Novel Polymorphic Region of the *rpoB* Gene Containing *Mycobacterium* Species-Specific Sequences and Its Use in Identification of Mycobacteria

Hyeyoung Lee,¹ Hye-Eun Bang,² Gill-Han Bai,³ and Sang-Nae Cho^{2,4}*

Department of Biomedical Laboratory Science, College of Health Sciences, Yonsei University, Wonju 220-710,¹ and Department of Microbiology and Brain Korea 21 Project for Medical Sciences, Yonsei University College of Medicine, Seoul 120-752,² and Department of Microbiology, Korean Institute of Tuberculosis, Seoul 137-140,³ and The International Vaccine Institute, Seoul 151-600,⁴ Republic of Korea

Received 5 August 2002/Returned for modification 23 September 2002/Accepted 23 February 2003

Sequence analysis of a specific region of the mycobacterium *rpoB* gene in 35 mycobacterial strains representing 26 different mycobacterial species of clinical importance showed that there exists a highly polymorphic region. Based on the sequences of the polymorphic region, the oligonucleotide probes of 14 mycobacterial species with relatively high clinical importance were designed and shown to be specific to their corresponding mycobacterial species by dot blot hybridization. The results showed that the probes designed in this study are highly specific to each mycobacterial species, which suggests that these sequences may be useful for the species identification of mycobacteria.

Tuberculosis (TB) is still a major public health problem in the world, with about 8 million new cases and over 2 million deaths reported annually. With the recent dissemination of human immunodeficiency virus infection throughout the world, infections with nontuberculous mycobacteria as well as TB have increased in many parts of the world over the last decade. For example, Mycobacterium avium infections accounted for almost 50% of mycobacterial infections among AIDS patients in certain geographical areas (4, 7). In order to provide proper drug regimens to patients with mycobacterial infections, it is important that species be identified correctly and rapidly, because drug regimens for TB may differ from those for other mycobacterial infections. In addition, with the wide use of liquid culture systems such as the BACTEC system (Becton Dickinson Diagnostic Instrument Systems, Sparks, Md.) and the MB/BacT system (Organon Teknika Corp., Boxtel, The Netherlands), the rapid identification of mycobacterial species has become even more important (3, 17).

Although biochemical tests are available for *Mycobacterium* species identification, it has proven to be difficult to use these tests because of time-consuming and often incorrect identification. In order to overcome such difficulties, high-performance liquid chromatography has been widely used for species identification based on mycolic acid analysis (1, 14, 25). In addition, with recent developments in molecular techniques and the availability of genome sequencing data, several molecular tests have been developed and are used in clinical mycobacterial laboratories. rRNA sequences, notably that of 16S rRNA, have been most widely used for mycobacterium species identification (18), and commercial kits based on such sequences are available (AccuProbe; Gen-Probe Inc., LiPA, In-

nogenetics N. V., Zwijnaarde, Szinjdrecht, Belgium). In addition, the *hsp65* gene (8, 15, 16, 23), the intergenic region between 16S and 23S rRNA (19), and the *rpoB* gene (9, 10, 13) are among the targets for molecular technique-based species identification. Sequencing (11, 15, 18, 20–22), DNA hybridization (6, 12), PCR-restriction fragment length polymorphism analysis (RFLP) (8, 9, 13, 16, 23), and microarray technology have also been employed to differentiate *Mycobacterium* species (5, 24).

Among the target genes, we were particularly interested in the rpoB gene. PCR sequence analysis of a region of the rpoB gene was suggested as a possible means of differentiating 44 species (9, 10, 13). We have also reported a new RFLP method (13) based on a different region of the *rpoB* gene, which is located between the first variable region (V1) and the second conserved region (C2), as determined using the genetic information of the Escherichia coli rpoB gene. The 360-bp region of the rpoB gene (bases 902 to 1261 and codons 302 to 420 of the rpoB gene of M. tuberculosis; GenBank accession number P47766) was found to be useful in the differentiation of more than 50 Mycobacterium species by a simple RFLP using two restriction enzymes. This clearly indicates that this 360-bp region of the *rpoB* gene contains highly informative sequences. In the present study, we analyzed sequences of this rpoB region of 35 mycobacterial strains representing 26 different mycobacterial species and prepared DNA probes that can be used in simple DNA hybridization tests for the identification of Mycobacterium species.

