JOURNAL OF CLINICAL MICROBIOLOGY, May 2003, p. 1996–2001 Vol. 41, No. 5
0095-1137/03/$08.00+0 DOI: 10.1128/JCM.41.5.1996–2001.2003

Usefulness of the MicroSeq 500 16S Ribosomal DNA-Based Bacterial Identification System for Identification of Clinically Significant Bacterial Isolates with Ambiguous Biochemical Profiles

Patrick C. Y. Woo,[1] Kenneth H. L. Ng,[2] Susanna K. P. Lau,[1] Kam-tong Yip,[2] Ami M. Y. Fung,[1] Kit-wah Leung,[1] Dorothy M. W. Tam,[1] Tak-lun Que,[2] and Kwok-yung Yuen[1,3]*

Department of Microbiology, The University of Hong Kong, Queen Mary Hospital,[1] Department of Microbiology, Tuen Mun Hospital,[2] and HKU-Pasteur Research Centre,[3] Hong Kong

Received 10 December 2002/Returned for modification 28 January 2003/Accepted 14 February 2003

Due to the inadequate automation in the amplification and sequencing procedures, the use of 16S rRNA gene sequence-based methods in clinical microbiology laboratories is largely limited to identification of strains that are difficult to identify by phenotypic methods. In this study, using conventional full-sequence 16S rRNA gene sequencing as the "gold standard," we evaluated the usefulness of the MicroSeq 500 16S ribosomal DNA (rDNA)-based bacterial identification system, which involves amplification and sequencing of the first 527-bp fragment of the 16S rRNA genes of bacterial strains and analysis of the sequences using the database of the system, for identification of clinically significant bacterial isolates with ambiguous biochemical profiles. Among 37 clinically significant bacterial strains that showed ambiguous biochemical profiles, representing 37 nonduplicating aerobic gram-positive and gram-negative, anaerobic, and *Mycobacterium* species, the MicroSeq 500 16S rDNA-based bacterial identification system was successful in identifying 30 (81.1%) of them. Five (13.5%) isolates were misidentified at the genus level (*Granulicatella adiacens* was misidentified as *Abiotrophia defectiva*, *Helcococcus kunzii* was misidentified as *Clostridium hastiforme*, *Olsenella uli* was misidentified as *Atopobium rimae*, *Leptotrichia buccalis* was misidentified as *Fusobacterium mortiferum*, and *Bergeyella zoohelcum* was misidentified as *Rimerella anatipestifer*), and two (5.4%) were misidentified at the species level (*Actinomyces odontolyticus* was misidentified as *Actinomyces meyeri* and *Arcobacter cryaerophilus* was misidentified as *Arcobacter butzleri*). When the same 527-bp DNA sequences of these seven isolates were compared to the known 16S rRNA gene sequences in the GenBank, five yielded the correct identity, with good discrimination between the best and second best match sequences, meaning that the reason for misidentification in these five isolates was due to a lack of the 16S rRNA gene sequences of these bacteria in the database of the MicroSeq 500 16S rDNA-based bacterial identification system. In conclusion, the MicroSeq 500 16S rDNA-based bacterial identification system is useful for identification of most clinically important bacterial strains with ambiguous biochemical profiles, but the database of the MicroSeq 500 16S rDNA-based bacterial identification system has to be expanded in order to encompass the rarely encountered bacterial species and achieve better accuracy in bacterial identification.

Identification of bacteria in clinical microbiology laboratories is traditionally performed by isolation of the organisms and study of their phenotypic characteristics, including Gram staining, morphology, culture requirements, and biochemical reactions. However, these methods of bacterial identification have major drawbacks. First, they cannot be used for noncultivable organisms. Second, we are occasionally faced with organisms exhibiting biochemical characteristics that do not fit into patterns of any known genus and species. Third, identification of slow-growing organisms would be extremely slow and difficult.

Since the discovery of PCR and DNA sequencing, comparison of the gene sequences of bacterial species showed that the 16S rRNA gene is highly conserved within a species and among species of the same genus, and hence can be used as the new "gold standard" for identification of bacteria to the species level. Using this new standard, phylogenetic trees, based on base differences between species, are constructed, and bacteria are classified and reclassified into new genera (8). Recently we have reported the use of this technique for the identification to species level of bacterial strains that have posed problems of identification in our clinical microbiology laboratory, as well as the clinical impact of accurate identification of such isolates (1, 3, 4, 5, 6, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25–29, 31; P. C. Y. Woo, J. H. C. Li, W. M. Tang, and K. Y. Yuen, Letter, N. Engl. J. Med. **345:**842-843, 2001).

The MicroSeq 500 16S ribosomal DNA (rDNA)-based bacterial identification system (Perkin-Elmer Applied Biosystems Division, Foster City, Calif.) has been designed for rapid and accurate identification of bacterial pathogens. In this system, the first 527-bp fragment of the 16S rRNA gene of the bacterial strain is amplified, sequenced, and analyzed using the database of the system. It has been shown that the system is useful for the identification of aerobic pathogenic gram-negative bacilli, *Mycobacterium* species, and coryneform bacteria (9, 11, 12). However, due to the inadequate automation in the amplification and sequencing procedures, it is still very labor-intensive and not cost-effective to use this system for routine identification of all bacterial isolates in clinical microbiology laborato-

* Corresponding author. Mailing address: Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong. Phone: (852) 28554892. Fax: (852) 28551241. E-mail: hkumicro@hkucc.hku.hk.

