Sequence-Based Identification of *Mycobacterium* Species Using the MicroSeq 500 16S rDNA Bacterial Identification System

JEAN BALDUS PATEL,^{1*} DEBRA G. B. LEONARD,¹ XAI PAN,² JAMES M. MUSSER,²
RICHARD E. BERMAN,³ AND IRVING NACHAMKIN¹

Department of Pathology and Laboratory Medicine, The University of Pennsylvania, Philadelphia,¹ and Pennsylvania State Public Health Laboratory, Pennsylvania Department of Health, Lionville,³ Pennsylvania, and Department of Pathology, Baylor College of Medicine, Houston, Texas²

Received 8 June 1999/Returned for modification 22 July 1999/Accepted 8 October 1999

We evaluated the MicroSeq 500 16S rDNA Bacterial Sequencing Kit (PE Applied Biosystems), a 500-bp sequence-based identification system, for its ability to identify clinical *Mycobacterium* isolates. The organism identity was determined by comparing the 16S rDNA sequence to the MicroSeq database, which consists primarily of type strain sequences. A total of 113 isolates (18 different species), previously recovered and identified by routine methods from two clinical laboratories, were analyzed by the MicroSeq method. Isolates with discordant results were analyzed by *hsp65* gene sequence analysis and in some cases repeat phenotypic identification, AccuProbe rRNA hybridization (Gen-Probe, Inc., San Diego, Calif.), or high-performance liquid chromatography of mycolic acids. For 93 (82%) isolates, the MicroSeq identity was concordant with the previously reported identity. For 18 (16%) isolates, the original identification was discordant with the MicroSeq identification. Of the 18 discrepant isolates, 7 (six unique sequences) were originally misidentified by phenotypic analysis or the AccuProbe assay but were correctly identified by the MicroSeq assay. Of the 18 discrepant isolates, 11 (seven unique sequences) were unusual species that were difficult to identify by phenotypic methods and, in all but one case, by molecular methods. The remaining two isolates (2%) failed definitive phenotypic identification, but the MicroSeq assay was able to definitively identify one of these isolates. The MicroSeq identification system is an accurate and rapid method for the identification of *Mycobacterium* spp.

The application of molecular techniques for identification of *Mycobacterium* spp. is becoming increasingly important. Isolate identification helps in evaluating the clinical significance of a positive culture and is important for predicting effective antibiotic therapy. Molecular identification provides two primary advantages to phenotypic identification: a more rapid turnaround time and improved accuracy in identification (28). Several clinical laboratories have developed their own molecular assays and databases for routine *Mycobacterium* identification. With these in-house-developed assays, the genetic targets vary, as does the method of target characterization. Two targets, which have proven to be effective, are the 16S rRNA gene (6, 9, 11, 17, 20, 23, 33–35) and the *hsp65* gene (2, 3, 5, 16, 18, 19, 32). The advantage of 16S rRNA gene analysis is that it can potentially be applied to the identification of all bacteria, whereas *hsp65* gene analysis is only useful for the identification of *Mycobacterium* spp. The most common methods of target characterization are amplification, followed by either probe hybridization, restriction length polymorphism analysis, or sequencing. Although sequence analysis requires more specialized equipment than the other methods, this technology is becoming less expensive. Sequencing also provides the highest level of resolution when looking for differences in a molecular target.

Sequencing the entire 16S rRNA gene (approximately 1,500 bp) is not a practical method for routine *Mycobacterium* spp.

identification. Rogall et al. (20) identified a hypervariable region (approximately 138 bp) in the 5' end of the gene which is sufficient for specific identification of most species. The MicroSeq 500 16S rDNA bacterial identification assay analyzes a larger portion (approximately 500 bp) of the same region. Unlike the assay developed by Rogall et al. (20), which uses primers specific for *Mycobacterium* spp., the primers used in the MicroSeq assay are generic for all bacteria, making this a potentially universal bacterial identification system.

The goal of the current study was to determine whether the MicroSeq 500 assay could be used to replace phenotypic characterization of most, if not all, *Mycobacterium* spp. isolated in a routine clinical laboratory. This assay was evaluated by blindly reidentifying a collection of previously identified *Mycobacterium* isolates representing a variety of species and including isolates that are difficult to identify by standard methods.

MATERIALS AND METHODS

Isolate selection and criteria for identification comparison. Isolates included in the study were previously recovered from clinical specimens and identified by either the AccuProbe rRNA hybridization assay (Gen-Probe, Inc., San Diego, Calif.), or phenotypic classification by using standard biochemical assays. One isolate was identified by high-performance liquid chromatography (HPLC) of mycolic acids in a reference laboratory (1). All isolates were identified in either the Clinical Microbiology Laboratory of the Hospital of the University of Pennsylvania (Philadelphia) or the Pennsylvania State Public Health Laboratory (Lionville). Both institutions have Level III Mycobacteriology Laboratories (24). Of the 113 isolates, 91 (15 species) came from the Hospital of the University of Pennsylvania, and 22 (9 species) came from the Pennsylvania State Public Health Laboratory. The isolates were chosen to represent a variety of different species and to include all difficult-to-identify species seen in these two laboratories during at least a 2-year period.

