

## Rapid Identification of Mycolic Acid Patterns of Mycobacteria by High-Performance Liquid Chromatography Using Pattern Recognition Software and a *Mycobacterium* Library

SUZANNE E. GLICKMAN,<sup>1\*</sup> JAMES O. KILBURN,<sup>1</sup> W. RAY BUTLER,<sup>1</sup> AND L. SCOTT RAMOS<sup>2</sup>

*Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333,<sup>1</sup> and Infometrix, Inc., Seattle, Washington 98121<sup>2</sup>*

Received 7 September 1993/Returned for modification 6 November 1993/Accepted 13 December 1993

**Current methods for identifying mycobacteria by high-performance liquid chromatography (HPLC) require a visual assessment of the generated chromatographic data, which often involves time-consuming hand calculations and the use of flow charts. Our laboratory has developed a personal computer-based file containing patterns of mycolic acids detected in 45 species of *Mycobacterium*, including both slowly and rapidly growing species, as well as *Tsukamurella paurometabolum* and members of the genera *Corynebacterium*, *Nocardia*, *Rhodococcus*, and *Gordona*. The library was designed to be used in conjunction with a commercially available pattern recognition software package, Pirouette (Infometrix, Seattle, Wash.). Pirouette uses the *K*-nearest neighbor algorithm, a similarity-based classification method, to categorize unknown samples on the basis of their multivariate proximities to samples of a preassigned category. Multivariate proximity is calculated from peak height data, while peak heights are named by retention time matching. The system was tested for accuracy by using 24 species of *Mycobacterium*. Of the 1,333 strains evaluated,  $\geq 97\%$  were correctly identified. Identification of *M. tuberculosis* ( $n = 649$ ) was 99.85% accurate, and identification of the *M. avium* complex ( $n = 211$ ) was  $\geq 98\%$  accurate;  $\geq 95\%$  of strains of both double-cluster and single-cluster *M. gordonae* ( $n = 47$ ) were correctly identified. This system provides a rapid, highly reliable assessment of HPLC-generated chromatographic data for the identification of mycobacteria.**

Reverse-phase high-performance liquid chromatography (HPLC) of mycolic acid esters has been demonstrated to be a rapid, reproducible, species-specific method for the identification of *Mycobacterium* species (2-5, 8-10, 14). HPLC analysis can also be used to differentiate other mycolic acid-containing bacteria (1, 6), specifically, members of the mycolic acid-containing genera *Nocardia*, *Rhodococcus*, *Gordona*, and *Corynebacterium*, from *Mycobacterium* species. Mycobacterial identification by HPLC is relatively inexpensive and has been found to be more accurate than the use of commercial nucleic acid probes (10, 14). Also, commercial probes are not available for many *Mycobacterium* species and do not recognize some *M. avium* complex strains (13). As a result, HPLC is quickly becoming a procedure of choice for the routine identification of mycobacterial species in many laboratories.

Current methods for the interpretation of HPLC-generated chromatographic data do require some expertise. Minimally, each chromatogram must be visually assessed, and often, hand calculations and flow charts are necessary to validate the identifications for many mycobacterial species (2, 4, 8). These methods for evaluating chromatographic data can become tedious and time-consuming for laboratories that process large numbers of samples.

Our laboratory has developed a personal computer-based file (library) consisting of 45 species of *Mycobacterium*, including both slowly and rapidly growing organisms. This library is to be used in conjunction with a commercially available pattern recognition software package. The pattern recognition software, Pirouette (Infometrix, Seattle, Wash.), is able to process

chromatographic data by using the *K*-nearest neighbor (KNN) algorithm. A similarity-based classification method, the KNN algorithm calculates an *N*-dimensional matrix where *N* is the number of variables in the measured data, i.e., the number of peaks found. Each sample is represented as a point in the multidimensional matrix. Once a matrix is formed, an unknown sample can be plotted among the known points and its distance to the known points in the matrix, i.e., the *Mycobacterium* library, can be measured. A classification of the unknown is then made on the basis of the identities of its nearest neighbors in the matrix. The number of neighbors (*K*) assessed is predetermined by the user in the design of the method so as to achieve optimal segregation of like samples (12). This system provides a means for a very rapid assessment of chromatographic data while maintaining a high degree of accuracy ( $\geq 97\%$ ) for the identification of a broad range of mycobacterial species.

