JOURNAL OF CLINICAL MICROBIOLOGY, Jan. 2000, p. 301–308 Vol. 38, No. 1
0095-1137/00/$04.00+0
Copyright © 2000, American Society for Microbiology. All Rights Reserved.

Differentiation of Phylogenetically Related Slowly Growing Mycobacteria by Their *gyrB* Sequences

HIROAKI KASAI,[1]* TAKAYUKI EZAKI,[2] AND SHIGEAKI HARAYAMA[1]

Marine Biotechnology Institute, Kamaishi Laboratories, 3-75-1 Heita, Kamaishi, Iwate 026-0001,[1] and Department of Microbiology, Gifu University School of Medicine, Gifu, Gifu 500-8705,[2] Japan

Received 27 May 1999/Returned for modification 18 August 1999/Accepted 4 October 1999

The conventional methods for identifying mycobacterial species are based on their phenotypic characterization. Since some problematic species are slow growers, their taxonomy takes several weeks or months to identify. The ribosomal DNA (rDNA) sequence-based identification strategy has been adopted to solve this problem. More recently, the *gyrB* sequences have been shown to be useful phylogenetic markers for the identification of species. We determined the *gyrB* sequences of 43 slowly growing strains belonging to 15 species in the genus *Mycobacterium*. The frequencies of base substitutions in the *gyrB* sequences were comparable to those in the 16S-23S rDNA internal transcribed spacer (ITS) sequences. The ITS sequences of four species belonging to the *M. tuberculosis* complex (*M. tuberculosis*, *M. bovis*, *M. africanum*, and *M. microti*) were 100% identical, while four synonymous substitutions were found in the *gyrB* sequences of these strains. Based on the differences found in the *gyrB* sequences, we developed PCR and PCR-restriction fragment length polymorphism methods to discriminate these species.

The increase in the incidence of infection by mycobacteria in humans, especially among immunocompromised patients, is a matter of serious concern to the public. In order to control this infection, rapid identification of the carriers of mycobacteria is a high priority. Intensive research efforts were made to develop rapid methods for identifying mycobacteria. Numerical taxonomic matrices and 16S ribosomal DNA (rDNA)-based phylogenetic analyses have provided a great deal of information on the systematics of mycobacteria (19, 24, 37). The 16S rDNA-based methods are currently widely accepted as rapid and accurate means for identifying mycobacteria (14, 15, 28). However, it was found that the distinctions between *Mycobacterium kansasii* and *M. gastri*, between *M. malmoense* and *M. szulgai*, between *M. marinum* and *M. ulcerans*, and between the members of the *M. tuberculosis* complex are difficult to make based on their 16S rDNA sequences since they are almost identical between these strains. To solve this problem, the sequences of the 16S-23S rDNA internal transcribed spacer (ITS) have been used to distinguish between *M. kansasii* and *M. gastri* (20). However, the ITS sequences are identical among the *M. tuberculosis* complex strains.

The *M. tuberculosis* complex consists of four closely related groups, *M. tuberculosis*, *M. bovis*, *M. africanum*, and *M. microti*, and they may be considered to be subspecies of *M. tuberculosis* (37). Their close relatedness has been demonstrated by DNA-DNA hybridization (1), isozyme analysis (5), and electrokaryotyping (5). However, their host range and pathogenicity are quite different. The host of *M. tuberculosis* and *M. africanum* is limited to humans (37), while *M. bovis* causes disease in a wide range of wild and domestic mammals, as well as in humans (18). *M. microti* has recently been reported to infect not only voles but also humans (32).

Several efforts have been made to differentiate slowly growing mycobacterial species by using protein-encoding genes such as the 32-kDa protein gene (22), *dnaJ* (26, 27, 34), *hsp 65* (17, 25), the superoxide dismutase gene (21, 44), *recA* (31), and *rpoV* (3). However, DNA sequence analyses of these genes revealed them to be invariant among the members of the *M. tuberculosis* complex.

Yamamoto and Harayama (40) have proposed that *gyrB* could be a suitable phylogenetic marker for the identification and classification of bacteria (10, 41, 42). They have shown that the divergence of *gyrB* sequences reflected the taxonomical relationships in the genera *Acinetobacter* (42) and *Pseudomonas* (43). They have also shown that the average base substitution rate of 16S rDNA was 1% per 50 million years, while that of the *gyrB* genes at synonymous sites was 0.7 to 0.8% per one million years (41). The *gyrB* analyses of other bacterial genera have also resolved closely related strains (33, 35, 36, 39).

In this report, we analyzed the *gyrB* sequences of type strains and clinical isolates of slowly growing mycobacteria, including the *M. tuberculosis* complex, *M. kansasii*, *M. gastri*, *M. avium*, *M. intracellulare*, *M. malmoense*, and *M. simiae*. In parallel, we analyzed the ITS sequences of these strains and compared the results with those of the *gyrB* analysis. Furthermore, we developed PCR-based methods to differentiate four species of the *M. tuberculosis* complex.

MATERIALS AND METHODS

Strains and cultivation. The strains used in this study are shown in Table 1. They were cultivated by using Ogawa's medium (Nissui Pharmaceutical Co.).

Determination of *gyrB* gene sequences. The protocol for determining the *gyrB* sequences was almost the same as that described by Yamamoto and Harayama except for the PCR primers (40). The primers used here are listed in Table 2, the location of each primer sequence being indicated by the numbering system for the *M. tuberculosis gyrB* sequence. A 1.2-kb *gyrB* segment was amplified by PCR by using the universal primers UP1TL and UP2rTL. To determine the nucleotide sequence of both strands of the *gyrB* segment, shorter fragments were amplified by using the PCR-amplified 1.2-kb *gyrB* segment as a template and two pairs of primers, either UP1EM13r plus QTK-21 or GGTH-r plus UP2r-21. The amplification reaction was subjected to 35 cycles with a denaturation step at 95°C for 1 min, an annealing step at 63°C for 1 min, and an extension step at 72°C for 2 min. Sequences were determined by using the universal primers for the M13 phage vector.