TABLE 1. PCR primers used for conventional 16S rRNA gene sequencing

Strain no.	Primer	
	Forward	Backward
1	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
2	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
3	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
4	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
5	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
6	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
7	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
8	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
9	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
10	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
11	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
12	LPW200 5′-GAGTTGCGAACGGGTGAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
13	LPW398 5′-GGCGTGCTTACCACATG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
14	LPW398 5′-GGCGTGCTTACCACATG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
15	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
16	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW325 5′-CGGATACCTTGTTACGACT-3′
17	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW282 5′-GCTTCGGGTGTYRCCAACTTTC-3′
18	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW205 5′-CTTGTTACGACTTCACCC-3′
19	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW325 5′-CGGATACCTTGTTACGACT-3′
20	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW325 5′-CGGATACCTTGTTACGACT-3′
21	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW325 5′-CGGATACCTTGTTACGACT-3′
22	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW325 5′-CGGATACCTTGTTACGACT-3′
23	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW324 5′-TTGTTACGACTTCACCCCA-3′
24	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
25	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
26	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
27	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW307 5′-TAGCGATTCCGACTTCAT-3′
28	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
29	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW182 5′-AGTCGCTGATTCCACTGTGG-3′
30	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW182 5′-AGTCGCTGATTCCACTGTGG-3′
31	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW182 5′-AGTCGCTGATTCCACTGTGG-3′
32	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
33	LPW55 5′-AGTTTGATCCTGGCTCAG-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
34	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
35	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
36	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′
37	LPW81 5′-TGGCGAACGGGTGAGTAA-3′	LPW58 5′-AGGCCCGGGAACGTATTCAC-3′

ries. At the moment, the use of this system or other 16S rRNA gene sequence-based identification methods for bacterial identification in clinical microbiology laboratories is largely limited to the identification of strains that are difficult to identify by phenotypic methods. In this study, using DNA sequencing of the complete 16S rRNA gene as the gold standard, we evaluated the usefulness of this system in the identification of 37 clinically significant bacterial strains that showed ambiguous biochemical profiles. These strains represented 37 nonduplicating aerobic gram-positive and gram-negative, anaerobic, and *Mycobacterium* species. The potential for 16S rRNA gene sequencing for general use in clinical microbiology laboratories is also discussed.

MATERIALS AND METHODS

Bacterial strains. The bacterial strains used in this study were isolated from patient specimens and obtained from the Clinical Microbiology Laboratory of Queen Mary Hospital in Hong Kong (1995 to 2001). Based on the Gram smear appearances, growth requirements, colonial morphologies, and the results of other simple phenotypic tests, such as motility, catalase, and cytochrome oxidase, appropriate strips or cards of the API system (bioMerieux Vitek, Hazelwood, Mo.) and Vitek system (bioMerieux Vitek) and/or additional conventional biochemical methods were used for identification of the bacterial strains (7). An ambiguous biochemical profile is defined as disagreement between the results provided by the API and Vitek systems or a biochemical profile that did not fit the typical profiles of known bacterial species (7). All bacterial strains that were clinically significant but showed ambiguous biochemical profiles were subject to conventional 16S rRNA gene sequencing. After excluding novel bacterial species, 37 strains, representing 37 nonduplicating aerobic gram-positive and gram-negative, anaerobic, and *Mycobacterium* species, were selected for DNA sequencing of the first 527-bp fragment of the 16S rRNA gene and analysis by the MicroSeq 16S rDNA-based bacterial identification system. Among the 37 strains, 24 (64.9%) were isolated from blood, four (10.8%) were isolated from stool, three (8.1%) were isolated from pus, two (5.4%) were isolated from biopsy specimens, one (2.7%) was isolated from bile, one (2.7%) was isolated from bronchoalveolar lavage, one (2.7%) was isolated from an intrauterine contraceptive device, and one (2.7%) was isolated from a cochlear implant.

Extraction of bacterial DNA. Bacterial DNA extraction was modified from our previous published protocol (21). Briefly, 80 μl of NaOH (0.05 M) was added to 20 μl of bacterial cells suspended in distilled water, and the mixture was incubated at 60°C for 45 min, followed by addition of 6 μl of Tris-HCl (pH 7.0), achieving a final pH of 8.0. The resultant mixture was diluted 100-fold, and 5 μl of the diluted extract was used for PCR.