### MATERIALS AND METHODS

**Bacterial strains.** Selected strains from the American Type Culture Collection (ATCC) and the Trudeau Mycobacterial Culture Collection (TMC) as well as clinical and laboratory isolates submitted to the Mycobacteriology Laboratory, Centers for Disease Control and Prevention, Atlanta, Ga., were examined by HPLC. Those isolates used for the construction of the *Mycobacterium* library are described in detail in Table 1. An evaluation set, composed of different clinical and laboratory isolates, was used to test the *Mycobacterium* library. Isolates included in the evaluation set are listed in Table 2. All strains had been previously identified at the Centers for Disease Control and Prevention by conventional biochemical methods (11), by the use of commercially available DNA

\* Corresponding author. Mailing address: Mycobacteriology Laboratory, Mailstop F08, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, Chamblee, GA 30341. Phone: (404) 488-4181. Fax: (404) 488-4670.

TABLE 1. Strains used to construct the *Mycobacterium* library

Species	Strains
<i>M. tuberculosis</i>	TMC107, TMC109, TMC110, TMC111, TMC112, TMC116, TMC120, TMC124, TMC125, TMC302, TMC305, TMC307, TMC309, TMC314, TMC320, TMC321, TMC323, TMC326, TMC331, ATCC 27244, and 10 clinical isolates
<i>M. avium</i>	TMC701, TMC706, TMC715, TMC716, TMC1464, ATCC 25291, and 55 clinical isolates
<i>M. intracellulare</i>	TMC1411, TMC1469, TMC1473, TMC1476, ATCC 13950, and 22 clinical isolates
<i>M. scrofulaceum</i>	TMC1302, TMC1303, TMC1305, TMC1306, TMC1307, TMC1309, TMC1312, TMC1314, TMC1316, TMC1320, TMC1321, TMC1323, and 18 clinical isolates
<i>M. goodii</i>	TMC1319, TMC1324, TMC1325, TMC1327, ATCC 14490, and 20 laboratory isolates
<i>M. kansasii</i>	TMC1203, TMC1204, TMC1214, TMC1217, ATCC 12478, and 10 clinical isolates
<i>M. terrae</i>	TMC1450, ATCC 15755, ATCC 25147, ATCC 25149, ATCC 25217, ATCC 25267, ATCC 25268, ATCC 25269, and 8 laboratory isolates
<i>M. nonchromogenicum</i>	TMC1481, ATCC 3578, ATCC 19531, ATCC 19532, ATCC 19533, ATCC 25142, ATCC 25143, ATCC 25144, ATCC 25145, ATCC 25216, ATCC 25218, and 6 clinical isolates
<i>M. xenopi</i>	TMC1470, TMC1482, ATCC 19970, and 18 clinical isolates
<i>M. fortuitum</i>	TMC1529, TMC1530, and 25 clinical isolates
<i>M. peregrinum</i>	TMC1545, TMC1547, and 9 clinical isolates
<i>M. chelonae</i>	TMC1524 and 16 clinical isolates
<i>M. abscessus</i>	TMC1542, TMC1543, and 29 clinical isolates
<i>M. marinum</i>	TMC1218 and 17 clinical isolates
<i>M. chelonae-like</i>	18 clinical isolates
<i>M. flavescens</i>	TMC1541 and 5 laboratory isolates
<i>M. malmoense</i>	TMC802 and 14 clinical isolates
<i>M. simiae</i>	TMC1226 and 14 clinical isolates
<i>M. szulgai</i>	TMC1328 and 14 clinical isolates
<i>M. bovis</i> BCG	TMC1002, TMC1010, TMC1011, TMC1020, TMC1021, TMC1024, TMC1025, TMC1030, TMC1103, TMC1108, and 2 clinical isolates
<i>M. haemophilum</i>	ATCC 29548 and 10 clinical isolates