Determination of ITS sequences. The ITS sequences were determined by using the primers Ec16S.1390p (5′-TTGTACACACCGCCCGTC-3′) and Mb23S.44n (5′-TCTCGATGCCAAGGCATCCACC-3′) (20). The amplification reaction

* Corresponding author. Mailing address: Marine Biotechnology Institute, Kamaishi Laboratories, 3-75-1 Heita, Kamaishi, Iwate 026-0001, Japan. Phone: 81-193-26-6544. Fax: 81-193-26-6584/6592. E-mail: hiroaki.kasai@kamaishi.mbio.co.jp.

TABLE 1. Source of DNA samples and accession numbers of *gyrB* sequences

Strain	Genus and species	Remarks	Accession no.	
			gyrB	ITS
KPM T801	*M. africanum*	Type strain (ATCC 25420)	AB014192	AB026699
ATCC 25274	*M. asiaticum*		AB014206	AB026703
KPM 3012	*M. avium*	Type strain (ATCC 25291)	AB014189	AB026690
KPM T704	*M. bovis*	Type strain (ATCC 19210)	AB014184	AB026693
KPM T702	*M. bovis*	BCG Japanese strain	AB014193	
IKEDA	*M. bovis*	Cow isolate, Hokkaido	AB018554	
KPM 3504	*M. gastri*	Type strain (ATCC 15754)	AB014202	AB026697
KPM 3502	*M. gastri*	Human isolate, 1982, Shimane	AB014294	
KPM 3503	*M. gastri*	Human isolate, 1982, Kyoto	AB014295	
KPM 2201	*M. gordonae*	Type strain (ATCC 14470)	AB014191	AB026692
KPM 3101	*M. intracellulare*	Type strain (ATCC 13950)	AB014188	AB026691
KPM 1001	*M. kansasii*	Type strain (ATCC 12478)	AB014204	AB026695
KPM 1004	*M. kansasii*	Human clinical, 1990, Osaka	AB014301	
KPM 1007	*M. kansasii*	Human clinical, 1990, Osaka	AB014302	
KPM KY256	*M. kansasii*	Human clinical, 1991, Kyoto	AB014304	
KPM KY761	*M. kansasii*	Human clinical, 1991, Kyoto	AB014305	
KPM KY798	*M. kansasii*	Human clinical, 1991, Kyoto	AB014306	
KPM 1988-1	*M. kansasii*	Human clinical, 1988, Osaka	AB014307	
KPM 3401	*M. malmoense*	Type strain (ATCC 29571)	AB014187	AB026696
KPM 1201	*M. marinum*	Type strain (ATCC 927)	AB014203	AB026701
KPM T901	*M. microti*	Type strain (NCTC 8710)	AB014205	AB026700
KPM 2027	*M. scrofulaceum*	Type strain (ATCC 19981)	AB014207	AB026702
KPM 1403	*M. simiae*	Type strain (ATCC 25275)	AB014182	AB026694
KPM 2403	*M. szulgai*	Type strain (NCTC 10831)	AB014185	AB026704
KPM T021	*M. tuberculosis*	Type strain (ATCC 27294)	AB014194	AB026698
H37Rv	*M. tuberculosis*		Z80233	
H37Ra	*M. tuberculosis*		X78888	
KPM KY590	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014209	
KPM KY631	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014210	
KPM KY643	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014211	
KPM KY673	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014212	
KPM KY677	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014213	
KPM KY678	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014214	
KPM KY679	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014215	
KPM KY682	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014216	
KPM KY686	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014217	
KPM KY697	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014218	
KPM KY698	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014219	
KPM KY699	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014220	
KPM KY708	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014221	
KPM KY709	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014230	
KPM KY713	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014231	
KPM KY715	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014240	
KPM KY721	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014241	
KPM KY741	*M. tuberculosis*	Human clinical, 1987, Kyoto	AB014242	

was subjected to 35 cycles with a denaturation step at 95°C for 1 min, an annealing step at 63°C for 1 min, and an extension step at 72°C for 1 min. The amplified DNA was purified, and its sequence was determined by using Ec16S.1390p and Mb23S.44n as sequence primers.

ETR typing. Exact tandem repeat (ETR) typing was performed by using the combination of primers designated ETR-A and heat-treated crude lysates as DNA templates (7). The amplification reaction was subjected to 35 cycles with a denaturation step at 95°C for 1 min, an annealing step at 60°C for 1 min, and an extension step at 72°C for 2 min. Amplified fragment was analyzed by 1.5% agarose gel electrophoresis for typing.

Analysis of sequence data. The *gyrB* sequences were aligned by using CLUSTAL W 1.7 (29), and the alignment was manually corrected. Phylogenetic analyses were performed by the PHYLIP version 3.5c package (6). Distance matrices based on Kimura's two-parameter model (12) were produced by using the DNADIST program, and a neighbor-joining tree was constructed by using the NEIGHBOR program. The resulting trees were depicted by using the Tree-View version 1.5 package (16). The stability of the grouping was assessed by bootstrapping with the SEQBOOT, DNADIST, NEIGHBOR, and CONSENSE programs. A total of 1,000 bootstrapped trees were generated. Pairwise distances of the *gyrB* or ITS sequences were calculated by using the MEGA version 1.0 package (13).

***gyrB*-based species-specific PCR for the *M. tuberculosis* complex.** The *gyrB* sequences of four type strains belonging to the *M. tuberculosis* complex were aligned, and the four sites at which base substitutions occurred were found. Species-specific primers whose 3′ ends were at the base substitution sites were designed and named 675-T, 756-G, 756-A, 1410-G, 1410-A, 1450-C, and 1450-A (Table 2). The PCR reactions were performed in a final volume of 20 μl containing 1 μl of a boiled bacterial suspension, a reaction buffer (GeneAmp Kit; Perkin-Elmer), 1.25 U of AmpliTaq Gold DNA polymerase, a 0.1 mM concentration of each deoxynucleoside triphosphate, and 100 pmol of each primer. The reaction was subjected to 35 cycles of amplification, with a denaturation step at 95°C for 1 min and an annealing-extension step at 72°C for 1.5 min, by using a Progene thermal cycler (Techne). The amplified fragments were analyzed by 1.5% agarose gel electrophoresis.