PCR, gel electrophoresis, and conventional 16S rRNA gene sequencing. PCR amplification and DNA sequencing of the full 16S rRNA genes were performed according to our previous publications (1, 3, 4, 5, 6, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25–29, 31; Woo et al., letter). Briefly, DNase I-treated distilled water and PCR master mix (which contains deoxynucleoside triphosphates [NTPs], PCR buffer, and *Taq* polymerase) were used in all PCRs by adding 1 U of DNase I (Pharmacia, Sweden) to 40 μl of distilled water or PCR master mix, incubating the mixture at 25°C for 15 min, and subsequently at 95°C for 10 min to inactivate

TABLE 2. Identification of clinical bacterial isolates with ambiguous biochemical profiles by conventional 16S rRNA gene sequencing, commercially available bacterial identification systems, and the Microseq 500 16S rDNA-based bacterial identification system

Strain no. (reference)	Sex[a]/age (yr) of patient	Clinical specimen or specimen source	Identification by conventional 16S rRNA gene sequencing	Identification by commercially available bacterial identification system				Identification by MicroSeq 500 16S rDNA-based bacterial identification system	
				Vitek		API		Identity	Percentage difference between sequence of isolates and that of identified bacterium in database
				Card used for identification	Identity[c]	Strips for identification	Identity		
1 (20)	F/43	Blood	*Staphylococcus aureus*	GPI	95% *Staphylococcus simulans*, 3% *Staphylococcus haemolyticus*	STAPH	87% *Staphylococcus aureus*, 5% *Staphylococcus warneri*	*Staphylococcus aureus*	0
2	M/3	Cochlear implant	*Staphylococcus epidermidis*	GPI	53% *Staphylococcus epidermidis*, 27% *Staphylococcus capitis*	STAPH	99% *Staphylococcus chromogenes*	*Staphylococcus epidermidis*	0
3	F/3	Blood	*Micrococcus luteus*	GPI	99.9% *Micrococcus* sp.	STAPH	99% *Staphylococcus auricularis*	*Micrococcus luteus*	0.2
4 (16)	F/69	Blood	*Streptococcus dysgalactiae*	GPI	Unidentified	20 STREP	77% *Streptococcus dysgalactiae*	*Streptococcus dysgalactiae*	0.6
5 (6)	M/81	Blood	*Streptococcus iniae*	GPI	Unidentified	20 STREP	99% *Streptococcus dysgalactiae*	*Streptococcus iniae*	0
6	M/33	Bronchoalveolar lavage	*Streptococcus salivarius*	GPI	Unidentified	20 STREP	Unidentified	*Streptococcus salivarius*	0.4
7	F/31	Blood	*Streptococcus anginosus*	GPI	Unidentified	20 STREP	70% *Streptococcus intermedius*, 30% *Streptococcus anginosus*	*Streptococcus anginosus*	0
8 (19)	M/62	Blood	*Granulicatella adiacens*	GPI	98% *Gemella morbillorum* or *Streptococcus agalactiae* or *Streptococcus acidominimus*	20 STREP	80% *Granulicatella adiacens*, 16% *Gemella morbillorum*	*Abiotrophia defectiva*	9.9
9 (19)	M/15	Blood	*Abiotrophia defectiva*	GPI	Unidentified	20 STREP	Unidentified	*Abiotrophia defectiva*	1.1
10	M/66	Blood	*Gemella morbillorum*	GPI	63% *Gemella morbillorum* or *Streptococcus agalactiae* or *Streptococcus acidominimus*, 22% *Streptococcus pneumoniae*	20 STREP	60% *Leuconostoc* sp., 28% *Streptococcus mitis*	*Gemella morbillorum*	0.1
11	F/41	Blood	*Gemella haemolysans*	GPI	68% *Gemella morbillorum* or *Streptococcus agalactiae* or *Streptococcus acidominimus*, 24% *Streptococcus pneumoniae*	20 STREP	95% *Gemella haemolysans*, 6% *Gemella morbillorum*	*Gemella haemolysans*	0.6
12	M/41	Blood	*Helcococcus kunzii*	GPI	Unidentified	20 STREP	50% *Gemella haemolysans*, 30% *Gemella morbillorum*, 18% *Abiotrophia adiacens*	*Clostridium hastiforme*	13.9
13 (4)	F/39	Blood	*Microbacterium* sp.	GPI	Unidentified	CORYNE	99% *Aureobacterium* or *Corynebacterium aquaticum*	*Microbacterium liquefaciens*	0.8
14	M/22	Blood	*Gordona terrae*	GPI	Unidentified	CORYNE	98% *Rhodococcus* sp.	*Gordona terrae*	0
15	M/38	Pus swab	*Tsukamurella* sp.	GPI	Unidentified	CORYNE	83% *Rhodococcus* sp., 12% *Aureobacterium* or *Corynebacterium aquaticum*	*Tsukamurella pulmonis*	0.2
16	M/71	Blood	*Clostridium ramosum*	ANI	86% *Clostridium barati*, 7% *Lactobacillus catenaforme*, 2% *Clostridium paraputrificum*	20A	97% *Bifidobacterium* sp.	*Clostridium ramosum*	0
17 (18)	F/36	Intrauterine contraceptive device	*Actinomyces odontolyticus*	ANI	99% *Propionibacterium granulosum*	20A	78% *Actinomyces meyeri* or *odontolyticus*, 19% *Actinomyces naeslundii*	*Actinomyces meyeri*	2.1
18 (15)	M/70	Blood	*Lactobacillus salivarius*	ANI	Unidentified	20A	70% *Actinomyces naeslundii*, 30% *Bifidobacterium* sp.	*Lactobacillus salivarius*	
19	M/43	Blood	*Olsenella uli*	ANI	81% *Propionibacterium granulosum*, 11% *Corynebacterium pseudotuberculosis*	20A	43% *Gemella morbillorum*, 37% *Lactobacillus fermentum*, 10% *Proionibacterium propionicum or avidum*	*Atopobium rimae*	11.1
20	F/87	Blood	*Eggerthella lenta*	ANI	83% *Corynebacterium jeikeium*, 7% *Clostridium hastiforme*, 7% *Clostridium histolyticum*	20A	92% *Eggerthella lenta*, 8% *Actinomyces viscosus*	*Eggerthella lenta*	0
21	M/69	Blood	*Eubacterium tenue*	ANI	98% *Clostridium difficile*	20A	76% *Clostridium bifermentens*, 18% *Clostridium cadaveris*	*Eubacterium tenue*	0
22	F/70	Blood	*Leptotrichia buccalis*	ANI	99% *Propionibacterium granulosum*	20A	98% *Bifidobacterium* sp.	*Fusobacterium mortiferum*	14.6
23	M/36	Blood	*Neisseria elongata*	NHI	91% *Kingella denitrificans*	NH	99% *Neisseria gonorrhoeae*	*Neisseria elongata*	0