A total of 48 mycobacterial reference strains representing 39 *Mycobacterium* species were used for the PCR amplification of the 360-bp region of the *rpoB* gene in the present study (Table 1). Among them, 39 mycobacterial strains were obtained from the Korean Institute of Tuberculosis (KIT), Seoul, Korea, and three species were obtained from the Korean Collection for Type Cultures (KCTC) at the Korean Research Institute of Bioscience and Biotechnology (KRIBB). *M. abscessus*, which

^{*} Corresponding author. Mailing address: Department of Microbiology, Yonsei University College of Medicine, 134 Shinchon-dong, Seoul 120-752, Republic of Korea. Phone: 822-361-5282. Fax: 822-392-9310. E-mail: raycho@yumc.yonsei.ac.kr.

TABLE 1. Bacterial strains used in this study

	Species	Strain	Source ^a
М.	africanum	ATCC 25420	КІТ
М.	scrofulaceum	ATCC 19981	KIT
М.	gilvum	ATCC 43909	KIT
М.	gastri	ATCC 15754	KIT
М.	asiaticum	ATCC 25276	KIT
М.	aurum	ATCC 23366	KIT
М.	avium	ATCC 25291	KIT
М.	moriokaense	ATCC 43059	KRIBB
М.	abscessus	Pettenkofer Institute	YUMC
М.	celatum type I	ATCC 51130	KIT
М.	celatum type II	ATCC 51131	KIT
М.	chelonae	ATCC 35749	KIT
М.	bovis	ATCC 19210	KIT
М.	flavescens	ATCC 14474	KIT
М.	fortuitum type I	ATCC 6841	KIT
М.	fortuitum type II	ATCC 49404	KIT
М.	gallinarum	ATCC 19710	KRIBB
М.	genavense	ATCC 51233	KIT
М.	microti	ATCC 19422	KIT
М.	gordonae type I	ATCC 14470	KIT
М.	gordonae type II		KIT
М.	gordonae type III		KIT
М.	gordonae type IV		KIT
М.	haemophilum	ATCC 29548	KIT
М.	intracellulare	ATCC 13950	KIT
М.	interjectum	ATCC 51457	KIT
М.	intermedium	ATCC 51848	KIT
М.	kansasii type I		Pasteur Institute
М.	kansasii type II		Pasteur Institute
<i>M</i> .	kansasii type III		Pasteur Institute
<i>M</i> .	kansasu type IV		Pasteur Institute
<i>M</i> .	kansasu type V	1000 1000	Pasteur Institute
<i>M</i> .	тисодепісит	ATCC 49650	KIT
<i>M</i> .	neoaurum	ATCC 25795	KII
M.	nonchromogenicum	ATCC 19530	
M.	parafortuttum	ATCC 19686	
M.	peregrinum	ATCC 14467	
M.	phiel	ATCC 11/38	
M.	puiveris	ATCC 35154	KKIBB
M. M	maimoense	ATCC 29571	
M. M	marinum	ATCC 25700	
1VI. M	szuigui tarraa	ATCC 15755	KII KIT
1VI. M	thermoresistibile	ATCC 15755	KII KIT
1VI. M	triviala	ATCC 1932/	KII KIT
1VI. M	ulcarans	ATCC 10/22	KII KIT
1V1. M	vaccaa	ATCC 15425	KIT
1V1. M	vaccae	ATCC 19405	KIT
M	tuberculosis H37Py	ATCC 17250	KIT
		AICC 2/294	1311

was recently separated from *M. chelonae* as an independent new species, was obtained from Department of Clinical Pathology at Yonsei University Medical Center (YUMC). Finally, five subtypes of *M. kansasii* were generously provided by V. Vincent at the Laboratoire de Référence des Mycobactéries, Institut Pasteur, Paris, France. Clinical isolates that were subjected to dot blot hybridization to evaluate the specificity of each mycobacterial species-specific probes were obtained from the KIT. All clinical isolates used in this study were identified on the basis of conventional tests that included microbiological characterization and biochemical tests and an *rpoB*-based RFLP method (13) to precisely identify the clinical isolates.