***gyrB*-based species-specific PCR for *M. kansasii* and *M. gastri*.** The *gyrB* sequences of the type strains of *M. kansasii* and *M. gastri* were aligned with those of other mycobacterial and related species. The sequences specific to *M. kansasii* and *M. gastri* were then found. Species-specific primers whose 3′ ends corresponded to the species-specific sequences were designed and named KG632f, MK962r, and MG962r (Table 2). The PCR reactions were performed in a final volume of 20 μl containing 1 μl of a boiled bacterial suspension, a reaction buffer (GeneAmp Kit), 1.25 U of AmpliTaq Gold DNA polymerase, a 0.1 mM concentration of each deoxyribonucleoside triphosphate, and 100 pmol of each

TABLE 2. List of PCR primers used for *gyrB* analysis

Primer	Nucleotide sequence (5′ to 3′)	Location of primers[a]
UP1TL	CAy GCn GGn GGn AAr TTy GA	553–572
UP2rTL	TCn ACr TCn GCr TCn GTC AT	1831–1850
UP1EM13r	CAG GAA ACA GCT ATG ACC AyG snG GnG GnA ArT Tyr A	553–572
QTK-21	TGT AAA ACG ACG GCC AGT Ary TTn kyy TTn GTy TG	1351–1367
GGTH-r	CAG GAA ACA GCT ATG ACC GAn GGn GGn ACn CA	1177–1190
UP2r-21	TGT AAA ACG ACG GCC AGT rTC nAC rTC nGC rTC nGT CAT	1831–1850
675-T	AGA TCA AGC GCG ACG GGT AT	656–675
756-G	GAA GAC GGG GTC AAC GGT G	738–756
756-A	GAA GAC GGG GTC AAC GGT A	738–756
1410-G	CCA GTG GGT CAG CTG TTC G	1410–1428
1410-A	CCA GTG GGT CAG CTG TTC A	1410–1428
1450-C	CCT TGT TCA CAA CGA CTT TCG C	1450–1471
1450-A	CCT TGT TCA CAA CGA CTT TCG A	1450–1471
MTUBf	TCG GAC GCG TAT GCG ATA TC	574–593
MTUBr	ACA TAC AGT TCG GAC TTG CG	1594–1613
KG632f	GGT GTC TCG GTG GTC AAC GC	613–632
MK962r	GAC CTT GTG CGG GGC GGC GG	962–981
MG962r	CAC CTT GTG GGG GGC GGT GA	962–981

[a] Locations of primers are shown by the nucleotide number of *gyrB* of *M. tuberculosis* (GenBank accession number L27512).

primer. The reaction was subjected to 30 cycles of amplification, with a denaturation step at 95°C for 1 min and an annealing-extension step at 68°C for 1.5 min, by using a Progene thermal cycler (Techne). The amplified fragments were analyzed by 1.5% agarose gel electrophoresis.

PCR-restriction fragment length polymorphism (RFLP) analysis. The mycobacterial GyrB amino acid sequences were aligned, and two regions specific to the *M. tuberculosis* complex were found. Primers, which were designed from the specific amino acid sequences were named MTUBf and MTUBr (Table 2) and

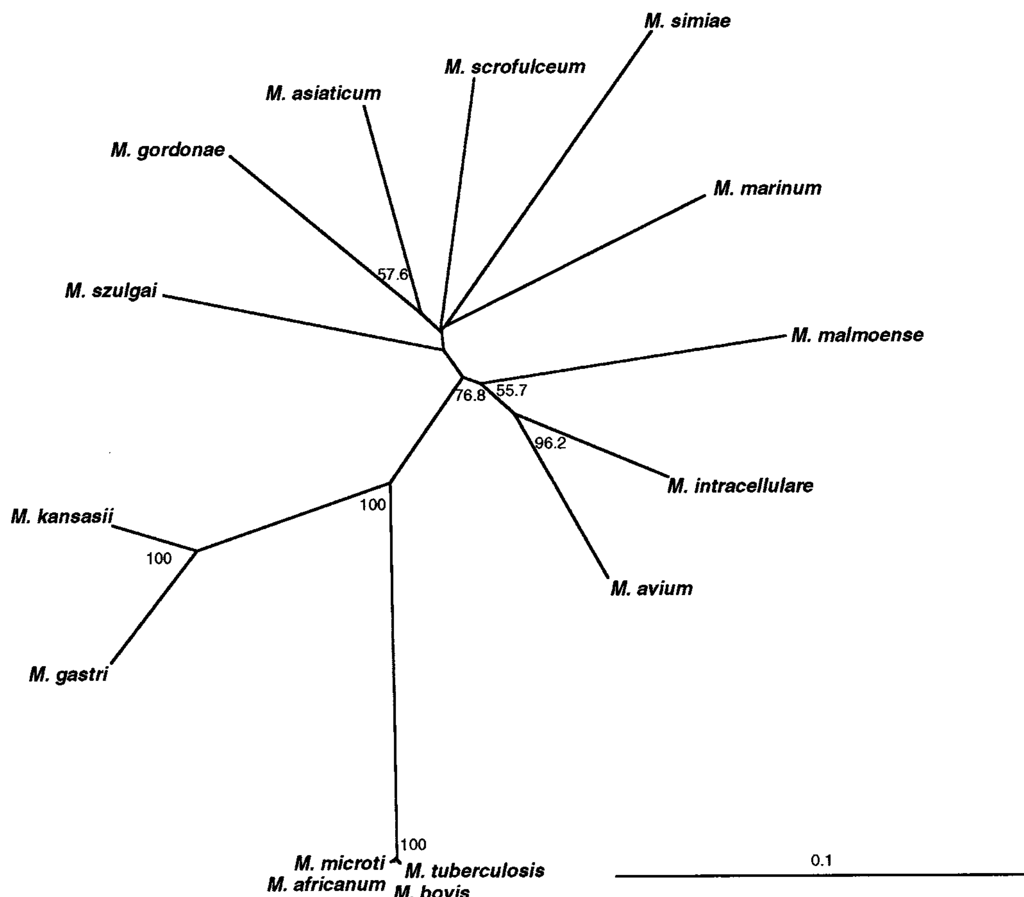


FIG. 1. Unrooted tree based on the *gyrB* sequences and showing the relationship between type strains of slowly growing mycobacteria. A neighbor-joining dendrogram was constructed. The numbers on the dendrogram are the percentages of occurrence in 1,000 bootstrapped trees; only values of >50% are shown.

used for PCR. The PCR reaction was performed under the same conditions as those for species-specific PCR already described. The amplified fragments were digested with *Rsa*I and *Taq*I, the digested fragments being analyzed by 2.0% agarose gel electrophoresis.