All isolates with discordant results between the original identification and the MicroSeq identification were retested by sequence analysis of the *hsp65* gene (16). Some of these isolates were also retested by either repeat phenotypic identification (7, 10, 14, 27), repeat probe hybridization (13), or HPLC analysis

* Corresponding author. Mailing address: Clinical Microbiology Laboratory, Department of Pathology and Laboratory Medicine, 4th Floor, Gates Bldg., 3400 Spruce St., Philadelphia, PA 19104-4283. Phone: (215) 662-6651. Fax: (215) 662-6655. E-mail: jbpatel@mail.med.upenn.edu.

of mycolic acids (1). The persons performing the MicroSeq assay or other assays for resolution of discrepancies were blinded to the previous results for the isolates. Identifications were counted as correct if two methods provided the same answer. A discordant result was defined as a MicroSeq database match with a species other than the species assigned by conventional identification, as long as the original species was present in the MicroSeq database. For example, if an isolate could not be definitively assigned to a species based upon biochemical classification, the MicroSeq identification was not counted as discordant. In some cases the MicroSeq identification was able to discriminate species within a complex of organisms that are not readily subdivided by the battery of biochemical tests used. These results were counted as concordant with the conventional identification as long as the species assigned by the MicroSeq assay was a known member of the complex.

Extraction, amplification, and sequencing of mycobacterial DNA. DNA extracts were made from pure cultures of mycobacteria as previously described (15). Briefly, a 1- μ l loopful of cells was suspended in 200 μ l of TE buffer (10 mM Tris, 1 mM EDTA; pH 8.0), heat killed at 95°C for 15 min, and mechanically disrupted with glass beads. The bacterial extract was separated from the beads by centrifugation and stored at -20°C until needed. A 500-bp 16S ribosomal DNA (rDNA) fragment was amplified from the 5' end of the gene in a reaction volume of 50 μ l (25 μ l of MicroSeq PCR master mix, 24 μ l of sterile H₂O, and 1 μ l of bacterial extract). Amplified products were purified with the Qiagen PCR Purification Kit (Valencia, Calif.), according to the manufacturer's recommendations, prior to sequencing. Forward and reverse sequencing reactions were performed for each amplified product. The sequencing reactions consisted of 13 μ l of MicroSeq sequencing mix, 4 μ l of sterile distilled water, and 3 μ l of purified amplified product. Sequencing reactions were purified with Centri-Sep columns (Princeton Separation, Princeton, N.J.) according to the manufacturer's instructions, and all sequence analysis was performed on an ABI PRISM 310 Genetic Analyzer (Perkin-Elmer/Applied Biosystems, Foster City, Calif.).

Sequence data analysis. All of the sequencing data was analyzed with the MicroSeq software version 1.36. The analysis steps were as follows: (i) assembly of the reverse and forward sequences into a consensus sequence; (ii) editing of the consensus sequence to resolve discrepancies between the two strands by evaluation of the electropherograms; and (iii) comparison of the consensus sequence to the *Mycobacterium* entries in the MicroSeq database. The database comparison, using the Full Alignment Tool of the MicroSeq software, generated a list of the closest matches with a distance score (40). This distance score indicated the percent difference between the unknown sequence and the database sequence. For the purpose of comparing an isolate's original identification to its MicroSeq identification, the MicroSeq identity was considered to be the closest match in the MicroSeq database no matter what the distance score was.

RESULTS

Performance of the MicroSeq 500 assay for identification of *Mycobacterium* isolates. A total of 113 isolates, representing 18 different species, were subject to 16S rDNA sequence analysis by the MicroSeq 500 assay. Of the 113 isolates, 93 (82%) had concordant results between the original identification and the MicroSeq identification, and 18 (16%) isolates had a molecular identification that was discordant with the original identification. The results for two isolates (2%) were not counted as concordant or discordant because the isolates could not be classified as recognized species by phenotypic analysis (Table 1).

The MicroSeq 500 assay was better able to distinguish between species than phenotypic identification in some cases. For example, the MicroSeq assay was able to subdivide isolates in the *M. avium-intracellulare* complex into species. In addition, all isolates previously grouped in the *M. terrae-triviale* complex were identified as *M. nonchromogenicum*. Other studies have reported that most clinical isolates reported as *M. terrae-triviale* complex are actually *M. nonchromogenicum* (21). Finally, one isolate originally reported as *M. fortuitum* was identified as *M. peregrinum* by 16S rDNA sequence. *M. peregrinum* was previously classified as a subspecies of *M. fortuitum* (12). However, the assay was unable to distinguish between some species. There was no separation of species within the *M. tuberculosis* complex. Likewise, the 500 bp of 16S rDNA sequence was insufficient to distinguish between *M. chelonae* and *M. abscessus*, between *M. genavensae* and *M. simiae*, and between *M. kansasii* and *M. gastri*. This finding is consistent with previous reports of *Mycobacterium* spp. identification by 16S rDNA sequence (Table 2) (9). It should be noted that the entire 16S