24 (28)	M/31	Stool	*Escherichia coli*	GNI+	73% *Salmonella arizonae*, 17% *Salmonella* spp.	20E	76% *Escherichia coli*, 23% *Salmonella arizonae*	*Escherichia coli*	0.1
25	M/48	Stool	*Kluyvera ascorbata*	GNI+	57% *Kluyvera* sp., 36% *Enterobacter intermedius*	20E	57% *Enterobacter amnigenus*, 24% *Enterobacter intermedius*, 19% *Kluyvera* sp.	*Kluyvera ascorbata*	0.2
26	M/60	Bile	*Shewanella alga*	GNI+	76% nonfermenting gram-negative bacillus, 11% *Comamonas acidovorans*	20NE	Unidentified	*Shewanella alga*	0.4
27	M/4	Pus	*Bergeyella zoohelcum*	GNI+	53% nonfermenting gram-negative bacillus (asaccharolytic), 39% *Myroides* sp.	20NE	51.2% *Psychrobacter phenylpyruvicus*, 36.5% *Bergeyella zoohelcum*, 8.5% *Methylobacterium mesophilicum*	*Riemerella anatipestifer*	7.4
28 (3)	M/32	Blood	*Haemophilus segnis*	NHI	56% *Actinobacillus actinomycetemcomitans*	NH	58% *Haemophilus aphrophilus* or *Haemophilus paraphrophilus*	*Haemophilus segnis*	0.2
29 (26)	M/66	Blood	*Campylobacter fetus*	ND[b]	ND	ND	ND	*Campylobacter fetus*	0
30	M/9m	Stool	*Campylobacter jejuni*	ND	ND	ND	ND	*Campylobacter jejuni* or *Campylobacter coli*	0
31	F/41	Stool	*Campylobacter hyointestinalis*	ND	ND	ND	ND	*Campylobacter hyointestinalis*	0
32 (25)	M/7	Blood	*Arcobacter cryaerophilus*	ND	ND	ND	ND	*Arcobacter butzleri*	2.6
33 (5)	F/69	Blood	*Arcobacter butzleri*	ND	ND	ND	ND	*Arcobacter butzleri*	0
34 (27)	F/79	Tissue biopsy	*Mycobacterium chelonae*	ND	ND	ND	ND	*Mycobacterium chelonae* or *Mycobacterium abscessus*	0
35 (24)	F/9	Blood	*Mycobacterium neoaurum*	ND	ND	ND	ND	*Mycobacterium neoaurum*	0
36 (27)	M/62	Tissue biopsy	*Mycobacterium nonchromogenicum*	ND	ND	ND	ND	*Mycobacterium nonchromogenicum*	1.6
37	M/77	Pus	*Mycoplasma hominis*	ND	ND	ND	ND	*Mycoplasma hominis*	0.1

[a] M, male; F, female.
[b] ND, not done.
[c] Percent identity and organism name are given. In cases where identity with more than one organism was found, both or all are listed.

the DNase I. The bacterial DNA extracts and control were amplified with 0.5 μM primers (Table 1) (Gibco BRL, Rockville, Md.). The PCR mixture (50 μl) contained bacterial DNA, PCR buffer (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 2 mM $MgCl_2$, 0.01% gelatin), a 200 μM concentration of each dNTP, and 1.0 U of *Taq* polymerase (Boehringer, Mannheim, Germany). The mixtures were amplified in 40 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min, with a final extension at 72°C for 10 min, in an automated thermal cycler (Perkin-Elmer Cetus, Gouda, The Netherlands). DNase I-treated distilled water was used as the negative control. 10 μl of each amplified product was electrophoresed in 1.0% (wt/vol) agarose gel, with a molecular size marker (Lambda DNA *Ava*II digest; Boehringer) in parallel. Electrophoresis in Tris-borate-EDTA buffer was performed at 100 V for 1.5 h. The gel was stained with ethidium bromide (0.5 μg/ml) for 15 min, rinsed, and photographed under UV light illumination.