The primer sets used to amplify the target *rpoB* gene were 5'-TCAAGGAGAAGCGCTACGA-3' (RPO5') and 5'-GGA

TGTTGATCAGGGTCTGC-3' (RPO3'), which resulted in a 360-bp PCR product (13). PCR was carried out in a final volume of 50 µl with 10 µl of DNA supernatant containing approximately 10 ng of genomic DNA, 10 pmol of each primer, 2 mM MgCl₂, 200 µM concentrations of deoxynucleoside triphosphates, and 1 U of DyNAzymeII DNA polymerase (Finnzymes, Espoo, Finland). DNA samples were first denatured completely by incubation at 94°C for 5 min and then amplified using 35 cycles of (i) denaturation at 94°C for 1 min, (ii) primer annealing at 58°C for 1 min, and (iii) elongation at 72°C in a thermocycler (Perkin-Elmer model 9600; Applied Biosystems, Foster City, Calif.). After the last amplification cycle, the samples were incubated further at 72°C for 7 min to obtain complete elongation of the final PCR products. Positive and negative controls were always included in each PCR. The positive control was the PCR mix with the DNA of the reference strain, M. bovis, and the negative control was a PCR mix without any DNA. After the PCR, the amplification results were visualized using 1.5% agarose gel electrophoresis and ethidium bromide staining.

For sequencing, PCR products were purified using the GeneClean III kit (Bio 101, Vista, Calif.) and cloned into a PCR-TOPO vector in the TOPO TA cloning kit (Invitrogen Co., Carlsbad, Calif.). The TOPO vectors containing PCR products were used for transformation of TOP10 competent cells (Invitrogen Co.). Plasmids containing inserts were purified from broth cultures with a Qiagen (Valencia, Calif.) plasmid kit and sequenced with the AutoRead sequencing kit and ALF DNA sequencer (Pharmacia Biotech, Uppsala, Sweden). Sequences were aligned using the Multialign program developed by F. Corpet (2).

In order to characterize the genetic nature of the 360-bp region of the *rpoB* gene, a total of 35 reference strains representing 26 different mycobacterial species were sequenced. Some species such as M. gordonae, M. kansasii, M. celatum, and M. fortuitum are known to have several subspecies, and thus these subspecies were also included in the sequence analysis. Among the 360-bp region sequenced in this study, sequences of 216 bp that have not been reported elsewhere are shown in Fig. 1, which shows that there exists a highly polymorphic region (black letters) flanked by highly conserved regions (red letters). Interestingly, the highly polymorphic region seemed to be suitable for the differentiation of mycobacterial species. For example, species differentiation between M. kansasii and M. gastri was possible since the sequences of M. kansasii are different from those of M. gastri, whose differentiation is not possible by 16S rRNA sequence analysis (18). Moreover, these polymorphic sequences were different even between highly closely related species, such as M. abscessus and M. chelonae or M. fortuitum and M. peregrinum, whose exact species identification has been extremely difficult by conventional culturebased microbiological and biochemical tests. In addition, the sequences of this polymorphic region in subspecies of M. kansasii, M. fortuitum, and M. gordonae were also differentiable, suggesting that this region of the rpoB gene may be used as a molecular signature for the differentiation of mycobacteria to the species or even to the subspecies level. However, there was no sequence difference in this region among species of the M. tuberculosis complex, including M. tuberculosis, M. bovis, M. microti, and M. africanum.