TABLE 1. Summary of isolates and identification comparison^a

Isolate analyzed ^b	No. tested	No. concordant (%)	No. discordant (%)
<i>M. tuberculosis</i> complex	5	5 (100)	0 (0)
<i>M. avium-intracellulare</i> complex	6	6 (100)	0 (0)
<i>M. kansasii</i>	6	6 (100)	0 (0)
<i>M. gordonae</i>	6	6 (100)	0 (0)
<i>M. xenopi</i>	11	9 (82)	2 (18)
<i>M. marinum</i>	9	9 (100)	0 (0)
<i>M. terrae-triviale</i> complex	9	9 (100)	0 (0)
<i>M. szulgai</i>	6	4 (67)	2 (33)
<i>M. scrofulaceum</i>	13	6 (46)	7 (54)
<i>M. flavescens</i>	3	0 (0)	3 (100)
<i>M. alcapulcensis</i>	1	0 (0)	1 (100)
<i>M. celatum</i>	1	1 (100)	0 (0)
<i>M. chelonae</i>	8	8 (100)	0 (0)
<i>M. abscessus</i>	7	6 (86)	1 (14)
<i>M. mucogenicum</i>	11	9 (82)	2 (18)
<i>M. fortuitum</i>	9	9 (100)	0 (0)
<i>M. asiaticum</i> -like	1	NA	NA
<i>M. scrofulaceum</i> , <i>M. intracellulare</i> -like	1	NA	NA
Total	113	93 (82)	18 (16)

^a See Materials and Methods for definitions of concordant and discordant. NA, not applicable.

^b Listed by original identification results.

rDNA sequence is identical for *M. genavensae* and *M. simiae*, as well as for *M. kansasii* and *M. gastri*. There are 16S rDNA sequence differences between *M. chelonae* and *M. abscessus*, but these occur in the 5' region of the gene.

The 18 discrepant isolates, representing 13 unique 16S rDNA sequences, could be grouped into two categories: (i) commonly encountered species that were originally misidentified by phenotypic identification or by initial probe hybridization, and (ii) rarely isolated species that are difficult to identify by phenotypic methods (Table 3). Seven discrepancies (six unique sequences) fell into the first category. Two isolates identified as *M. xenopi* by phenotypic analysis were identified as *M. avium* by the MicroSeq assay. Repeat AccuProbe assays and *hsp65* sequence confirmed the *M. avium* species assignments. Both isolates were originally tested with the *M. avium-intracellulare* probe but were negative. Similarly, two isolates of *M. gordonae* were misidentified as *M. scrofulaceum* based upon phenotypic analysis. Once again, the original *M. gordonae* AccuProbe assays were negative, but *hsp65* sequence and repeat AccuProbe assays confirmed the *M. gordonae* identifications. One of these isolates was resubmitted for identification by phenotypic analysis. Like the original analysis, it was reported to be *M. scrofulaceum*. This isolate had a positive 14-day aryl-sulfatase test like the majority of *M. gordonae* isolates, but it was also positive for urease and negative for Tween hydrolysis, suggesting *M. scrofulaceum* (7). Two rapidly growing mycobacteria isolates were misidentified. An isolate originally identified as *M. abscessus* was determined to be *M. fortuitum* by molecular identification, and an isolate first identified as *M. mucogenicum* was identified as *M. abscessus* or *M. chelonae* by the MicroSeq assay and as *M. chelonae* by *hsp65* sequence. Repeat phenotypic identification agreed with the MicroSeq identification for both of these isolates. The original misidentifications were attributed to misinterpretation of either a polymyxin B susceptibility (the former misidentification) or a cephalothin susceptibility (the latter misidentification) (39). Finally, an isolate related to *M. nonchromogenicum* was originally identified as *M. mucogenicum*. Repeat phenotypic classification again

TABLE 2. Concordant results between conventional identification and MicroSeq identification techniques

Conventional identification	No. concordant	MicroSeq identification	No.	Difference (%) ^a	Comment
<i>M. tuberculosis</i> complex ^b	5	<i>M. tuberculosis</i> complex	5	0.00	
<i>M. avium-intracellulare</i> complex	6	<i>M. intracellulare</i>	1	0.00	
		<i>M. avium</i>	5	0.00	
<i>M. kansasii</i> ^b	6	<i>M. kansasii</i> - <i>M. gastri</i>	6	0.00	
<i>M. gordonae</i> ^b	6	<i>M. gordonae</i>	2	0.00	
		<i>M. gordonae</i>	2	0.20	
		<i>M. gordonae</i>	1	0.78	
		<i>M. gordonae</i>	1	0.98	
<i>M. xenopi</i>	9	<i>M. xenopi</i>	1	0.00	
		<i>M. xenopi</i>	2	0.10	
		<i>M. xenopi</i>	1	0.19	
		<i>M. xenopi</i>	5	0.29	All have same sequence
<i>M. marinum</i>	9	<i>M. marinum</i>	9	0.00	
<i>M. terrae-triviale</i> complex	9	<i>M. nonchromogenicum</i>	7	0.97	All have same sequence
		<i>M. nonchromogenicum</i>	2	1.16	Two different sequences
<i>M. szulgai</i>	4	<i>M. szulgai</i>	4	0.00	
<i>M. scrofulaceum</i>	6	<i>M. scrofulaceum</i>	3	0.00	
		<i>M. scrofulaceum</i>	3	0.59	All have same sequence
<i>M. celatum</i>	1	<i>M. celatum</i>	1	0.19	
<i>M. chelonae</i>	8	<i>M. abscessus</i> - <i>M. chelonae</i>	7	0.00	
		<i>M. abscessus</i> - <i>M. chelonae</i>	1	1.00	
<i>M. abscessus</i>	6	<i>M. abscessus</i> - <i>M. chelonae</i>	6	0.00	
<i>M. mucogenicum</i>	9	<i>M. mucogenicum</i>	7	0.00	
		<i>M. mucogenicum</i>	1	0.10	
		<i>M. mucogenicum</i>	1	0.15	
<i>M. fortuitum</i>	9	<i>M. fortuitum</i> subsp. <i>fortuitum</i>	8	0.00	
		<i>M. peregrinum</i>	1	0.00	