The PCR products were gel purified using the QIAquick PCR purification kit (QIAgen, Hilden, Germany). Both strands of the PCR products were sequenced twice with an ABI 377 automated sequencer according to manufacturers' instructions (Perkin-Elmer Applied Biosystems Division), using the PCR primers and additional primers designed from the first round of sequencing results. The sequences of the PCR products were compared with known 16S rRNA gene sequences in the GenBank by multiple sequence alignment using the CLUSTAL W program (14).

PCR amplification and DNA sequencing of the first 527-bp fragment of the 16S rRNA gene and analysis by the MicroSeq 500 16S rDNA-based bacterial identification system. Bacterial DNA extracts were amplified with 0.5 μM primers (005F and 531R). The PCR mixture (50 μl) contained bacterial DNA, PCR buffer (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 3 mM $MgCl_2$, 0.01% gelatin), a 200 μM concentration of each dNTP, and 1.0 U of *Taq* polymerase (Boehringer Mannheim, Germany). The mixtures were amplified in 30 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 45 s, with a final extension at 72°C for 10 min, in an automated thermal cycler (Perkin-Elmer Cetus). The amplified products were purified and sequenced as described above. The DNA sequences were analyzed using the database provided by the MicroSeq 500 16S rDNA-based bacterial identification system.

RESULTS

Conventional 16S rRNA gene sequencing. PCR of the 16S rRNA genes of the 37 isolates with ambiguous biochemical profiles showed bands at about 1,400 to 1,500 bp. For all 37 isolates, there was <1% difference between the 16S rRNA gene sequences of the isolates and the most closely matched sequence in the GenBank.

DNA sequencing of the first 527-bp fragment of the 16S rRNA gene and analysis by the MicroSeq 500 16S rDNA-based bacterial identification system. PCR amplification of the first 527-bp fragments of the 16S rRNA genes of the 37 isolates showed bands at about 500 bp. Analysis of the 37 sequences using the MicroSeq 500 16S rDNA-based bacterial identification database showed that the identities of 30 (81.1%) strains were the same as those obtained by conventional 16S rRNA gene sequencing (Table 2). For the remaining seven (18.9%) sequences, five (13.5%) isolates were misidentified at the genus level (case 8, *Granulicatella adiacens* misidentified as *Abiotrophia defectiva*; case 12, *Helcococcus kunzii* misidentified as *Clostridium hastiforme*; case 19, *Olsenella uli* misidentified as *Atopobium rimae*; case 22, *Leptotrichia buccalis* misidentified as *Fusobacterium mortiferum*; and case 27, *Bergeyella zoohelcum* misidentified as *Rimerella anatipestifer*), whereas two (5.4%) were misidentified at the species level (case 17, *Actinomyces odontolyticus* misidentified as *Actinomyces meyeri*; case 32, *Arcobacter cryaerophilus* misidentified as *Arcobacter butzleri*).

Identification by commercially available bacterial identification systems. Phenotypic identification using API and Vitek systems were performed in 28 of the 37 isolates. Using full 16S rRNA gene sequencing as the gold standard, the API system correctly identified seven (25%) of the 28 isolates at >70%

TABLE 3. Analysis of DNA sequences of strains identified incorrectly using database of Microseq 500 16S rDNA bacterial identification system

Strain no.	Identification by conventional 16S rRNA gene sequencing	Identification by DNA sequencing of first 527-bp fragment of 16S rRNA gene				
		Analysis by using database of MicroSeq 500 16S rDNA bacterial identification system	Analysis by using database of GenBank			
			Best match	No. of base differences (%) between strain and best match	Second best match	No. of base differences (%) between strain and second best match
8	*Granulicatella adiacens*	*Abiotrophia defectiva*	*Granulicatella adiacens*	0 (0)	*Abiotrophia paraadiacens*	1 (0.2)
12	*Helcococcus kunzii*	*Clostridium hastiforme*	*Helcococcus kunzii*	10 (1.9)	*Sedimentibacter hydroxybenzoicus*	91 (17.3)
17	*Actinomyces odontolyticus*	*Actinomyces meyeri*	*Actinomyces odontolyticus*	0 (0)	*Actinomyces meyeri*	4 (0.8)
19	*Olsenella uli*	*Atopobium rimae*	*Olsenella uli*	0 (0)	*Olsenella profusa*	21 (4.0)
22	*Leptotrichia buccalis*	*Fusobacterium mortiferum*	*Leptotrichia buccalis*	2 (0.4)	*Streptobacillus moniliformis*	73 (13.6)
27	*Bergeyella zoohelcum*	*Riemerella anatipestifer*	*Bergeyella zoohelcum*	2 (0.4)	*Riemerella anatipestifer*	22 (4.2)
32	*Arcobacter cryaerophilus*	*Arcobacter butzleri*	*Arcobacter cryaerophilus*	0 (0)	*Arcobacter butzleri*	12 (2.3)

confidence, whereas the Vitek system only correctly identified one (3.6%) of the 28 isolates at >70% confidence (Table 2).