<i>M. asiaticum</i>	TMC803, ATCC 25274, and 10 clinical isolates
<i>M. shimoidei</i>	8 laboratory isolates
<i>M. triviale</i>	TMC1453, ATCC 23292, and 6 laboratory isolates
<i>M. gastri</i>	TMC1456, ATCC 25157, ATCC 25159, ATCC 25162, and 6 clinical isolates
<i>M. celatum</i>	10 clinical isolates
<i>M. aichiense</i>	ATCC 27280 and 3 laboratory isolates
<i>M. komossense</i>	ATCC 33013
<i>M. duvalii</i>	British Culture Collection NCTC 358
<i>M. sphagni</i>	ATCC 33027
<i>M. smegmatis</i>	TMC1515, TMC1533, and 7 clinical isolates
<i>M. gadium</i>	ATCC 22726
<i>M. vaccae</i>	TMC1526 and ATCC 15483
<i>M. austroafricanum</i>	ATCC 33464
<i>M. chubuense</i>	ATCC 27278 and 3 laboratory isolates
<i>M. aurum</i>	ATCC 25792 and ATCC 25793
<i>M. gilvum</i>	1 laboratory isolate
<i>M. parafortuitum</i>	ATCC 19696
<i>M. phlei</i>	TMC1516 and TMC1548
<i>M. neoaurum</i>	ATCC 25790, ATCC 25791, and ATCC 25794
<i>M. thermoresistibile</i>	ATCC 19527 and 9 clinical isolates
<i>M. fallax</i>	3 laboratory isolates
<i>Tsukamurella paurometabolum</i>	10 laboratory isolates
<i>M. diernhoferi</i>	ATCC 19340, ATCC 19341, ATCC 19344, ATCC 25958, and ATCC 25959
<i>M. tokaiense</i>	2 laboratory isolates
<i>Nocardia brasiliensis</i>	5 laboratory isolates
<i>Nocardia asteroides</i>	6 laboratory isolates
<i>Gordona sputi</i>	ATCC 29627
<i>Gordona bronchialis</i>	1 laboratory isolate
<i>Gordona rubropertincta</i>	ATCC 14352
<i>Gordona terrae</i>	ATCC 25594
<i>Rhodococcus erythropolis</i>	ATCC 25544
<i>Rhodococcus equi</i>	ATCC 6939, ATCC 1621, and 3 laboratory isolates
<i>Rhodococcus maris</i>	1 laboratory isolate
<i>Rhodococcus rhodni</i>	2 laboratory isolates
<i>Rhodococcus rhodochrous</i>	ATCC 13808
<i>Rhodococcus luteus</i>	1 laboratory isolate
<i>Rhodococcus coprophilus</i>	1 laboratory isolate
<i>Corynebacterium pseudotuberculosis</i>	ATCC 19410
<i>Corynebacterium diphtheriae</i>	1 clinical isolate
<i>Corynebacterium xerosis</i>	ATCC 3373
<i>Corynebacterium ulcerans</i>	1 clinical isolate
<i>Corynebacterium renale</i>	ATCC 1941

TABLE 2. Clinical and laboratory strains used to evaluate the *Mycobacterium* library and percentage of accurate identifications

Species	No. of strains evaluated	No. of incorrect identifications	% Accuracy
<i>M. tuberculosis</i>	649	1	99
<i>M. avium</i>	157	13	92
<i>M. intracellulare</i>	56	4	93
<i>M. scrofulaceum</i>	40	7	83
<i>M. gordonae</i>	47	2	96
<i>M. kansasii</i>	54	2	96
<i>M. terrae</i>	11	1	91
<i>M. nonchromogenicum</i>	20	1	95
<i>M. xenopi</i>	35	0	100
<i>M. fortuitum-M. peregrinum</i>	42	3	93
<i>M. chelonae</i>	30	1	97
<i>M. abscessus</i>	54	2	96
<i>M. marinum</i>	26	1	96
<i>M. chelonae-like</i>	16	4	75
<i>M. malmoense</i>	10	0	100
<i>M. simiae</i>	11	1	91
<i>M. szulgai</i>	23	0	100
<i>M. bovis</i> BCG	11	0	100
<i>M. haemophilum</i>	6	0	100
<i>M. asiaticum</i>	7	0	100
<i>M. shimoidei</i>	2	1	50
<i>M. gastri</i>	6	0	100
<i>M. celatum</i>	10	1	90
<i>T. paurometabolum</i>	4	0	100
Non- <i>Mycobacterium</i>	6	0	100