^a Percent difference between the isolate's sequence and the sequence of the type strain as calculated by the Full Alignment Tool of the MicroSeq software (see Materials and Methods).

^b Identified by AccuProbe assay. All other isolates were identified by biochemical analysis.

resulted in an identification of *M. mucogenicum*, but HPLC analysis of the mycolic acids put this isolate in the *M. terrae-triviale* complex.

The remaining 11 discrepancies (seven unique sequences) consisted of rarely isolated species that were difficult to identify by biochemical classification. These isolates were first identified phenotypically as *M. scrofulaceum*, *M. acapulcensis*, *M. szulgai*, and *M. flavescens* (Table 3). All but one isolate failed definitive identification by sequence analysis of the *hsp65* gene and had 16S rRNA gene sequences that were between 0.80 and 2.79% different from the sequence of the most closely related type strain. Based upon the 500-bp 16S rDNA sequence, these isolates were related to *M. genavense*-*M. simiae*, *M. malmoense*, *M. flavescens*, *M. triplex*, *M. gadium*, or *M. thermoresistibile*. The one exception was an isolate originally called *M. flavescens*. This isolate's 16S rDNA sequence was only 0.20% different from the sequence of the *M. phlei* type strain, and the *hsp65* sequence also matched that of *M. phlei*.

Two isolates could not be identified by phenotypic criteria, and only one of them had an exact match in the MicroSeq database. One isolate, called *M. asiaticum*-like, was equally related to *M. interjectum*, *M. triplex*, and *M. genavense*-*M. simiae*. The remaining isolate was first reported as *M. scrofulaceum*, *intracellulare*-like based upon biochemical identification but it was definitively identified as *M. triplex* by the MicroSeq assay (Table 3). *M. triplex* is a recently recognized species which was not on the identification chart, but review of the isolate's biochemical profile was consistent with a report characterizing this species (4).

DISCUSSION

Rapid identification of *Mycobacterium* spp. is becoming increasingly important for good patient care. Phenotypic identi-

fication can take between 2 and 8 weeks to complete. In the meantime, the patient may receive unnecessary or insufficient antimicrobial therapy. For this reason alone, molecular identification of *Mycobacterium* spp. is likely to become the standard of care. Commercial probe hybridization assays already provide a rapid identification with a turnaround time of less than 1 day, but these assays can only test for one species at a time, and probes are available for only four species. Until now, all molecular assays for the universal identification of *Mycobacterium* isolates have been in-house-developed assays. The MicroSeq 500 assay is a commercial bacterial identification assay which can be applied to the routine identification of clinical *Mycobacterium* isolates. The turnaround time for this assay is 2 days, requiring approximately 4 h of a technologist's time. A detailed description of the MicroSeq 1500 assay based on a 1,500-bp region was previously published (31). The MicroSeq 500 assay has the same steps, but sequences 500 bp by using only one amplification reaction and two sequencing reactions, whereas the 1,500-bp assay requires 1 amplification reaction and 12 sequencing reactions.

In addition to providing a rapid identification, the MicroSeq identification also proved to be more accurate than the phenotypic identification. This finding is consistent with a previous report comparing phenotypic identification to molecular identification (28). In the current study, seven isolates were misidentified by phenotypic analysis. In addition, four of these isolates, two *M. avium* and two *M. gordonae* isolates, failed identification by the AccuProbe hybridization assay during the initial work-up. Such misidentification could potentially have serious clinical consequences. For example, isolation of *M. scrofulaceum* is more likely to be considered clinically significant than isolation of *M. gordonae*. In addition, confusing *M. fortuitum* with *M. abscessus* may unnecessarily limit the choice

TABLE 3. Discordant results between conventional identification and MicroSeq identification techniques