DNA-DNA hybridization. The levels of DNA-DNA relatedness were determined fluorometrically by the method of Ezaki et al. (4) by using biotinylated DNA.

Nucleotide sequence accession numbers. The *gyrB* and ITS sequences determined in this study have been deposited in DDBJ (DNA DataBank of Japan) under the accession numbers shown in Table 1.

RESULTS AND DISCUSSION

Phylogenetic analysis of slowly growing mycobacteria based on their *gyrB* sequences. Primers UP1TL and UP2rTL were used to amplify ~1.2-kb *gyrB* fragments by PCR from 15 mycobacterial species, and their nucleotide sequences were determined. The sizes of the amplified *gyrB* fragments were 1,257 bp for the *M. tuberculosis* complex, *M. kansasii*, and *M. gastri*, while those for the other species were longer by 6 bp. The alignment of the translated sequences showed that two residues were inserted in the N-terminal region: the amino acid sequences of the N terminus were Ser-Asp-(Ser or Ala)-Tyr-Ala-Ile-Ser for *M. kansasii*, *M. gastri*, and the *M. tuberculosis* complex, whereas they were Gly-Glu-Asn-Ser-Gly-Tyr-(Asn or Thr)-Val-Ser for the other species. The overall amino acid sequence identities ranged from 82 to 99%. The pairs showing more than 90% identity were *M. kansasii* and *M. gastri* (94%), *M. avium* and *M. intracellulare* (91%), and four species of the *M. tuberculosis* complex (99%). The other pairs showed <90% identity. The unrooted phylogenetic tree (Fig. 1) showed that the 15 mycobacterial species were well resolved by their *gyrB* sequences. Especially, branches of medically important species, such as *M. kansasii*, *M. gastri*, *M. avium*, *M. intracellulare*, and *M. tuberculosis* complex, were supported by high bootstrap values. The results indicated that the *gyrB* sequences would be useful for differentiation of those species.

As shown in Table 3, the resolution in the *gyrB*-based phylogeny was better than that in the ITS-based phylogeny, with some exceptions. The ITS sequence can evolve faster than 16S rDNA, as this region does not encode any biologically functional molecules, except for a role in pre-rRNA processing, and this sequence has been proposed to be suitable for differentiating closely related organisms (2, 9). In fact, Roth et al. (20) have shown that *M. kansasii* and *M. gastri* could be distinguished by using their ITS sequences. However, the ITS sequences of four species belonging to the *M. tuberculosis* complex were 100% identical. This means that the four species of the *M. tuberculosis* complex appear to have diverged quite recently, as has been indicated by Sreevatsan et al. (23).

Identification of *M. kansasii* and *M. gastri* by *gyrB*-targeted species-specific PCR. The *gyrB* sequences of the type strains of *M. gastri* and *M. kansasii* were 94.7% identical (1,191 of 1,257 bp). All except five substitutions were synonymous. The *gyrB* sequences of eight clinical isolates were determined to identify intraspecies variations in the *gyrB* sequences. The taxonomic identification of these isolates was done by DNA-DNA hybridization experiments: six of them were identified to be *M. kansasii*, and the remaining two were determined to be *M. gastri*. The *gyrB* sequences of these isolates were 100% identical to those of the type strains. Two stretches of DNA were found to be specific to *M. kansasii* and/or *M. gastri* when their *gyrB* sequences were aligned with those of 350 strains of high G+C gram-positive bacteria (H. Kasai and S. Harayama, unpublished data). Their locations corresponded to positions 613 to 632 and positions 962 to 981 of the *M. tuberculosis gyrB* sequence (GenBank accession number L27512). A set of primers adapted to these two specific sequences was designed. The

TABLE 3. Pairwise distance values for the *gyrB* sequences (lower left) and 16S-23S rDNA spacer (ITS) sequences (upper right)

Strain	Pairwise distance in strain:														
	T704	T021	T801	T901	KPM 2201	KPM 2403	KPM 1001	KPM 3504	KPM 1201	ATCC 25274	KPM 3012	KPM 3101	KPM 2027	KPM 3401	KPM 1403
M. bovis T702		0.0000	0.0000	0.0000	0.1888	0.1205	0.1205	0.1285	0.1285	0.1526	0.2048	0.2016	0.1486	0.1526	0.1727
M. tuberculosis T021	0.0024		0.0000	0.0000	0.1888	0.1205	0.1205	0.1285	0.1395	0.1526	0.2048	0.2016	0.1486	0.1526	0.1727
M. africanum T801	0.0016	0.0008		0.0000	0.1888	0.1205	0.1205	0.1285	0.1395	0.1526	0.2048	0.2016	0.1486	0.1526	0.1727
M. microti T901	0.0024	0.0016	0.0008		0.1888	0.1205	0.1205	0.1285	0.1395	0.1526	0.2048	0.2016	0.1486	0.1526	0.1727
M. gordonae KPM 2201	0.1756	0.1739	0.1748	0.1756		0.1325	0.1205	0.1365	0.1004	0.1365	0.2008	0.1727	0.1727	0.1767	0.1446
M. szulgai KPM 2403	0.1715	0.1699	0.1707	0.1715	0.1230		0.0602	0.0683	0.0723	0.0964	0.1526	0.1406	0.1124	0.1165	0.1285
M. kansasii KPM 1001	0.1472	0.1448	0.1456	0.1448	0.1594	0.1610		0.0643	0.0683	0.0964	0.1687	0.1446	0.1205	0.1285	0.1245
M. gastri KPM 3504	0.1505	0.1497	0.1505	0.1497	0.1739	0.1675	0.0526		0.0723	0.1044	0.1365	0.1285	0.1004	0.1044	0.1205
M. marinum KPM 1201	0.1780	0.1764	0.1772	0.1780	0.1254	0.1262	0.1586	0.1683		0.1004	0.1687	0.1325	0.1245	0.1285	0.1084
M. asiaticum ATCC 25274	0.1739	0.1723	0.1731	0.1739	0.1011	0.1100	0.1537	0.1586	0.1173		0.1566	0.1285	0.1365	0.1526	0.1084
M. avium KPM 3012	0.1667	0.1650	0.1659	0.1667	0.1262	0.1295	0.1392	0.1553	0.1262	0.1214		0.0482	0.0884	0.1044	0.1205
M. intracellulare KPM 3101	0.1642	0.1626	0.1634	0.1642	0.1270	0.1141	0.1481	0.1602	0.1254	0.1116	0.0801		0.0803	0.0924	0.0924
M. scrofulaceum KPM 2027	0.1739	0.1723	0.1731	0.1739	0.1116	0.1214	0.1561	0.1675	0.1214	0.1133	0.1157	0.1084		0.0522	0.0884
M. malmoense KPM 3401	0.1650	0.1634	0.1642	0.1650	0.1481	0.1343	0.1578	0.1659	0.1497	0.1351	0.1222	0.1116	0.1303		0.1205
M. simiae KPM 1403	0.1820	0.1804	0.1812	0.1804	0.1343	0.1497	0.1634	0.1715	0.1392	0.1311	0.1432	0.1481	0.1319	0.1610	