probes, with HPLC flow charts (2, 8), or by a combination of these techniques.

**Mycolic acid sample preparation and HPLC.** Cells were harvested from Lowenstein-Jensen slants as described previously (8). Briefly, cells were saponified with 25% potassium hydroxide in 50% methanol, extracted into chloroform, and then derivatized to UV-absorbing esters with *p*-bromophenacyl bromide. Mycolic acids were separated by chromatography as described previously (2). Briefly, samples were injected by using an Altex injector (model 210A) with a 20- $\mu$ l sample loop, and mycolic acid samples were separated by using a Beckman C-18 ultrasphere-XL analytical cartridge column with a particle size of 3  $\mu$ m (Beckman Instruments, Fullerton, Calif.). UV-absorbing esters were detected with a Beckman model 166 detector set at 260 nm. The column was equilibrated with 98% methanol–2% methylene chloride. Over a 1-min period, solvent concentrations were changed to 80% methanol–20% methylene chloride. Over the next 9 min, the solvent composition was changed linearly to 35% methanol–65% methylene chloride, with a flow rate of 2.5 ml/min. Over the next 0.5 min, the mobile phase was changed to the initial 98% methanol–2% methylene chloride condition and HPLC was continued for 1.5 min.

**Chromatographic peak labeling.** A high-molecular-weight standard (Ribi ImmunoChem Research, Inc., Hamilton, Mont.), which has been described previously (4), was used as an internal standard. Thirty-seven different peaks, including the internal standard, were labeled with an arbitrary naming scheme (Fig. 1). Each peak was automatically named by using peak identification tables tailored for variable ranges in internal standard (ISTD) elution times (Table 3). The peak identification tables were controlled by HPLC software (Beckman Instruments, Fullerton, Calif.).

**Library design.** A training set, or a group of strains compiled to construct the library, was developed with 45 species of *Mycobacterium*, 1 species of *Tsukamurella*, 2 species of *Nocar-*

*dia*, 7 species of *Rhodococcus*, 4 species of *Gordona*, and 5 species of *Corynebacterium*, for a total of 577 strains (Table 1). After HPLC analysis, chromatographic data from each sample were analyzed by using the Pirouette pattern recognition software. Training set samples of *Mycobacterium* species and *Tsukamurella paurometabolum* were defined by a species-specific numerical class; members of the non-*Mycobacterium* genera were combined and assigned to a single class. Peak height data were normalized for each sample used in the training set. By using peak heights to calculate a position, each sample was plotted into a matrix by the KNN algorithm (12). Once samples were defined into the matrix, the number of neighbors (*K*) that had to be assessed when predicting the identities of unknown samples was set to achieve the optimal segregation of each species.

**Library evaluation.** An independent set of 1,333 strains, including 24 species of *Mycobacterium*, was used to evaluate the accuracy of the *Mycobacterium* library. Following HPLC analysis, the identity of each sample was predicted by using the KNN algorithm in the Pirouette software. The distances between the unknown sample and each of the known samples in the precalculated matrix (library) were measured. An identification of the unknown sample was made by identifying the class of the nearest known samples and assigning the class that was shared by the majority of nearest neighbors to the unknown. Each class was representative of a specific species contained in the *Mycobacterium* library. The optimal number of neighbors (*K*) used to predict an unknown was determined previously and was implemented in the design of the library.