Conventional identification	No. discordant	MicroSeq identification	No.	Difference (%) ^a	Identity by alternative method ^f
Commonly encountered Species originally misidentified					
<i>M. mucogenicum</i>	2	<i>M. abscessus</i> - <i>M. chelonae</i>	1	0.20	<i>M. chelonae</i> (hsp65)
		<i>M. nonchromogenicum</i> ^c	1	0.97	<i>M. nonchromogenicum</i> (hsp65), <i>M. terrae</i> complex (HPLC analysis), <i>M. mucogenicum</i> (rpt phenotypic ID)
<i>M. xenopi</i>	2	<i>M. avium</i>	2	0.00	<i>M. avium</i> (probe and hsp65)
<i>M. abscessus</i>	1	<i>M. fortuitum</i>	1	0.00	<i>M. fortuitum</i> (hsp65, rpt phenotypic ID)
<i>M. scrofulaceum</i>	2	<i>M. gordonae</i>	1	0.00	<i>M. gordonae</i> (probe and hsp65)
		<i>M. gordonae</i>	1	0.98	<i>M. gordonae</i> (probe and hsp65), <i>M. scrofulaceum</i> (rpt phenotypic ID)
Rarely isolated species difficult to identify by phenotypic methods					
<i>M. scrofulaceum</i>	5	<i>M. genavensae</i> - <i>M. simiae</i> ^d	4	1.39	No close relative (hsp65)
		<i>M. malmoense</i>	1	0.97	Related to <i>M. interjectum</i> (hsp65)
<i>M. alcapulcensis</i> ^b	1	<i>M. flavescens</i> ^e	1	2.79	No close relative (hsp65)
<i>M. szulgai</i>	2	<i>M. flavescens</i> ^e	1	2.79	No close relative (hsp65)
		<i>M. triplex</i>	1	1.39	No close relative (hsp65)
<i>M. flavescens</i>	3	<i>M. gadium</i>	1	0.80	<i>M. aurum</i> -like (hsp65)
		<i>M. thermoresistibile</i>	1	2.39	No close relative (hsp65)
		<i>M. phlei</i>	1	0.20	<i>M. phlei</i> (hsp65)
Isolates without definitive phenotypic identification					
<i>M. asiaticum</i> -like	NA	<i>M. interjectum</i> ; <i>M. triplex</i> ; <i>M. genavense</i> - <i>M. simiae</i>	1	1.79	No close relative (hsp65)
<i>M. scrofulaceum</i> , <i>M. intracellulare</i> -like	NA	<i>M. triplex</i>	1	0.00	related to <i>M. genavense</i> (hsp65)

^a Percent difference between the isolate's sequence and the sequence of the type strain as calculated by the Full Alignment Tool of the MicroSeq software (see Materials and Methods).

^b Identified by HPLC analysis of mycolic acids. All other isolates were identified by biochemical analysis.

^c This sequence is identical to the seven other sequences that are most closely related to *M. nonchromogenicum* with a 0.97% difference (see Table 2).

^d All of these sequences are identical.

^e This sequence is identical to the one other sequence that is most closely related to *M. flavescens* with a 2.79% difference.

^f hsp65, sequence analysis; rpt phenotypic ID, biochemical profiling; probe, AccuProbe assay.

of antibiotic therapy. *M. fortuitum* is more likely to respond to the quinolones, sulfonamide, and tetracycline than is *M. abscessus* (29, 30, 36). In contrast, mistaking *M. chelonae* with *M. mucogenicum* may result in the use of ineffective antibiotic therapy. *M. chelonae* is resistant to cefoxitin and sulfonamide, whereas *M. mucogenicum* is usually susceptible (29, 37, 38). Inaccurate phenotypic identifications most likely result from normal phenotypic variation within a species and lack of experience among technologists who perform and read the biochemical assays. The reason for the AccuProbe failures is not clear. All of the isolates in question were correctly identified by AccuProbe upon repeat testing. Since there is no evidence that the original cultures were mixed, the most likely explanations for the false-negative results are either insufficient inoculum or incomplete cell lysis. Identification by 16S rDNA sequence avoids some of the problems associated with these methods. The 16S rDNA sequence is relatively conserved, so there is little intraspecies variation, but in most cases it contains enough heterogeneity to distinguish between species. In addition, sequence analysis requires less judgment on the part of technologists for interpretation and, unlike identification by probe hybridization, sequencing is not subject to false-negative results.

Another important advantage of molecular identification is the classification of unusual isolates. Unusual isolates included in this study either failed definitive phenotypic identification or were misidentified as other species, most commonly *M. scrofulaceum*, *M. flavescens*, or *M. szulgai*. Only two of these isolates, identified as *M. triplex* and *M. phlei* based on the 16S rDNA

sequence, had close matches in the MicroSeq database. Although the other isolates did not have close matches in the database, the 500 bases of the sequence provided important information. Based upon this sequence, other databases can be searched and the sequence can be compared to other sequences from clinical isolates. In this study, we found two isolates, with identical sequences, that were most closely related to *M. flavescens* (2.79% difference). Originally these two isolates were assigned to different species. Without sequence analysis we may not have recognized these isolates as closely related and, very likely, identical species. This sequence was also used to search GenBank and the Ribosomal Database library (www.cme.msu.edu/RDP) for related sequences. The closest match from both databases was to a sequence from a recently reported isolate, *M. novocastrense* (0.70% difference between sequences) (25).

For all isolates that do not have an exact database match, it can be helpful to construct a phylogenetic tree with the sequence data. The spacial relationship of the unknown sequence to known sequences helps to determine if the isolate represents a novel species. The MicroSeq software provides two tree-making tools, one that is based upon the Unweighted Pair Group Method using Averages (UPGMA) (26) and another that uses the neighbor-joining pair group method (22). However, tree analysis will provide only limited clinical and phylogenetic information. For unusual isolates (i.e., isolates that are not typically recovered from clinical specimens), location within a tree does not necessarily predict clinical relevance or susceptibility. In addition, the short sequence analyzed by

this assay may not provide enough information for an accurate phylogenetic classification.