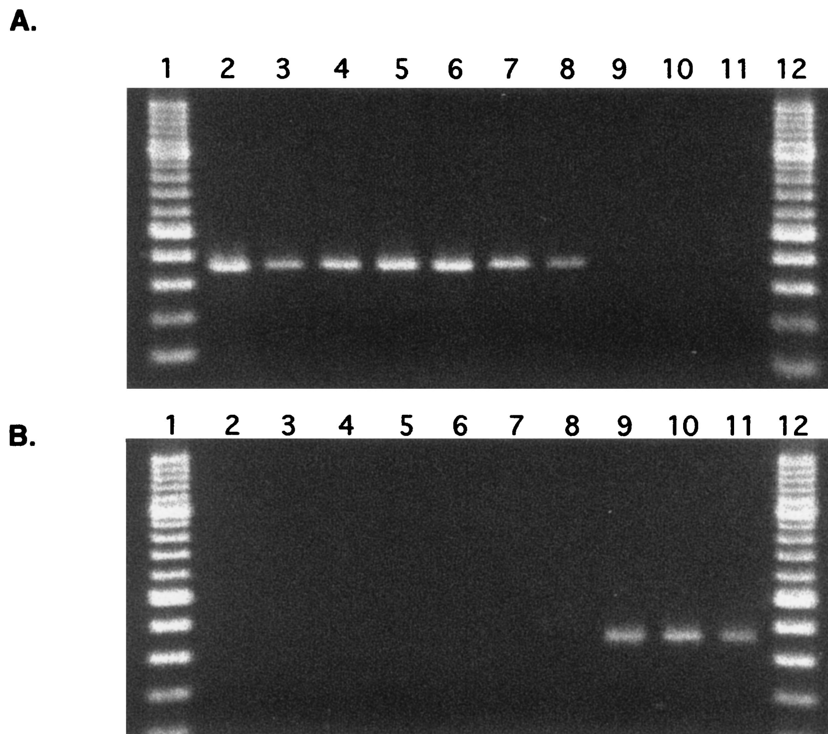


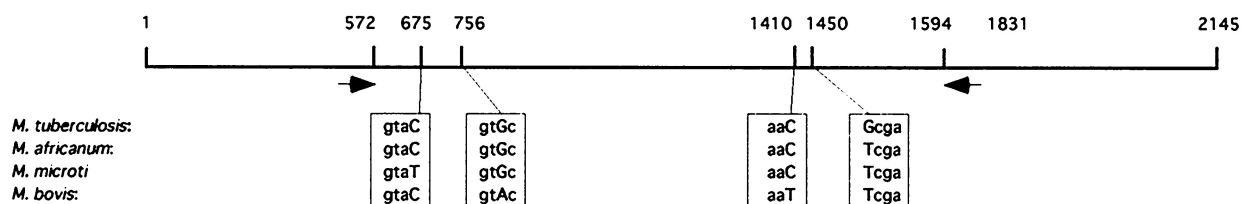
FIG. 2. PCR for the amplification of *gyrB* to detect *M. kansasii* and *M. gastri*. Panels A and B show the specificity of *M. kansasii*-specific primers and of *M. gastri*-specific primers, respectively. Lanes: 1 and 12, 100-bp ladder molecular size markers; 2 and 9, PCR products from DNA of the type strains of *M. kansasii* and *M. gastri*; 3 to 8, PCR products from the cell lysate of clinical isolates identified as *M. kansasii*; 10 and 11, PCR products from the cell lysate of clinical isolates identified as *M. gastri*. The length of each amplified fragment was 368 bp.

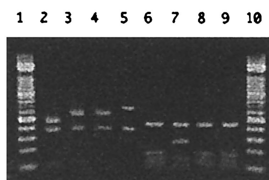
forward primer was complementary to both the *M. kansasii* and *M. gastri* sequences, while two types of reverse primers, one specific to *M. kansasii* and the other specific to *M. gastri*, were prepared. As shown in Fig. 2, *M. kansasii* and *M. gastri* could be distinguished by PCR with these primer sets.

***gyrB* sequences of strains in the *M. tuberculosis* complex.** We determined the *gyrB* sequences of the type strains of *M. tuberculosis*, *M. microti*, *M. africanum*, and *M. bovis*. The obtained sequences were compared with each other, and we found substitutions at four different sites (Fig. 3A). At position 675, the *gyrB* sequence of *M. microti* was T, and the *gyrB* sequences of the other strains were C. At position 756, the *gyrB* sequence of *M. bovis* was A, and the *gyrB* sequences of the other strains were G. At position 1410, the *gyrB* sequence of *M. bovis* was T, and the *gyrB* sequences of the other strains were C. At position 1450, the *gyrB* sequence of *M. tuberculosis* was G, and the *gyrB* sequences of the other strains were T. All the substitutions were synonymous, three being transitions and one being a transversion. These substitutions are unlikely to have developed under the selective pressure of therapeutic drug treatment because drug resistance is acquired by nonsynonymous substitution. Therefore, these substitutions can be regarded as the result of naturally occurring divergent evolution of these members. Synonymous substitution may thus be useful to deduce the natural relationship among pathogenic bacteria since they are not subjected to selection by antibiotics.