	1 10	20	30	40	50	60	70	80	90	1	100	110
M.gordonaeIV	CCTGGCCCGTGTCGG	CCGCTACAAGO	TCAACAAGA	AGCTGGGCCTG	CATGT	-CG-GCGATC	CGATCAC	CAGCT	CGACGC1	GACCGA	GAGGACGT	CGTCO
M.gordonaeI	CCTGGCCCGGGTAGG	CCGCTACAAG	itcaacaaga	AGCTCGGCCTG	CACGT	-CG-GCGATC	CGATCAC	CAGCT	CCACGC1	GACCGAC	GAAGACGT	CGTCO
M.gordonaeIII	CCTGGCCCGTGTCGG	CCGCTACAAGO	TCAACAAGA	AGCTCGGCCTG	CACGT	-CG-GCGATC	CGATCAC	CAGCT	CCACGC1	GACCGA	IGAAGACGT	CGTCG
M.aviun M.internellulare	CETECCECCTETECC	ILLULI HLHHUU	TCOOCOOCO	HULTLUULLTU	LHLUL	-66-666666	CCOTCOC	LHGLI	CCOC	COCCCO	COOCOCCT	LUILU
		CCGCTACAAGE	TCAACAAGA	AGCTCGGTCTG	AACGT	-CG-GCAAGC	CGATCAC	CAGCT	CGACGC1	GACCGA	GAAGACGTI	CGTAG
M.szulgai	CCTGGCTCGCGTCGG	CCGTTACAAGO	TCAACAAAA	AGCTCGGTCTG	AACGT	-CG-GCGAGC	CGATCAC	CAGTT	CGACGCT	GACCGA	GAGGATGT	CGTCG
M.ulcerans	CCTGGCTCGCGTGGG	itcggta <mark>ca</mark> age	itcaacaaga	AGCTCGGCCTG	AACGC	-CG-GCCAGC	CCATCAC	CAGCT	CGACGC1	GACCGAC	GAAGAC GT(CGTCG
M.marinum	CCTGGCCCGGGTGGG	CCGGTACAAGG	TCAACAAGA	AGCTCGGCCTG	AACGC	-CG-GCCAGC	CCATCAC	CAGCT	CGACGC1	GACCGA	GAAGACGT	CGTCG
M.peregrinum	CUTGGUUUGUGTUGG	ICCGCTOCOOCO	TCOOCOOCO	HGUIGGGUUIG	HHCGC	-CG-CCCCCC		GILGI	CGHCCCI	GHUUGHU	COOCOCCT	
M fortuitumTT		CCGCTACAAGA	TCAACAAGA	AGETGGGEETG	AACGC	-CG-GCCAGC	CGATCAC		CGACTC1	GACCGA	GAAGACGTI	CGTCG
M.genavense	CCTGGCCCGCGTCGG	CCGCTACAAGO	TCAACAAGA	AGCTGGGGCTG	CACGC	-CG-GCGAGC	CGATCAC	GTCGT	CGACGTT	GACCGAC	GAAGACGT	CGTCG
M.simiae	CTTGGCCCGCGTCGG	CCGCTACAAG	TCAACAAGA	AGCTGGGGCTG	CACGC	-CG-GCGAGC	CGATCAC	GTCGT	CGACGT1	GACCGAC	GAAGAC GT(CGTCG
M.tuberculosis	CCTGGCCCGCGTCGG	TCGCTATAAGO	TCAACAAGA	AGCTCGGGCTG	CATGT	-CG-GCGAGC	CCATCAC	GTCGT	CGACGC1	GACCGA	GAAGACGT	CGTGG
H bouis	CCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	TCCCTATAACO	TCOOCOOCO	HULILUUULIU ACCTCCCCCTC	CATCT		CCATCAC		LUHLUL CCACCC1	GACCEAL	IUNHUNUU II ICAACACCTI	LUIUU CETEE
M.africanum	CCTGGCCCGCGTCGG	TCGCTATAAGE	TCAACAAGA	AGCTCGGGCTG		-CG-GCGAGC	CCATCAC	GTCGT	CGACGC1	GACCGA	GAAGACGT	CGTGG
M.celatumII	CCTCGCGCGGGGTGGG	CCGCTACAAGO	ITCAACAAGA	AGCTCGGCCTG	AACAC	-CG-CGTCCC	CGATCAC	GACGA	CCACTC1	GACCGAR	GAGGACGT	CGTCG