In many cases the MicroSeq assay is better able to discriminate between species than phenotypic analysis. However, the portion of sequence analyzed by this assay is not sufficient to distinguish between *M. chelonae* and *M. abscessus*, between *M. genavensae* and *M. simiae*, and between *M. kansasii* and *M. gastri*. Clinically, the most important pairs to distinguish from each other are *M. chelonae* and *M. abscessus*. This distinction is important because antibiotic susceptibility can vary between these two species and, depending upon the site of infection, these species vary in their clinical significance (24). *M. chelonae* and *M. abscessus* can be distinguished from each other by phenotypic assays (e.g., the salt tolerance test), molecular characterization of the 16S rRNA gene 3' end where sequence differences occur, or sequence analysis of the *hsp65* (19, 28).

The MicroSeq assay has two advantages that in-house-developed assays do not provide: commercially prepared reagents and a commercially prepared database. The premade amplification and sequencing master mixes significantly decrease technologist time for reagent preparation and quality control. The MicroSeq database, composed primarily of sequences from type strains, has several advantages over in-house-developed databases and public databases. First, the database provides a practical alternative to in-house database development. The development of a database would require either sequencing a collection of isolates which have been identified by an alternative method or compiling sequences from public databases. Both processes are labor-intensive. Second, the MicroSeq database is likely to be more accurate than a database which was developed by using only biochemical profiling as the reference identity. The type strains in the MicroSeq database are classified by using a polyphasic approach and are considered to be the prototype for a species. But sequences deposited in public databases and strains (other than type strains) deposited in culture collections are not monitored. Therefore, species assignments are made by any criteria that the depositor chooses. As a result, caution must be exercised when using this information for clinical purposes. Finally, the MicroSeq database is very extensive, as it contains a total of 63 unique sequences.

A weakness of the MicroSeq database is that it has only one entry per species. This is a problem when the unknown isolate does not have an exact match in the database. The software is capable of comparing an unknown isolate to previously sequenced isolates, thereby expanding the database through use and collected experience. However, this is most helpful for addressing possible epidemiological issues. The problem of species assignment and reporting still remains. The genetic difference between closely related species of *Mycobacterium* in the database is quite variable. For example, the genetic difference between the *M. mucogenicum* sequence and the *M. farcinogenes* sequence is only 0.40%. In contrast, the most similar sequence to the *M. xenopi* sequence belongs to *M. shimoides*, but these two sequences differ by 4.36%. To implement this assay in a clinical laboratory, it is helpful to establish reporting criteria. At the Hospital of the University of Pennsylvania, we chose to establish three categories of reports. We will report an isolate either as a distinct species, as "related to" a species, or as "most closely related to" a species depending upon the amount of sequence difference between the unknown isolate and the database entries. A cutoff of <0.80% difference was chosen for species identity. Some of the rapidly growing *Mycobacterium* spp. differ by less than 0.80%, but we saw little variation among the rapid growers included in this study. An isolate is reported as "related to" the closest database match if

the genetic difference is ≥ 0.80 and $\leq 1.50\%$. In these cases, the isolates may be the same species. For example, sequence diversity has been noted within the species *M. gordonae* (8). But for some species, like *M. szulgai*, a 0.97% difference can indicate a distinct species. Finally, in cases where the genetic difference is $>1.50\%$, the isolate is reported as a unique isolate that is "most closely related to" the best database match. These isolates most likely represent novel species.

The primary disadvantage to the MicroSeq assay is the cost. Although the price of PCR and automated sequencing technology are decreasing, the reagents, equipment, and software necessary for this assay are significantly more expensive than biochemical profiling or the ribosomal probe hybridization. This will limit the ability of hospital-based laboratories to utilize this assay. However, the application of this assay is not limited to *Mycobacterium* spp. It can also be employed for the identification of other unusual or slow-growing bacteria (31). A broader application of this technology will likely make it more cost effective in the future.

ACKNOWLEDGMENTS

We thank the technologists at the Hospital of the University of Pennsylvania and the Pennsylvania State Public Health Laboratories for their assistance in collecting isolates and repeating phenotypic identifications.

This study was supported in part by Perkin-Elmer/Applied Biosystems.