***M. tuberculosis* complex-specific PCR and differentiation of the members by PCR-RFLP.** A two-step method for differentiating the strains in the *M. tuberculosis* complex was developed. The first step involves detecting the *M. tuberculosis* complex, while the second step involves the differentiation of four species of the *M. tuberculosis* complex. The GyrB sequences of the *M. tuberculosis* complex were aligned with those of other mycobacterial species, and two unique sequences for the *M. tuberculosis* complex were found. A pair of primers, MTUB-f and MTUB-r, was then designed for specific amplification of the partial *gyrB* sequences from the *M. tuberculosis* complex. With these primers, the *gyrB* sequences were amplified from *M. tuberculosis*, but not from *M. kansasii*, *M. gastri*, *M. abscessus*, *M. chelonae*, or *M. triviale* (data not shown). The length of the amplified fragments was 1,020 bp. At the second step, the amplified DNA fragments were digested by *Rsa*I or *Taq*I. As shown in Fig. 3B, *M. bovis* (lane 2) and *M. microti* (lane 5) could be differentiated from the other two species by the electrophorogram of the *Rsa*I-digested fragments. Fragments of 500 and 700 bp can be found in the cases of *M. bovis* and *M. microti*, respectively, while 600-bp fragments were generated by *Rsa*I digestion of the PCR products from *M. tuberculosis* and *M. africanum. Taq*I digestion of the *M. tuberculosis gyrB* fragment generated a 300-bp fragment, while a 150-bp fragment was generated from the other three species. Thus, RFLP analysis of the *gyrB* fragment could distinguish four species belonging to the *M. tuberculosis* complex.

Differentiation of members of the *M. tuberculosis* complex by species-specific PCR. We were also able to differentiate the members of the *M. tuberculosis* complex by species-specific PCR. We designed PCR primers that allowed selective amplification of the *gyrB* fragments from each species of the *M. tuberculosis* complex. The sequences and their positions are listed in Table 2. According to the designs of these primers, the *gyrB* fragment of *M. tuberculosis* could be amplified only by using the primer set of 756-G and 1410-C, the *gyrB* fragment of *M. bovis* could be amplified only by using the primer set of 756-A and 1410-A, the *gyrB* fragment of *M. africanum* could be





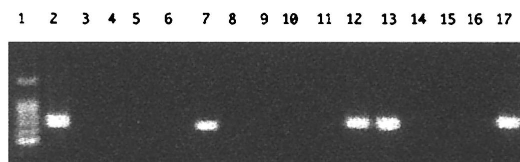


FIG. 3. Differentiation of members of the *M. tuberculosis* complex by using PCR and PCR-RFLP. (A) *gyrB* sequences of the members of the *M. tuberculosis* complex. The sequences of polymorphic loci are enclosed in boxes, and capital letters indicate substituted sequences. Black arrows show the positions of primers for the PCR-RFLP analysis. (B) PCR-RFLP patterns obtained with the type strains of the *M. tuberculosis* complex. Lanes: 1 and 10, 100-bp ladder molecular size markers; 2 to 9, RFLP patterns of PCR products from *M. microti* (lanes 2 and 6), *M. tuberculosis* (lanes 3 and 7), *M. africanum* (lanes 4 and 8), and *M. bovis* (lanes 5 and 9). Lanes 2, 3, 4, and 5 show RFLP patterns obtained by *Rsa*I digestion; lanes 6, 7, 8, and 9 show RFLP patterns obtained by *Taq*I digestion. (C) Species-specific PCR amplification. Lanes: 1, 100-bp ladder molecular size markers; 2, 6, 10, and 14, PCR products from *M. tuberculosis* DNA; 3, 7, 11, and 15, PCR products from *M. bovis* DNA; 4, 8, 12, and 16, PCR products from *M. microti* DNA; 5, 9, 13, and 17, PCR products from *M. africanum* DNA; 2, 3, 4, and 5, PCR with 210-G and 442-C; 6, 7, 8, and 9, PCR with 756-A and 1410-A; 10, 11, 12, and 13, PCR with 756-G and 1410-A; 14, 15, 16, and 17, PCR with 675-T and 1410-G.

amplified only by using the primer set of 756-G and 1410-A, and the *gyrB* fragment of *M. microti* could be amplified by using both the 756-G and 1450-A set and the 675-T and 1410-G set. Combination of these primers enabled four species of the *M. tuberculosis* complex to be differentiated (Fig. 3C).

***gyrB* sequences of clinically isolated *M. tuberculosis* strains.** To evaluate the general applicability of the *gyrB*-based identification method, we analyzed the *gyrB* sequences of 18 *M. tuberculosis* strains isolated from patients, one *M. bovis* strain isolated from a cow, and the Japanese reference strain of BCG. Each *M. tuberculosis* strain was isolated from a different patient. They were classified into five different groups based on the variable numbers of tandem repeats in the ETR-A locus (7). Group 1 contained two strains (KPM KY590 and KPM 631), group 2 also contained two strains (KPM KY682 and KPM 721), and groups 3 and 4 each contained a single strain (KPM KY673 and KPM 679, respectively). The remaining 12 strains belonged to group 5. Thus, the strains studied were derived from at least five independent origins. The *M. tuberculosis*-specific substitutions found in the *gyrB* sequence of the type strain were conserved in all of the clinical isolates. Both the *gyrB* sequences of *M. bovis* isolated from a cow and the Japanese BCG strain were identical to that of the type strain of *M. bovis*, although the origins of these three strains were different. These results suggest that the synonymous substitutions found in *M. tuberculosis* and *M. bovis* are species specific. Besides these three common substitutions, we found three strain-specific base substitutions in the four *M. tuberculosis* strains. In the *gyrB* sequence of strain KPM KY673, the A at position 883 was found to be substituted by C, provoking the amino acid substitution from Ala to Pro. The A at position 1114 was found to be substituted by G in strain KPM KY679, while the G at position 1116 was found to be substituted by C in strains KPM KY678 and KPM H37Rv (GenBank accession number L27512). The latter two changes occurred in the same codon, ATG, provoking a change in the amino acid sequence from Met to Val or Ile. These two amino acid substitutions are found at positions that are highly conserved among the GyrB proteins of high G+C gram-positive bacteria. It is thus possible that the amino acid changes at these sites could confer some drug resistance to the strains.