M.celatumI	CCTGGCCCGGGTGGG	ICCGCTACAAGO	itcaacaaga	AGCTCGGCCTG	AACAC	-CG-AGAGCC	CAATCAC	CACCA	CGACGC1	CACCGAF	GAGGACGT	CGACO
M.kansasiiI	CCTGGCCCGTGTCGG	CCGATACAAGO	ITCAACAAGA	AGCTGGGCCTG	AACAC	-CA-ATCATC	CGATCAC	CACGA	CGACGC1	GACCGA	IGAAGACGT	CGTCG
M,kansasii¥	CCTCCCCCCCCCCTCC	TCCTTOCOOCO	TCOOCOOCO	HULIUUULIU	HHLHL COCCC	-LG-HILHIL	CCOTCOC	LHLUH	CCOCCC1	COCCCO	COCCOCCT	LUIUU CCTCC
M. snegnatis	CCTGGCCCGTGTCGG	CCGTTACAAGE	TCAACAAGA	AGCTGGGCCTG	AAC6C	-66-6C886C	CGATCAC	CAGCT	CGAC6C1	GACCGAR	GAGGACGTI	
M.nalmoense	CCTGGCCAGGGTTGG	CCGTTACAAGO	TCAACAAGA	AGCTCGGGCTG	CCGGCGG	CCG-AGTCGG	CCGTACC	CGCCT	CGACCACGCT	GACCGAF	GCGGATGT	CGTCG
M.xenopi	CCTGGCCCGGGTGGG	CCGCTACAAG	TCAACAAGA	AACTCGGGCTG	AACAC	-CG-AGAATG	CGCCAAC	CACCA	CGACCC1	GACCGAR	IGAGGAC GTI	CGTCG
M.chelonae	CCTGGCCCGCGTGGG	CCGGTACAAGO	TGAACAAGA	AGCTGGGTCTT	GG <mark>CG</mark>	-GT-GCCAAC	CCGGCTC	-TGGTGACTG	CCACCACGCT	CACCGAC	GAAGACGT	CGTCG
M. Clausessus	TUT66UUU6U6T666	TCCCTOCOOCO	I GHHCHHGH	HULTUUUUU	66666	-GC-HCCHHI		-HGGIGHCCH		COCCCO	COCCOCCT	
H.terrae		TCGCTACAAGE	TCAACAAGA	AGCTGGGCCTG	CC66C		CGACGTC	GAGACGTCGC	CGACCACGCI	GACCGAR	GAGGACATI	CGTGG
M.scrofulaceum	CCTGGCCCGCGTCGG	CCGCTACAAGO	TCAACAAGA	AGCTGGGTCTG	CACGC	-CG-GCGAGC	CGATC	ACGT	CGTCCACGCT	GACCGAC	GAAGACGT	CGTCG
M.gastri	CCTGGCCCGCGTCGG	CCGCTACAAGO	itcaacaaga	AGCTGGGCCTG	AACAC	-CG-ATCATC	CGATC	ACCA	CCACGACGC1	GACCGAR	IGAAGAC GT(CGTCG
M.kansasiiIY	CCTGGCCCGAGTTGG	CCGCTATAAGO	TCAACAAGA	AGCTGGGCTTG	AACAC	-CG-ATCACC	CGATC	ACCA	CGACGACGCT	GACGGAF	IGAAGACGT	CGTCG
M.kansasiill	CETECCECCECCECCE	100000000000000000000000000000000000000	TCOOCOOCO	HUL1UU-LL1L	HHLHL	-CC-OTCOTC	LUHIL CCOTC	HLLH 0CCO	CCOCCOCCCI	GACCCO	COOCOCCT	COTCO
Consensus		ic-ac inclinat	TCAACAAGA	ACTEGGECTE	aacge		Ceat.caC	acct.	CeACeC1	GACCGA	GAaGACGTI	CGTcG
							0					
	444 400	400	4.40	450	400	470		00 4	0 0 (00	94.6	94.0
	111 120	130	140	150	160	170	1	80 1 -+	90 2	200	210	216 1
M.gordonaeI\	111 120 + / ccaccatcgagta	130 ICCTGGTCCGCC	140 	150 	160 ••• 66tca6c	170 AcacgAt	1 Gaccgtt	80 1 • •	90 2 C <mark>CGAGGTTCC</mark>	:00 :GGTGGAC	210 a	216 1 CAT
