiological and morphological characters. An identification by FT-IR spectroscopy was considered to be successful if the d value of the first recommended reference strain in the hit list (Table 1) was below 1.5 and, in addition, either the next similar hits were three strains of the same species or the distance between the first hit and the next species was larger than 0.25. Table 1 shows example hit lists as a result of three identity tests with three unknown isolates.

Five of 722 isolates were not identifiable by conventional methods. These strains may be mixed cultures which are difficult to purify, defective mutants, or novel species. They have not been investigated in further detail. Twelve isolates (1.7% of 717 strains) could be identified only when a subjective decision based on personal experience with yeast taxonomy (habitats and morphology, etc.) was used to evaluate the unclear FT-IR hit list. Another 6 of 717 strains could not be identified by FT-IR at all. In summary, 699 of 717 strains were identified correctly by FT-IR spectroscopy, which corresponds to an identification rate of 97.5%.

Conclusion. With an identification time of 24 to 26 h starting from a single colony and an identification rate of about 97%, FT-IR spectroscopy provides a superior, rapid alternative to conventional identification systems for food-borne yeasts, which take several days. Identification is limited only by the quality of the reference spectrum library, which can be improved steadily by adding further yeast isolates to the database. The method is easy to use, and we now routinely identify yeasts by FT-IR.

ACKNOWLEDGMENTS

This work was supported by the Bundesministerium für Wirtschaft through the Arbeitsgemeinschaft industrieller Forschungsvereinigungen "Otto von Guericke" e.V. (AiF), grant no. AiF-FV-10768N.

The comments of three reviewers led to a significantly improved presentation of the data.

REFERENCES

1. **Augustyn, O. P. H., J. L. F. Kock, and D. Ferreira.** 1992. Differentiation between yeast species, and strains within a species, by cellular fatty acid analysis. Syst. Appl. Microbiol. **15:**105–115.
2. **Baleiras Couto, M. M., J. T. W. E. Vogels, H. Hofstra, J. H. J. Huis in't Veld, and J. M. B. M. van der Vossen.** 1995. Random amplified polymorphic DNA and restriction enzyme analysis of PCR amplified rDNA in taxonomy: two identification techniques for food-borne yeasts. J. Appl. Bacteriol. **79:**525–535.
3. **Baleiras Couto, M. M., B. Eijsma, H. Hofstra, J. H. J. Huis in't Veld, and J. M. B. M. van der Vossen.** 1996. Evaluation of molecular typing techniques to assign genetic diversity among *Saccharomyces cerevisiae* strains. Appl. Environ. Microbiol. **62:**41–46.
4. **Barnett, J. A., R. W. Payne, and D. Yarrow.** 1990. Yeasts: characteristics and identification, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.
5. **Barnett, J. A.** 1996. Yeast identification program. J. A. Barnett, Norwich, England.
6. **Boekhout, T., and C. P. Kurtzman.** 1996. Principles and methods in yeast classification, and an overview of currently accepted yeast genera, p. 1–81. *In* K. Wolf (ed.), Nonconventional yeasts in biotechnology. Springer-Verlag, Berlin, Germany.
7. **Curk, M. C., F. Peladan, and J. C. Hubert.** 1994. Fourier-transform infrared (FTIR) spectroscopy for identifying *Lactobacillus* species. FEMS Microbiol. Lett. **123:**241–248.
8. **Deak, T., and L. R. Beuchart.** 1987. Identification of foodborne yeasts. J. Food Prot. **50:**243–264.
9. **Deak, T., and L. R. Beuchart.** 1988. Evaluation of simplified and commercial systems for identification of foodborne yeasts. Int. J. Food Microbiol. **7:** 135–145.
10. **Donhauser, S., R. Springer, and G. Vogeser.** 1990. Identifizierung und Klassifizierung von Brauereihefen durch Chromosomenanalyse mit der Pulsfeldgelelektrophorese. Monatsschr. Brauwissensch. **12:**392–400.
11. **Franz, M.** 1994. Identifizierung von Clostridien mittels FT-IR-Spektroskopie. Dtsch. Milchwirtsch. **3:**130–132.
12. **FT-IR workgroup RKI/BIAM.** 1992. Anleitung zur Charakterisierung von Mikroorganismen mit IFS 25/B und Opus. Robert-Koch-Institut, Berlin, Germany.
13. **Goodacre, R., É. M. Timmins, P. J. Rooney, J. J. Rowland, and D. Kell.** 1996. Rapid identification of *Streptomyces* species using diffuse reflectance-absorbance Fourier-transform infrared spectroscopy and artificial neural networks. FEMS Microbiol. Lett. **140:**233–239.
14. **Günzer, H., and H. M. Heisse.** 1996. IR-Spektroskopie. 3. Auflage. VCH-Verlag, Weinheim, Germany.
15. **Haag, H., H.-G. Gremlich, R. Bergmann, and J.-J. Sanglier.** 1996. Characterization and identification of actinomycetes by FT-IR spectroscopy. J. Microbiol. Methods **27:**157–163.
16. **Helm, D., H. Labischinski, G. Schallehn, and D. Naumann.** 1991. Classification and identification of bacteria by Fourier-transform infrared spectroscopy. J. Gen. Microbiol. **13:**69–79.
17. **Henderson, D. O., R. Mu, and M. Gunasekaran.** 1996. A rapid method for the identification of *Candida* at the species level by Fourier-transform infrared spectroscopy. Biochem. Lett. **51:**223–228.
18. **Holt, C., D. Hirst, A. Sutherland, and F. MacDonald.** 1995. Discrimination of species in the genus *Listeria* by Fourier transform infrared spectroscopy and canonical variate analysis. Appl. Environ. Microbiol. **61:**377–378.
19. **Jakobsen, M., and J. Narvhus.** 1996. Yeasts and their possible beneficial and negative effects on the quality of dairy products. Int. Dairy J. **6:**755–768.
20. **James, S. A., J. Cai, I. N. Roberts, and M. D. Collins.** 1997. A phylogenetic analysis of the genus *Saccharomyces* based on 18S rRNA gene sequences: description of *Saccharomyces kunashirensis* sp. nov. and *Saccharomyces martiniae* sp. nov. Int. J. Syst. Bacteriol. **47:**453–460.
21. **Kosse, D., H. Seiler, R. I. Amann, W. Ludwig, and S. Scherer.** 1997. Identification of yoghurt-spoiling yeasts with 18S-rRNA-targeted oligonucleotide probes. Syst. Appl. Microbiol. **20:**468–480.
22. **Kreger-van Rij, N. J. W.** 1984. The yeasts, a taxonomic study. Elsevier Science Publishers, Amsterdam, The Netherlands.