Medical applications of *gyrB* sequences. Many clinically significant strains are classified as slowly growing mycobacteria. As shown in this study, more substitutions were found in the *gyrB* sequences than in the 16S rDNA and ITS sequences of slowly growing mycobacterial strains. With the benefit of higher base substitution frequencies, PCR or PCR-RFLP of *gyrB* could be used for the species identification of clinically isolated mycobacteria. These methods are highly applicable to clinical medicine because DNA sequencing is not required for rapid identification of these species.

The *gyrB*-based methods have already shown to be useful for differentiation of closely related strains of other bacteria, such as *Vibrio* (33) and *Bacillus* (39) spp. For quantitation analysis, too, the *gyrB*-based methods would be more useful than 16S rDNA-based methods because the copy number of *gyrB* is single in almost all of the bacteria examined (35, 36), while that of 16S rDNA is variable. We have determined the *gyrB* sequences of various bacteria and deposited them in the *gyrB* database (11), which recently became accessible via the internet (http://www.mbio.co.jp/icb/icb.html/). The sequence data accumulated in the *gyrB* database will be useful for the development of bacterial diagnostic systems based on molecular methods such as PCR, PCR-RFLP, and high-density DNA probe arrays (8, 30).

ACKNOWLEDGMENTS

We are grateful to Isao Kawamoto for helpful suggestions. We thank Yukiko Itazawa and Yuhka Takahashi for their technical assistance.

This work was performed as part of The Industrial Science and Technology Frontier Program supported by New Energy and Industrial Technology Development Organization.