## RESULTS

Because HPLC software is currently limited in its ability to use relative retention times for peak naming, the use of real retention times was necessary for automated peak labeling. To standardize labeling, several identification tables were devised to correctly identify peaks regardless of shifts in chromatography. The identification tables were calculated on the basis of the elution time of the ISTD. Tables were created to cover shifts in ISTD elution times ranging from 9.31 to 10.00 min. We found that the automated peak labeling tables became unreliable in chromatograms with ISTD elution times of more than 10.00 min.

Results obtained from the evaluation set are outlined in Table 2. Of the 1,333 isolates, representing 24 species of *Mycobacterium*, we were able to correctly identify 97% of the strains examined by using the *Mycobacterium* library.

Only one strain of *M. tuberculosis* was missed by using the library; it was misidentified as *M. triviale*. Further examination revealed that this culture gave a pattern with visually obvious deviations from the chromatograms normally produced by *M. tuberculosis*. Only one isolate, an *M. kansasii* isolate, was misidentified as an *M. tuberculosis* complex member, specifically *M. bovis* BCG.

Of the 157 *M. avium* strains evaluated, a total of 13 were misidentified; 11 were misidentified as *M. intracellulare*, and two were misidentified as *M. scrofulaceum*. Four of 56 strains of *M. intracellulare* were misidentified, with three misidentified as *M. avium* and one misidentified as *M. scrofulaceum*. Similarly, most misidentifications of the *M. scrofulaceum* strains evaluated occurred with either *M. avium* or *M. intracellulare*; i.e., five were misidentified as *M. intracellulare* and one was misidentified as *M. avium*. One strain of *M. scrofulaceum* was misidentified as *M. gordonae* (double cluster). As supported by these findings, identification of *M. avium* and *M. intracellulare*