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.
- Devallois, A., K. S. Goh, and N. Rastogi. 1997. Rapid identification of mycobacteria to species level by PCR-restriction fragment length polymorphism analysis of the *hsp65* gene and proposition of an algorithm to differentiate 34 mycobacterial species. *J. Clin. Microbiol.* **35**:2969-2973.
- Fiss, E. H., F. F. Chehab, and G. F. Brooks. 1992. DNA amplification and reverse dot blot hybridization for detection and identification of mycobacteria to the species level in the clinical laboratory. *J. Clin. Microbiol.* **30**:1220-1224.
- Floyd, M. M., L. S. Guthertz, V. A. Silcox, P. S. Duffey, Y. Jang, E. P. Desmond, J. T. Crawford, and W. R. Butler. 1996. Characterization of an SAV organism and proposal of *Mycobacterium triplex* sp. nov. *J. Clin. Microbiol.* **34**:2963-2967.
- Hance, A. J., B. Grandchamp, V. Levy-Frebault, D. Lecossier, J. Razulier, D. Bocart, and B. Gicquel. 1989. Detection and identification of mycobacteria by amplification of mycobacterial DNA. *Mol. Microbiol.* **3**:843-849.
- Hughes, M. S., R. A. Skuce, L. A. Beck, and S. D. Neill. 1993. Identification of mycobacteria from animals by restriction enzyme analysis and direct DNA cycle sequencing of polymerase chain reaction-amplified 16S rRNA gene sequences. *J. Clin. Microbiol.* **31**:3216-3222.
- Kent, P. T., and G. P. Kubica. 1985. Public health mycobacteriology: a guide for the level III laboratory. Centers for Disease Control, Atlanta, Ga.
- Kirschner, P., and E. C. Bottger. 1992. Microheterogeneity within rRNA of *Mycobacterium gordonae*. *J. Clin. Microbiol.* **30**:1049-1050.
- Kirschner, P., B. Springer, U. Vogel, A. Meier, A. Wrede, M. Kiekenbeck, F. C. Bange, and E. C. Bottger. 1993. Genotypic identification of mycobacteria by nucleic acid sequence determination: report of a 2-year experience in a clinical laboratory. *J. Clin. Microbiol.* **31**:2882-2889.
- Koneman, E. W., S. D. Allen, W. M. Janda, P. C. Schreckenberger, Jr., and W. C. Winn (ed.). 1997. Mycobacteria, p. 893-952. In E. W. Koneman, S. D. Allen, W. M. Janda, P. C. Schreckenberger, Jr., and W. C. Winn (ed.), Color atlas and textbook of diagnostic microbiology. Lippincott-Raven Publishers, Philadelphia, Pa.
- Kox, L. F., J. van Leeuwen, S. Knijper, H. M. Jansen, and A. H. Kolk. 1995. PCR assay based on DNA coding for 16S rRNA for detection and identification of mycobacteria in clinical samples. *J. Clin. Microbiol.* **33**:3225-3233.
- Kusunoki, S., and T. Ezaki. 1992. Proposal of *Mycobacterium peregrinum* sp. nov., nom. rev., and elevation of *Mycobacterium chelonae* subsp. *abscessus* (Kubica et al.) to species status: *Mycobacterium abscessus* comb. nov. *Int. J. Syst. Bacteriol.* **42**:240-245.
- Lebrun, L., F. Espinasse, J. D. Poveda, and V. Vincent-Levy-Frebault. 1992. Evaluation of nonradioactive DNA probes for identification of mycobacteria. *J. Clin. Microbiol.* **30**:2476-2478.