REFERENCES

1. **Baess, I.** 1979. Deoxyribonucleic acid relatedness among species of slowly-growing mycobacteria. Acta Pathol. Microbiol. Scand. **87:**221–226.
2. **Barry, T., G. Colleran, M. Glennon, L. K. Dunican, and F. Gannon.** 1991. The 16S/23S ribosomal spacer region as a target for DNA probes to identify eubacteria. PCR Methods Appl. **1:**51–56.
3. **Comicini, S., D. Barbarini, S. Telecco, L. Bonno, and P. Marone.** 1998. Rapid identification of *Mycobacterium tuberculosis* and *Mycobacterium avium* by polymerase chain reaction and restriction enzyme analysis within sigma factor regions. New Microbiol. **21:**391–395.
4. **Ezaki, T., Y. Hashimoto, and E. Yabuuchi.** 1989. Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to determine genetic relatedness among bacterial strains. Int. J. Syst. Bacteriol. **39:**224–229.
5. **Feizabadi, M. M., I. D. Robertson, D. V. Cousins, and D. J. Hampson.** 1996. Genomic analysis of *Mycobacterium bovis* and other members of the *Mycobacterium tuberculosis* complex by isozyme analysis and pulsed-field gel electrophoresis. J. Clin. Microbiol. **34:**1136–1142.
6. **Felsenstein, J.** 1993. PHYLIP (Phylogeny Inference Package), version 3.5c. Department of Genetics, University of Washington, Seattle.
7. **Frothingham, R., and W. A. Meeker-O'Connell.** 1998. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. Microbiology **144:**1189–1196.
8. **Gingeras, T. R., G. Ghandour, E. Wang, A. Berno, P. M. Small, F. Drobniewski, D. Alland, E. Desmond, M. Holodniy, and J. Drenkow.** 1998. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. Genome Res. **8:**435–448.
9. **Gürtler, V., and V. A. Stanisich.** 1996. New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. Microbiology **142:**3–16.
10. **Harayama, S., and S. Yamamoto.** 1996. Phylogenetic identification of *Pseudomonas* strains based on a comparison of *gyrB* and *rpoD* sequences, p. 250–258. *In* T. Nakazawa, K. Furukawa, D. Haas, and S. Silver (ed.), Molecular biology of pseudomonads. ASM Press, Washington, D.C.
11. **Kasai, H., K. Watanabe, G. Elisabeth, A. Bairoch, K. Isono, and S. Harayama.** 1998. Construction of the *gyrB* database for identification and classification of bacteria, p. 13–21. *In* Genome informatics 1998. Universal Academy Press, Tokyo, Japan.
12. **Kimura, M.** 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16:**111–120.
13. **Kumar, S., K. Tamura, and M. Nei.** 1994. MEGA: molecular evolutionary genetics analysis software for microcomputers. Comput. Appl. Biosci. **10:**189–191.
14. **Magdalena, J., P. Supply, and C. Locht.** 1998. Specific differentiation between *Mycobacterium bovis* BCG and virulent strains of the *Mycobacterium tuberculosis* complex. J. Clin. Microbiol. **36:**2471–2476.
15. **Noordhoek, G. T., A. H. Kolk, G. Bjune, D. Catty, J. W. Dale, and P. E. Fine.** 1994. Sensitivity and specificity of PCR for detection of *Mycobacterium tuberculosis* isolates: a blind comparison study among seven laboratories. J. Clin. Microbiol. **32:**277–284.
16. **Page, R. D. M.** 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12:**357–358.
17. **Plikaytis, B. B., B. D. Plikaytis, M. A. Yakurus, W. R. Bulter, 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. J. Clin. Microbiol. **30:**1815–1822.
18. **Rodriguez, J. G., G. A. Mejia, P. D. Portillo, M. E. Patarroyo, and L. A. Murillo.** 1995. Species-specific identification of *Mycobacterium bovis* by PCR. Microbiology **141:**2131–2138.
19. **Rogall, T., J. Wolters, T. Floher, and E. C. Böttger.** 1990. Towards a phylogeny and definition of the species at the molecular level within the genus *Mycobacterium*. Int. J. Syst. Bacteriol. **40:**323–330.
20. **Roth, A., M. Fischer, M. E. Hamid, S. Michalke, W. Ludwig, and H. Mauch.** 1998. Differentiation of phylogenetically related slowly growing mycobacteria based on 16S-23S rRNA gene internal transcribed spacer sequences. J. Clin. Microbiol. **36:**139–147.
21. **Shivannavar, C. T., V. M. Katoch, V. D. Sharma, M. A. Patil, K. Katoch, V. P. Bharadwaj, R. K. Sharma, A. S. Bhatia, and B. M. Agrawal.** 1996. Determination of mycobacterial phylogeny on the basis of immunological relatedness of superoxide dismutases. Int. J. Syst. Bacteriol. **46:**1164–1169.
22. **Soini, H., E. C. Böttger, and M. K. Viljanen.** 1994. Identification of mycobacteria by PCR-based sequence determination of the 32-kilodalton protein gene. J. Clin. Microbiol. **32:**2944–2947.
23. **Sreevatsan, S., X. Pan, K. E. Stockbauer, N. D. Connel, B. N. Kreiswirth, T. S. Whittam, and J. M. Musser.** 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. USA **94:**9869–9874.
24. **Stahl, D. A., and J. W. Urbance.** 1990. The division between fast- and slowly growing species corresponds to natural relationships among the mycobacteria. J. Bacteriol. **172:**116–124.
25. **Swanson, D. S., V. Kapur, K. Stockbauer, X. Pan, R. Frothingham, and J. M. Musser.** 1997. Subspecific differentiation of *Mycobacterium avium* complex strains by automated sequencing of a region of the gene (*hsp65*) encoding a 65-kilodalton heat shock protein. Int. J. Syst. Bacteriol. **47:**414–419.
26. **Takewaki, S., K. Okuzumi, H. Ishiko, K. Nakahata, A. Ohkubo, and R. Nagai.** 1993. Genus-specific polymerase chain reaction for the mycobacterial *dnaJ* gene and species-specific oligonucleotide probes. J. Clin. Microbiol. **31:**446–450.
27. **Takewaki, S., K. Okuzumi, I. Manabe, M. Tanimura, K. Miyamura, K. Nakahata, Y. Yazaki, A. Ohkubo, and R. Nagai.** 1994. Nucleotide sequence comparison of the mycobacterial *dnaJ* gene and PCR-restriction fragment length polymorphism analysis for identification of mycobacterial species. Int. J. Syst. Bacteriol. **44:**159–166.
28. **Telenti, A., F. Marchesi, M. Bals, F. Bally, E. C. Böttger, 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.
29. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
30. **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.
31. **van Soolingen, D., T. Hoogenboezem, P. E. W. de Haas, P. W. M. Hermans, M. A. Koedam, K. S. Teppema, P. J. Brennan, G. S. Besra, F. Portaels, J. Top, L. M. Shouls, and J. D. A. van Embden.** 1997. A novel pathogenic taxon of the *Mycobacterium tuberculosis* complex, Canetti: characterization of an exceptional isolate from Africa. Int. J. Syst. Bacteriol. **47:**1236–1245.
32. **van Soolingen, D., A. G. M. van der Zanden, P. E. W. de Haas, G. T. Noordhoek, A. Kiers, N. A. Foudraine, F. Portaels, A. H. J. Kolk, K. Kremer, and J. D. A. van Embden.** 1998. Diagnosis of *Mycobacterium microti* infections among humans by using novel gene markers. J. Clin. Microbiol. **36:**1840–1845.
33. **Venkateswaran, K., N. Dohmoto, and S. Harayama.** 1998. Cloning and nucleotide sequence of the *gyrB* gene of *Vibrio parahaemolyticus* and its application in detection of this pathogen in shrimp. Appl. Environ. Microbiol. **64:**681–687.
34. **Victor, T. C., A. M. Jordaan, E. J. van Schalkwyk, G. J. Coetzee, and P. D. van Helden.** 1996. Strain-specific variation in the *dnaJ* gene of mycobacteria. J. Med. Microbiol. **44:**332–339.
35. **Watanabe, K., S. Yamamoto, S. Hino, and S. Harayama.** 1998. Population dynamics of phenol-degrading bacteria in activated sludge determined by *gyrB*-targeted quantitative PCR. Appl. Environ. Microbiol. **64:**1203–1209.
36. **Watanabe, K., M. Teramoto, and S. Harayama.** 1999. An outbreak of nonflocculating catabolic populations caused by the breakdown of a phenol-digesting activated-sludge process. Appl. Environ. Microbiol. **65:**2813–2819.
37. **Wayne, L. G., and G. P. Kubica.** 1986. The mycobacteria, p. 1435–1457. *In* P. H. A. Sneath and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 2. Williams and Wilkins Co., Baltimore, Md.
38. **Wayne, L. G., R. C. Good, M. I. Krichevsky, Z. Blacklock, H. L. David, D. Dawson, W. Gross, J. Hawkins, P. A. Jenkins, I. Juhlin, W. Käppler, H. H. Kleeberg, V. Levy-Frebault, C. McDurmont, E. E. Nel, F. Portaels, S. Rüsch-Gerdes, K. H. Schröder, V. A. Silcox, I. Szabo, M. Tsukamura, L. van den Breen, B. Vergmann, and M. A. Yakrus.** 1989. Third report of the cooperative, open-ended study of slowly growing mycobacteria by the International Working Group on Mycobacterial Taxonomy. Int. J. Syst. Bacteriol. **39:**267–278.
39. **Yamada, S., E. Ohashi, N. Agata, and K. Venkateswaran.** 1999. Cloning and nucleotide sequence analysis of *gyrB* of *Bacillus cereus*, *B. thuringiensis*, *B. mycoides*, and *B. anthracis* and their application to the detection of *B. cereus* in rice. Appl. Environ. Microbiol. **65:**1483–1490.
40. **Yamamoto, S., and S. Harayama.** 1995. PCR amplification and direct sequencing of *gyrB* genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. Appl. Environ. Microbiol. **61:**1104–1109.
41. **Yamamoto, S., and S. Harayama.** 1996. Phylogenetic analysis of *Acineto-*

bacter strains based on the nucleotide sequences of *gyrB* genes and on the amino acid sequences of their products. Int. J. Syst. Bacteriol. **46:**506–511.

42. **Yamamoto, S., P. J. M. Bouvet, and S. Harayama.** 1999. Phylogenetic structures of the genus *Acinetobacter* based on the *gyrB* sequences: comparison with the grouping by DNA-DNA hybridization. Int. J. Syst. Bacteriol. **49:** 87–95.
43. **Yamamoto, S., and S. Harayama.** 1998. Phylogenetic relationships of *Pseudomonas putida* strains deduced from the nucleotide sequences of *gyrB*, *rpoD*, and 16S rRNA genes. Int. J. Syst. Bacteriol. **48:**813–819.
44. **Zolg, J. W., and S. Philippi-Schulz.** 1994. The superoxide dismutase gene, a target for detection and identification of mycobacteria by PCR. J. Clin. Microbiol. **32:**2801–2812.