DISCUSSION

In this study, we showed that the MicroSeq 500 16S rDNA-based bacterial identification system is useful for identification of most clinically important bacterial strains with ambiguous biochemical profiles, and hence would be a useful substitution for conventional full-sequence 16S rRNA gene sequencing in identification of bacterial strains that pose problems in clinical microbiology laboratories. Using conventional 16S rRNA gene sequencing as the gold standard, the MicroSeq 500 16S rRNA bacterial identification system is able to identify 32 (86.5%) of the 37 (including 15 aerobic or facultative anaerobic gram-positive; 11 aerobic, microaerophilic, facultative anaerobic gram-negative; seven anaerobic; three mycobacterial; and one *Mycoplasma*) isolates with ambiguous biochemical profiles to the genus level, and is able to identify 30 (81.1%) of these 37 isolates to the species level.

The most important reason for failure of the MicroSeq 500 16S rDNA-based bacterial identification system in identifying a bacterium is a lack of the 16S rRNA gene sequence of the particular bacterium in the database. PCR amplification of all 37 isolates using 005F and 531R as PCR primers were successful, yielding specific bands at about 500 bp. Furthermore, DNA sequencing of the corresponding PCR products using the same oligonucleotides as sequencing primers posed no problems. When the sequences of the 527 bp were aligned to the database of the MicroSeq 500 16S rDNA-based bacterial identification system, seven (18.9%) of the isolates did not yield the correct identity. The 16S rRNA gene sequences of all the seven isolates were not included in the MicroSeq 500 16S rDNA-based bacterial identification system database, probably because they were expected to be rarely encountered. On the other hand, when the same 527-bp DNA sequences of these seven isolates were compared to the known 16S rRNA gene sequences in the GenBank, five yielded the correct identity, with good discrimination between the best and second best match sequences For the remaining two strains, only full 16S rRNA gene sequencing correctly identified them with good discrimination, indicating that the first 527-bp fragments of the 16S rRNA genes of these species were not discriminative enough. This discrepancy between using the GenBank database and the MicroSeq 500 16S rDNA-based bacterial identification system database suggests that the database of the latter has to be expanded in order to encompass the rarely encountered bacterial species and achieve better accuracy in identification of bacteria with ambiguous biochemical profiles. If this limitation of the MicroSeq 500 16S rDNA-based bacterial identification system database is overcome, it would be a better choice than full 16S rRNA gene sequencing in clinical microbiology laboratories, as it involves amplification and sequencing of only about 500 bp. Therefore it would be less time consuming and expensive than full 16S rRNA gene sequencing.

16S rRNA gene sequencing will continue to be the working gold standard for the identification of most bacteria, and better automation of such a technique may put it into routine use in large clinical microbiology laboratories, especially those serving tertiary centers, replacing the traditional phenotypic tests. Compared to phenotypic tests, 16S rRNA gene sequence-based identification schemes are superior in the identification of strains considered unidentifiable due to atypical biochemical profiles, slow-growing bacteria, rarely encountered bacterial species, and noncultivable strains. Furthermore, such a technique will be applicable to not only pyogenic bacteria but also other organisms such as mycobacteria (24, 27; Woo et al., letter), of which the identification is not routinely performed in most clinical microbiology laboratories because special expertise and equipment such as gas liquid chromatography are required. Modern technologies have made it possible to construct a high density of oligonucleotide arrays on a chip with oligonucleotides representing the 16S rRNA gene sequence of various bacteria. Such a design will facilitate automation of the annealing and detection of the PCR products of 16S rRNA gene amplification and avoid the step of sequencing the amplified PCR products. Hence, the turnaround time can be even shorter. Since amplification of the 16S rRNA gene takes only 4 to 6 h, and the annealing and detection of PCR product takes only another few hours, theoretically the identification can be completed within 1 day. However, at the moment, due to the inadequate automation of the DNA amplification and sequencing steps, it would not be cost-effective to use the MicroSeq 500 16S rDNA-based bacterial identification system for identification of all bacterial isolates in clinical microbiology

laboratories. On the contrary, the use of this system for identification of clinically important bacteria with ambiguous biochemical profiles would be more cost-effective and the accuracy can be easily improved with expansion and regular updating of the database.

Despite the usefulness of 16S rRNA gene sequence-based identification in most circumstances, there are still problems in some situations that remain to be solved. These include sharing of similar 16S rRNA gene sequences among different bacterial species and too much variation of the 16S rRNA gene sequences among different strains within the same species. When two or more bacterial species such as *Streptococcus pneumoniae*, *Streptococcus oralis*, and *Streptococcus mitis*; *Burkholderia pseudomallei* and *Burkholderia thailandensis*; and some rapidly growing *Mycobacterium* species share similar 16S rRNA gene sequences, 16S rRNA gene sequence-based identification systems would be unable to differentiate the species. Additional sequencing systems based on other conserved gene sequences, such as *groEL* gene sequencing, has to be employed for the differentiation of these species (2, 10, 30). As for the problem of too much variation of the 16S rRNA gene sequences among different strains within the same species, such as in *Enterobacter*, *Pantoea*, and *Leclercia*, reclassification of these groups of bacteria may be necessary to achieve better identification using gene sequence-based bacterial identification systems. However, despite the impossibility to accurately assign a particular clinical isolate to a specific species, assigning the clinical isolate to a certain group can successfully assist the clinical management of the corresponding patient (23).