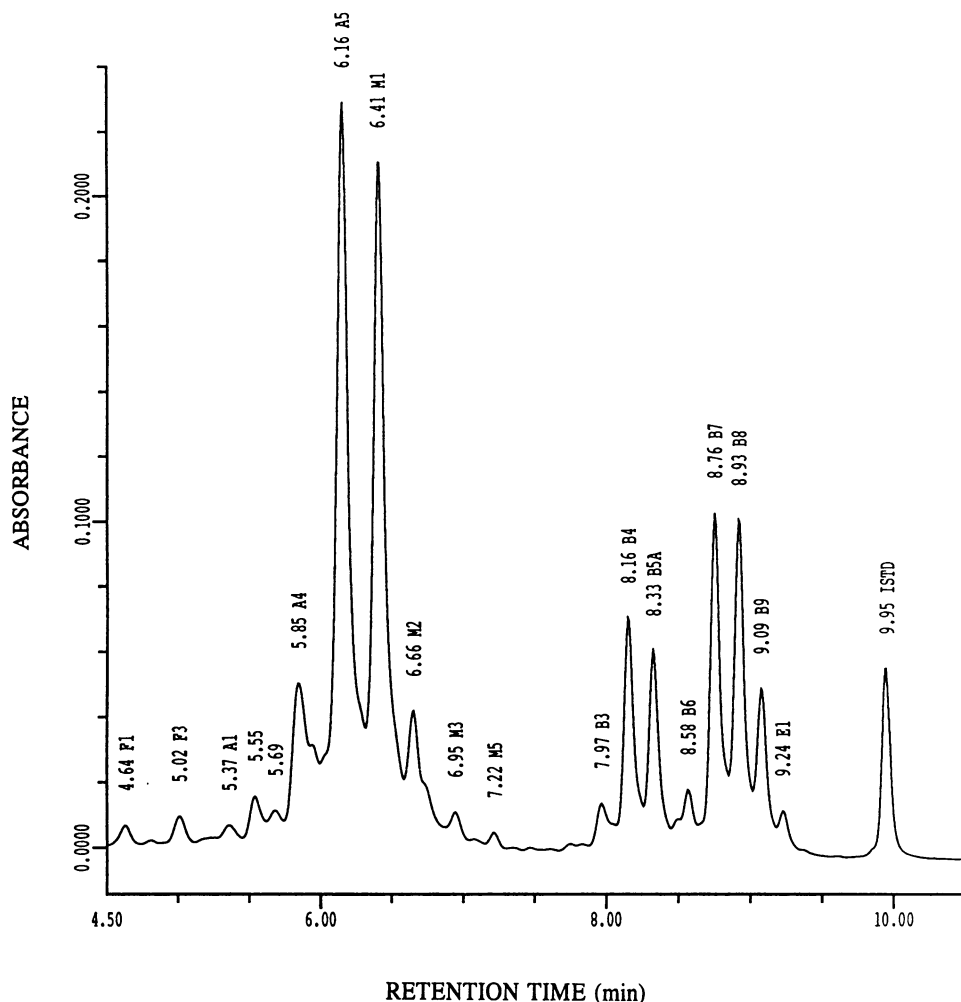


FIG. 1. Representative chromatogram of mycolic acid derivatives from *M. avium* illustrating many of the peaks labeled by the peak identification tables designed for use with the pattern recognition identification system.

to the *M. avium* complex level affords a higher degree of accuracy (98.5%).

Of the remaining slowly growing mycobacteria tested, one *M. celatum* (7) culture was misidentified as *M. avium*, one single-cluster *M. gordonae* was misidentified as *M. szulgai*, one double-cluster *M. gordonae* was misidentified as *M. intracellulare*, and one *M. kansasii* was misidentified as *M. bovis* BCG. Also, one *M. simiae* was misidentified as *M. flavescens*, one *M. nonchromogenicum* was misidentified as *M. terrae*, and one *M. terrae* was misidentified as *M. nonchromogenicum*. To summarize, 18 species of slowly growing mycobacteria were represented in the evaluation set. For the 1,155 slow growers examined, 97% were correctly identified by using the *Mycobacterium* library.

Six rapidly growing *Mycobacterium* species were included in the evaluation set. Ninety-three percent of the 168 rapid growers examined were correctly identified by using Pirouette software and the *Mycobacterium* library. Three strains of *M. fortuitum-M. peregrinum* were missed, with two misidentified as *M. haemophilum* and one misidentified as *M. chelonae*. One *M. chelonae* strain was misidentified as *M. abscessus*, and two strains of *M. abscessus* were misidentified as *M. chelonae*. Other misidentifications included one *M. marinum* being misidentified as a single-cluster *M. gordonae*.

TABLE 3. Peak identification for chromatograms with internal standard retention times ranging from 9.80 to 10.00 min

Peak name	Elution time (min)	Peak name	Elution time (min)
S1 .....	1.26-1.56	M2.....	6.27-6.57
S2 .....	1.39-1.69	M3.....	6.61-6.91
S3 .....	1.51-1.81	M4.....	6.80-7.10
S4A.....	1.61-1.91	M5.....	6.88-7.18
S4B.....	1.70-2.00	M6.....	7.10-7.40
S5 .....	1.85-2.15	B1.....	7.25-7.55
S6A.....	1.95-2.25	B2.....	7.45-7.75
S6B.....	2.02-2.32	B3.....	7.67-7.97
S7 .....	2.30-2.60	B4.....	7.85-8.15
S8 .....	2.63-2.93	B5A.....	8.00-8.30
S9 .....	2.98-3.28	B5B.....	8.15-8.45
F1 .....	4.29-4.59	B6.....	8.27-8.57
F2 .....	4.66-4.96	B7.....	8.45-8.75
F3 .....	4.85-5.15	B8.....	8.61-8.91
F4 .....	5.02-5.32	B9.....	8.77-9.07
A1 .....	5.21-5.51	E1.....	8.94-9.24
A4 .....	5.55-5.85	E2.....	9.09-9.39
A5 .....	5.82-6.12	ISTD.....	9.80-10.00
M1.....	6.08-6.38		

Ten samples of non-*Mycobacterium* species, including *T. paurometabolum* and strains of the genera *Corynebacterium*, *Rhodococcus*, *Gordona*, and *Nocardia*, were used to test the library. No isolates of these genera were misidentified, and none of the *Mycobacterium* isolates evaluated were categorized into the class representing these combined genera.