23. **Kurtzman, C. P.** 1994. Molecular taxonomy of yeasts. Yeasts **10:**1727–1740.
24. **Lischewski, A., R. I. Amann, D. Harmsen, H. Merkert, J. Hacker, and J. Morschhäuser.** 1996. Specific detection of *Candida albicans* and *Candida tropicalis* by fluorescent in situ hybridization with an 18S rRNA-targeted oligonucleotide probe. Microbiology **142:**2731–2740.
25. **Martini, A. V., and C. P. Kurtzman.** 1985. Deoxyribonucleic acid relatedness among species of the genus *Saccharomyces* sensu stricto. Int. J. Syst. Bacteriol. **35:**508–511.
26. **Meaden, P.** 1990. DNA fingerprinting of brewer's yeast: current perspectives. J. Inst. Brew. **96:**195–200.
27. **Messner, R., H. Prillinger, F. Altmann, K. Lopandic, K. Wimmer, O. Molnar, and F. Weigang.** 1994. Molecular characterization and application of random amplified polymorphic DNA analysis of *Mrakia* and *Sterigmatomyces* species. Int. J. Syst. Bacteriol. **44:**694–703.
28. **Naumann, D.** 1985. The ultra rapid differentiation and identification of pathogenic bacteria using FT-IR techniques. SPIE Fourier Comput. Infrared Spectrosc. **553:**268–269.
29. **Naumann, D., V. Fijala, H. Labischinski, and P. Giesbrecht.** 1988. The rapid differentiation and identification of pathogenic bacteria using Fourier-transform infrared spectroscopic and multivariate statistical analysis. J. Mol. Struct. **174:**165–170.
30. **Naumann, D., H. Labischinski, D. Helm, and P. Giesbrecht.** 1990. 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. **Noronha-da-Costa, P., C. Rodrigues, I. Spencer-Martins, and V. Loureiro.** 1996. Fatty acid patterns of film-forming yeasts and new evidence for the heterogeneity of *Pichia membranaefaciens*. Lett. Appl. Microbiol. **23:**79–84.
32. **Perkamps, H.-H.** 1993. Parat-Lexikon der Spektroskopie, 1st ed. VCH-Verlag, Weinheim, Germany.
33. **Praphailong, W., M. Van Gestel, G. H. Fleet, and G. M. Heard.** 1997. Evaluation of the Biolog system for identification of food and beverage yeasts. Lett. Appl. Microbiol. **24:**455–459.
34. **Roostita, R., and G. H. Fleet.** 1996. Growth of yeasts in milk and associated changes to milk composition. Int. J. Food Microbiol. **31:**205–219.
35. **Seiler, H., and M. Busse.** 1988. Identifizierung von Hefen mit Mikrotiterplatten. Forum Mikrobiol. **11:**505–509.
36. **Serfas, O., G. Standfuss, I. Flemming, and D. Naumann.** 1991. FT-IR-Spektroskopie in der Bioanalytik. BioTec Analytik **3:**42–47.
37. **Tichy, H.-V., and R. Simon.** 1994. Effiziente Analyse von Mikroorganismen mit PCR-Fingerprint-Verfahren. Bioforum **17:**499–505.
38. **Tötök, T., and A. D. King, Jr.** 1991. Comparative study on the identification of food-borne yeasts. Appl. Environ. Microbiol. **57:**1207–1212.
39. **Valdés-Stauber, N., S. Scherer, and H. Seiler.** 1997. Identification of yeasts and coryneform bacteria from the surface microflora of brick cheeses. Int. J. Food Microbiol. **34:**115–129.
40. **Welthagen, J. J., and B. C. Viljoen.** 1997. The value of certain chemotaxonomic methods in the identification of food related yeasts. Food Microbiol. **14:**231–245.