ACKNOWLEDGMENTS

This work is partly supported by the University Development Fund, University Research Grant Council, and the Committee for Research and Conference Grant, The University of Hong Kong.

REFERENCES

1. **Cheuk, W., P. C. Y. Woo, K. Y. Yuen, P. H. Yu, and J. K. C. Chan.** 2001. Intestinal inflammatory pseudotumor with regional lymph node involvement: Identification of a new bacterium as the etiologic agent. J. Pathol. **192:**289–292.
2. **Goh, S. H., S. Potter, J. O. Wood, S. M. Hemmingsen, R. P. Reynolds, and A. W. Chow.** 1996. HSP60 gene sequences as universal targets for microbial species identification: studies with coagulase-negative staphylococci. J. Clin. Microbiol. **34:**818–823.
3. **Lau, S. K. P., P. C. Y. Woo, B. Y. L. Chan, A. M. Y. Fung, T. L. Que, and K. Y. Yuen.** 2002. *Haemophilus segnis* polymicrobial and monomicrobial bacteremia identified by 16S ribosomal RNA gene sequencing. J. Med. Microbiol. **51:**635–640.
4. **Lau, S. K. P., P. C. Y. Woo, G. K. S. Woo, and K. Y. Yuen.** 2002. Catheter-related *Microbacterium* bacteremia identified by 16S ribosomal RNA gene sequencing. J. Clin. Microbiol. **40:**2681–2685.
5. **Lau, S. K. P., P. C. Y. Woo, J. L. L. Teng, K. W. Leung, and K. Y. Yuen.** 2002. Identification by 16S ribosomal RNA gene sequencing of *Arcobacter butzleri* bacteraemia in a patient with acute gangrenous appendicitis. Mol. Pathol. **55:**182–185.
6. **Lau, S. K. P., P. C. Y. Woo, H. Tse, K. W. Leung, and K. Y. Yuen.** 2003. Invasive *Streptococcus iniae* infections outside North America. J. Clin. Microbiol. **41:**1004–1009.
7. **Murray, P. R., E. J. Baro, M. A. Pfaller, F. C. Tenover, and. R. H. Yolken (ed.).** 1999. Manual of clinical microbiology. 7th ed. American Society for Microbiology, Washington, D.C.
8. **Olsen, G. J., and C. R. Woese.** 1993. Ribosomal RNA: a key to phylogeny. FASEB J. **7:**113–123.
9. **Patel, J. B., D. G. Leonard, X. Pan, J. M. Musser, R. E. Berman, and I. Nachamkin.** 2000. Sequence-based identification of *Mycobacterium* species using the Microseq 500 16S rDNA bacterial identification system. J. Clin. Microbiol. **38:**246–251.
10. **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.
11. **Tang, Y. W., A. Von Graevenitz, M. G. Waddington, M. K. Hopkins, D. H. Smith, H. Li, C. P. Kolbert, S. O. Montgomery, and D. H. Persing.** 2000. Identification of coryneform bacterial isolates by ribosomal DNA sequence analysis. J. Clin. Microbiol. **38:**1676–1678.
12. **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 technique for identification of unusual aerobic pathogenic gram-negative bacilli. J. Clin. Microbiol. **36:**3674–3679.
13. **Teng, J. L. L., P. C. Y. Woo, K. W. Leung, S. K. P. Lau, M. K. M. Wong, and K. Y. Yuen.** 2003. Pseudobacteraemia in a patient with neutropenic fever caused by a novel *Paenibacillus* species: *Paenibacillus hongkongensis* sp. nov. Mol. Pathol. **56:**29–35.
14. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
15. **Woo, P. C. Y., A. M. Y. Fung, S. K. P. Lau, and K. Y. Yuen.** 2002. Identification by 16S ribosomal RNA gene sequencing of *Lactobacillus salivarius* bacteremic cholecystitis. J. Clin. Microbiol. **40:**265–267.
16. **Woo, P. C. Y., A. M. Y. Fung, S. K. P. Lau, S. S. Y. Wong, and K. Y. Yuen.** 2001. Group G beta-hemolytic streptococcal bacteremia characterized by 16S ribosomal RNA gene sequencing. J. Clin. Microbiol. **39:**3147–3155.
17. **Woo, P. C. Y., A. M. Y. Fung, S. S. Y. Wong, H. W. Tsoi, and K. Y. Yuen.** 2001. Isolation and characterization of a *Salmonella enterica* serotype Typhi variant and its clinical and public health implications. J. Clin. Microbiol. **39:**1190–1194.