### DISCUSSION

The purpose of the endeavor described here was to provide a rapid, yet accurate, means of evaluating chromatographic data generated by HPLC for the identification of mycolic acid-containing genera, specifically *Mycobacterium* species. After an evaluation of the *Mycobacterium* library and the pattern recognition method, we concluded that the system is a rapid, reliable, relatively simple, and specific method for identification. Since this method reduces the amount of personnel time needed for evaluating chromatograms, it is cost-effective as well. Current pricing for the Pirouette software is about \$4,000, and the *Mycobacterium* library is available free of charge.

When compared with the flow chart method for evaluating chromatograms, our evaluation showed the pattern recognition method to be 99.7% specific and 100% sensitive for the *M. tuberculosis* complex and 95.7% specific and 100% sensitive for the *M. avium* complex. For rapidly growing mycobacteria, our evaluation of the *Mycobacterium* library showed 94% specificity for *M. chelonae*-*M. abscessus* strains and 88% specificity for *M. fortuitum*-*M. peregrinum* strains.

The chromatograms of the misidentified isolates were compared with the chromatograms used in the training set. In every case, the misidentified sample produced a chromatogram that was characteristic for its species yet that displayed chromatographic differences from those in the training set that were obvious with only a cursory visual comparison. For example, chromatograms of the misidentified strains displayed variations in peak heights that were not represented in the library. To avoid similar misidentifications in the future, the samples producing these chromatographic variations were added to the library.

Our laboratory was unable to evaluate every species included in the *Mycobacterium* library because of limited sample numbers in our inventory. In the case of species for which few samples were available, every sample was used to construct the library. Also, members of the non-*Mycobacterium* genera *Nocardia*, *Gordona*, *Rhodococcus*, and *Corynebacterium* were combined into a single category. Limited numbers of non-*Mycobacterium* samples prevented proper species definition and evaluation. However, efforts are under way to gather more samples in these poorly represented species for evaluation and further definition.

The Pirouette pattern recognition software was designed to analyze multivariate data sets and contains two classification algorithms, KNN and Soft Independent Modeling of Class Analogy (SIMCA) (15). The KNN algorithm was chosen as the classification technique best suited for the *Mycobacterium* library because of its greater prediction accuracy with sets of data containing both a high number of variables (peaks) and groups with small numbers of isolates. Also, better discrimination of strong subgroups or strain variations commonly found among mycobacterial species was achieved with the KNN algorithm than with the SIMCA algorithm. However, the KNN algorithm does not have confidence limits built into its algorithm as does the SIMCA algorithm. Thus far, attempts to use the SIMCA algorithm for verification of KNN algorithm identifications have been unsuccessful. Even with a high degree of confidence imposed upon the system, SIMCA algorithm

models used to verify KNN algorithm predictions were found to agree with incorrect KNN algorithm predictions.

Misclassifications can occur with the Pirouette software when a sample's chromatogram shows a variation not represented in the *Mycobacterium* library. Periodic upgrades have been and will be made to the library as unrepresented strains are found. Also, identification accuracy will be affected by poor chromatography and incorrect labeling of major peaks. However, a category was created to recognize poor samples and to classify them as such, and common variations in peak labeling have been included in the *Mycobacterium* library to minimize the number of inaccurate classifications caused by incorrectly labeled peaks.

Overall, the *Mycobacterium* library used in conjunction with the Pirouette software was found to provide a rapid, accurate system for identifying HPLC chromatograms produced by a broad spectrum of mycobacterial species.

### ACKNOWLEDGMENTS

We thank Eugene Bryant, South Carolina Department of Health and Environmental Control; Wayne Dupree, State of Louisiana Department of Health and Hospitals; Louise Thibert, Public Health Laboratory of Quebec; and Nancy Warren, Virginia State Health Department for culture contributions.