14. Metchock, B. G., F. S. Nolte, and R. J. Wallace. 1995. *Mycobacterium*, p. 399–437. In P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. H. Tenover (ed.), *Manual of clinical microbiology*. ASM Press, Washington, D.C.
15. Nachamkin, I., C. Kang, and M. P. Weinstein. 1997. Detection of resistance to isoniazid, rifampin, and streptomycin in clinical isolates of *Mycobacterium tuberculosis* by molecular methods. *Clin. Infect. Dis.* **24**:894–900.
16. Pai, S., N. Esen, X. Pan, and J. M. Musser. 1997. Routine rapid *Mycobacterium* species assignment based on species-specific allelic variation in the 65-kilodalton heat shock protein gene (*hsp65*). *Arch. Pathol. Lab. Med.* **121**:859–864.
17. Patel, S., M. Yates, and N. A. Saunders. 1997. PCR-enzyme-linked immunosorbent assay and partial rRNA gene sequencing: a rational approach to identifying mycobacteria. *J. Clin. Microbiol.* **35**:2375–2380.
18. Plikaytis, B. B., B. D. Plikaytis, M. A. Yakus, W. R. Butler, C. L. Woodley, V. A. Silcox, and T. M. Shinnick. 1992. Differentiation of slowly growing *Mycobacterium* species, including *Mycobacterium tuberculosis*, by gene amplification and restriction fragment length polymorphism analysis. *J. Clin. Microbiol.* **30**:1815–1822.
19. Ringuet, H., C. Akoua-Koffi, S. Honore, A. Varnerot, V. Vincent, P. Berche, J. L. Gaillard, and C. Pierre-Audigier. 1999. *hsp65* sequencing for identification of rapidly growing mycobacteria. *J. Clin. Microbiol.* **37**:852–857.
20. Rogall, T., T. Flohr, and E. C. Bottger. 1990. Differentiation of *Mycobacterium* species by direct sequencing of amplified DNA. *J. Gen. Microbiol.* **136**:1915–1920.
21. Rogall, T., J. Wolters, T. Flohr, and E. C. Bottger. 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium*. *Int. J. Syst. Bacteriol.* **40**:323–330.
22. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic Trees. *Mol. Biol. Evol.* **4**:406–425.
23. Sanguinetti, M., B. Posteraro, F. Ardito, S. Zanetti, A. Cingolani, L. Sechi, A. De Luca, L. Ortona, and G. Fadda. 1998. Routine use of PCR-reverse cross-blot hybridization assay for rapid identification of *Mycobacterium* species growing in liquid media. *J. Clin. Microbiol.* **36**:1530–1533.
24. Scientific Assembly on Microbiology, Tuberculosis, and Pulmonary Infections, American Thoracic Society. 1997. Diagnosis and treatment of disease caused by nontuberculous mycobacteria. *Am. J. Respir. Crit. Care Med.* **156**:S1–S25.
25. Shojai, H., M. Goodfellow, J. G. Magee, R. Freeman, F. K. Gould, and C. G. Brignall. 1997. *Mycobacterium novocastrense* sp. nov., a rapidly growing photochromogenic mycobacterium. *Int. J. Syst. Bacteriol.* **47**:1205–1207.
26. Sokal, R. R., and P. H. A. Sneath. 1963. *Principles of numerical taxonomy*. W. H. Freeman, San Francisco, Calif.
27. Springer, B., E. C. Bottger, P. Kirschner, and R. J. Wallace, Jr. 1995. Phylogeny of the *Mycobacterium chelonae*-like organism based on partial sequencing of the 16S rRNA gene and proposal of *Mycobacterium mucogenicum* sp. nov. *Int. J. Syst. Bacteriol.* **45**:262–267.
28. Springer, B., L. Stockman, K. Teschner, G. D. Roberts, and E. C. Bottger. 1996. Two-laboratory collaborative study on identification of mycobacteria: molecular versus phenotypic methods. *J. Clin. Microbiol.* **34**:296–303.
29. Swenson, J. M., C. Thornsberry, and V. A. Silcox. 1982. Rapidly growing mycobacteria: testing of susceptibility to 34 antimicrobial agents by broth microdilution. *Antimicrob. Agents Chemother.* **22**:186–192.
30. Swenson, J. M., R. J. Wallace, Jr., V. A. Silcox, and C. Thornsberry. 1985. Antimicrobial susceptibility of five subgroups of *Mycobacterium fortuitum* and *Mycobacterium chelonae*. *Antimicrob. Agents Chemother.* **28**:807–811.
31. Tang, Y. W., N. M. Ellis, M. K. Hopkins, D. H. Smith, D. E. Dodge, and D. H. Persing. 1998. Comparison of phenotypic and genotypic techniques for identification of unusual aerobic pathogenic gram-negative bacilli. *J. Clin. Microbiol.* **36**:3674–3679.
32. Telenti, A., F. Marchesi, M. Balz, F. Bally, E. C. Bottger, and T. Bodmer. 1993. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *J. Clin. Microbiol.* **31**:175–178.
33. Troesch, A., H. Nguyen, C. G. Miyada, S. Desvarenne, T. R. Gingeras, P. M. Kaplan, P. Cros, and C. Mabilat. 1999. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. *J. Clin. Microbiol.* **37**:49–55.
34. van der Vliet, G. M. E., R. A. F. Schukink, B. van Gemen, P. Schepers, and P. R. Klatser. 1993. Nucleic acid sequence-based amplification (NASBA) for the identification of mycobacteria. *J. Gen. Microbiol.* **139**:2423–2429.
35. Vaneechoutte, M., H. De Beenhouwer, G. Claeys, G. Verschraegen, A. De Rouck, N. Paepe, A. Elaichouni, and F. Portaels. 1993. Identification of *Mycobacterium* species by using amplified ribosomal DNA restriction analysis. *J. Clin. Microbiol.* **31**:2061–2065.
36. Wallace, R. J., Jr., G. Bedsole, G. Sumter, C. V. Sanders, L. C. Steele, B. A. Brown, J. Smith, and D. R. Graham. 1990. Activities of ciprofloxacin and ofloxacin against rapidly growing mycobacteria with demonstration of acquired resistance following single-drug therapy. *Antimicrob. Agents Chemother.* **34**:65–70.
37. Wallace, R. J., Jr., B. A. Brown, and G. O. Onyi. 1991. Susceptibilities of *Mycobacterium fortuitum* biovar. *fortuitum* and the two subgroups of *Mycobacterium chelonae* to imipenem, cefmetazole, cefaxitin, and amoxicillin-clavulanic acid. *Antimicrob. Agents Chemother.* **35**:773–775.
38. Wallace, R. J., Jr., B. A. Brown, and G. O. Onyi. 1992. Skin, soft tissue, and bone infections due to *Mycobacterium chelonae chelonae*: importance of prior corticosteroid therapy, frequency of disseminated infections, and resistance to oral antimicrobials other than clarithromycin. *J. Infect. Dis.* **166**:405–412.
39. Wallace, R. J., Jr., V. A. Silcox, M. Tsukamura, B. A. Brown, J. O. Kilburn, W. R. Butler, and G. Onyi. 1993. Clinical significance, biochemical features, and susceptibility patterns of sporadic isolates of the *Mycobacterium chelonae*-like organism. *J. Clin. Microbiol.* **31**:3231–3239.
40. Waterman, M. S. 1995. Introduction to computational biology, p. 201–202. Chapman & Hall, London, England.