18. **Woo, P. C. Y., A. M. Y. Fung, S. K. P. Lau, E. Hon, and K. Y. Yuen.** 2002. Diagnosis of pelvic actinomycosis by 16S ribosomal RNA gene sequencing and its clinical significance. Diagn. Microbiol. Infect. Dis. **43:**113–118.
19. **Woo, P. C. Y., A. M. Y. Fung, S. K. P. Lau, B. Y. L. Chan, S. K. Chiu, T. L. Que, R. W. H. Yung, and K. Y. Yuen.** 2003. *Granulicatella adiacens* and *Abiotrophia defectiva* bacteraemia characterised by 16S ribosomal RNA gene sequencing. J. Med. Microbiol. **52:**137–140.
20. **Woo, P. C. Y., A. S. P. Leung, K. W. Leung, and K. Y. Yuen.** 2001. Identification of slide-coagulase positive, tube-coagulase negative *Staphylococcus aureus* by 16S ribosomal RNA gene sequencing. Mol. Pathol. **54:**244–247.
21. **Woo, P. C. Y., C. Y. Lo, S. K. Lo, H. Siau, J. S. M. Peiris, S. S. Y. Wong, W. K. Luk, T. M. Chan, W. W. Lim, and K. Y. Yuen.** 1997. Distinct genotypic distributions of cytomegalovirus (CMV) envelope glycoprotein in bone marrow and renal transplant recipients with CMV disease. Clin. Diagn. Lab. Immunol. **4:**515–518.
22. **Woo, P. C. Y., D. M. W. Tam, K. W. Leung, S. K. P. Lau, J. L. L. Teng, M. K. M. Wong, and K. Y. Yuen.** 2002. *Streptococcus sinensis* sp. nov., a novel *Streptococcus* species isolated from a patient with infective endocarditis. J. Clin. Microbiol. **40:**805–810.
23. **Woo, P. C. Y., E. Y. L. Cheung, K. W. Leung, and K. Y. Yuen.** 2001. Identification by 16S ribosomal RNA gene sequencing of an *Enterobacteriaceae* species with ambiguous biochemical profile from a renal transplant recipient. Diagn. Microbiol. Infect. Dis. **39:**85–93.
24. **Woo, P. C. Y., H. W. Tsoi, K. W. Leung, P. N. L. Lum, A. S. P. Leung, C. H. Ma, K. M. Kam, and K. Y. Yuen.** 2000. Identification of *Mycobacterium neoaurum* isolated from a neutropenic patient with catheter-related bacteremia by 16S rRNA sequencing. J. Clin. Microbiol. **38:**3515–3517.
25. **Woo, P. C. Y., K. T. K. Chong, K. W. Leung, T. L. Que, and K. Y. Yuen.** 2001. Identification of *Arcobacter cryaerophilus* isolated from a traffic accident victim with bacteraemia by 16S ribosomal RNA gene sequencing. Diagn. Microbiol. Infect. Dis. **40:**125–127.
26. **Woo, P. C. Y., K. W. Leung, H. W. Tsoi, S. S. Y. Wong, J. L. L. Teng, and K. Y. Yuen.** 2002. Thermo-tolerant *Campylobacter fetus* bacteraemia identified by 16S ribosomal RNA gene sequencing: an emerging pathogen in immunocompromised patients. J. Med. Microbiol. **51:**740–746.
27. **Woo, P. C. Y., K. W. Leung, S. S. Y. Wong, K. T. K. Chong, E. Y. L. Cheung, and K. Y. Yuen.** 2002. Relative alcohol-resistant mycobacteria are emerging pathogens in patients receiving acupuncture treatment. J. Clin. Microbiol. **40:**1219–1224.
28. **Woo, P. C. Y., P. K. L. Leung, K. W. Leung, and K. Y. Yuen.** 2000. Identification by 16S ribosomal RNA gene sequencing of an *Enterobacteriaceae* species from a bone marrow transplant recipient. Mol. Pathol. **53:**211–215.
29. **Woo, P. C. Y., J. L. L. Teng, S. K. P. Lau, P. N. L. Lum, K. W. Leung, K. L. Wong, K. W. Li, K. C. Lam, and K. Y. Yuen.** 2003. Analysis of a viridans group strain reveals a case of bacteremia due to Lancefield group G alpha-hemolytic *Streptococcus dysgalactiae* subsp. *equisimilis* in a patient with pyomyositis and reactive arthritis. J. Clin. Microbiol. **41:**613–618.
30. **Woo, P. C. Y., G. K. S. Woo, S. K. P. Lau, S. S. Y. Wong, and K. Y. Yuen.** 2002. Single gene target bacterial identification: *groEL* gene sequencing for discriminating clinical isolates of *Burkholderia pseudomallei* and *Burkholderia thailandensis*. Diagn. Microbiol. Infect. Dis. **40:**4382–4387.
31. **Yuen, K. Y., P. C. Y. Woo, J. L. L. Teng, K. W. Leung, M. K. M. Wong, and S. K. P. Lau.** 2001. *Laribacter hongkongensis* gen. nov., sp. nov., a novel gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. J. Clin. Microbiol. **39:**4227–4232.