### REFERENCES

- Butler, W. R., D. G. Ahearn, and J. O. Kilburn. 1986. High-performance liquid chromatography of mycolic acids as a tool in the identification of *Corynebacterium*, *Nocardia*, *Rhodococcus*, and *Mycobacterium* species. *J. Clin. Microbiol.* **23**:182-185.
- Butler, W. R., K. C. Jost, and J. O. Kilburn. 1991. Identification of mycobacteria by high-performance liquid chromatography. *J. Clin. Microbiol.* **29**:2468-2472.
- Butler, W. R., and J. O. Kilburn. 1988. Identification of major slowly growing pathogenic mycobacteria and *Mycobacterium gordoniae* by high-performance liquid chromatography of their mycolic acids. *J. Clin. Microbiol.* **26**:50-53.
- Butler, W. R., and J. O. Kilburn. 1990. High-performance liquid chromatography patterns of mycolic acids as criteria for identification of *Mycobacterium chelonae*, *Mycobacterium fortuitum*, and *Mycobacterium smegmatis*. *J. Clin. Microbiol.* **28**:2094-2098.
- Butler, W. R., and J. O. Kilburn. 1992. Identification of difficult to grow mycobacteria with high-performance liquid chromatography, abstr. U-73, p. 177. Abstr. 92nd Gen. Meet. Am. Soc. Microbiol. 1992. American Society of Microbiology, Washington, D.C.
- Butler, W. R., J. O. Kilburn, and G. P. Kubica. 1987. High-performance liquid chromatography analysis of mycolic acids as an aid in the laboratory identification of *Rhodococcus* and *Nocardia* species. *J. Clin. Microbiol.* **25**:2126-2131.
- Butler, W. R., S. P. O'Connor, M. A. Yakrus, R. W. Smithwick, B. B. Plikaytis, C. W. Moss, M. M. Floyd, C. L. Woodley, J. O. Kilburn, F. S. Vadney, and W. M. Gross. 1993. *Mycobacterium celatum* sp. nov. *Int. J. Syst. Bacteriol.* **43**:539-548.
- Butler, W. R., L. Thibert, and J. O. Kilburn. 1992. Identification of *Mycobacterium avium* complex strains and some similar species by high-performance liquid chromatography. *J. Clin. Microbiol.* **30**:2698-2704.
- Floyd, M. M., V. A. Silcox, W. D. Jones, Jr., W. R. Butler, and J. O. Kilburn. 1992. Separation of *Mycobacterium bovis* BCG from *Mycobacterium tuberculosis* and *Mycobacterium bovis* by using high-performance liquid chromatography of mycolic acids. *J. Clin. Microbiol.* **30**:1327-1330.
- Guthertz, L. S., S. D. Lim, Y. Jang, and P. S. Duffey. 1993. Curvilinear-gradient high-performance liquid chromatography for the identification of mycobacteria. *J. Clin. Microbiol.* **31**:1876-1881.
- Kent, P. T., and G. P. Kubica. 1985. Public health mycobacteriology: a guide for the level III laboratory, p. 71-157. U.S. Department of Health and Human Services publication 86-8230. U.S.

- Department of Health and Human Services, Washington, D.C.
12. **Kowalski, B. R., and C. F. Bender.** 1972. Pattern recognition. A powerful approach to interpreting chemical data. *J. Am. Chem. Soc.* **94**:5633–5639.
  13. **Saito, H., H. Tomioka, K. Sato, H. Tasaka, and D. J. Dawson.** 1990. Identification of various serovar strains of *Mycobacterium avium* complex by using DNA probes specific for *Mycobacterium avium* and *Mycobacterium intracellulare*. *J. Clin. Microbiol.* **28**: 1694–1697.
  14. **Thibert, L., and S. LaPierre.** 1993. Routine application of high-performance liquid chromatography for identification of mycobacteria. *J. Clin. Microbiol.* **31**:1759–1763.
  15. **Wold, S.** 1976. Pattern recognition by means of disjoint principal components models. *Pattern Recognition* **8**:127–139.