H₊gordonael\ H₊gordonael	111 120 I CCACCATCGAGTA I CCACCATCGAGTA	130 CCTGGTCCGCC	140 TCCAC	150 GAG	160 GGTCAGC GGCCAGC	170 AcacgAt AcacgAt	1 Gaccgtt Gaccgtc	80 1 -+	90 2 CCGAGGTTCC CCGAGGTGCC	:00 : GGTGGA G :G GTTGA G	210 a	216 1 CAT CAT
H.gordonael\ H.gordonael H.gordonaelII H.gordonaelII	111 120 I I CCACCATCGAGTA I CCACCATCGAGTA I CCACCATCGAGTA I CCACCATCGAGTA I CCACCATCGAGTA I	130 ICCT66TCC6C0 ICCT66TCC6C0 ICCT66TCC6TC	140 	150 GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC	170 ACACGAT ACACGAT ACACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTT GACCGTT	80 1 CCGGGCGGGA CCGGGCGGCA CCGGGCGGCA CCCGGCGGCGGCA	90 2 CCGAGGTTCC CCGAGGTCC CCGAGGTCC TCGAGGTCC	:00 :GGTGGA0 :GGTTGA0 :GGTGGA0 :GGTGGA0	210 ACCGACGAC GACCGACGAC GACCGACGACGAC GACCGACGACGACGACGACGACGACGACGACGACGACGAC	216 1 CAT CAT CAT CAT
K.gordonael\ H.gordonael M.gordonaelII M.aviu M.aviu	111 120 1	130 CCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGTC ICCTGGTGCGCCC	140 	150 GAG GAG GAG 	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGTCAGC GGCCAGC	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTT GACCGTC GACCGTC	80 1 	90 2 CCGAGGTTCC CCGAGGTCC CCGAGGTCC TCGAGGTGCC TCGAGGTGCC	:00 :GGTGGAG :GGTTGAG :GGTGGAG :GGTGGAG	210 ACCGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGACGACGACGACGACGACGACGACGACGACG	216 1 CAT CAT CAT CAT CAT
N₊gordonael\ N₊gordonael N₊gordonaeIII N₊gordonaeIII N₊utracellulare M₊gordonaell	111 120 I I CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGCCCCCCCCCC	140 :TCCAC :TGCAC :TGCAC :TGCAC :TGCAC	150 GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGC GGTCAGT	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CGGCGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACGGTT	80 1 CCGGGCGGGA CCGGGCGGCA CCGGGCGGCA CCCGGCGGCA CCCGGCGGCGCA	90 2 CCGAGGTTCC CCGAGGTGCC CCGAGGTGCC TCGAGGTGCC TCGAGGTGCC CCGAGGTGCC	:00 :GGTTGGAG :GGTTGAG :GGTGGAG :GGTGGAG :GGTGGAG	210 GACCGACGAC GACCGACGACGAC GACCGACGACGACGACGACGACGACGACGACGACGACGAC	216 1 Cat Cat Cat Cat Cat Cat
H.gordonael H.gordonael H.gordonaelII H.aviu H.intracellularc H.gordonaelI 	111 120 12 CCACCATCGAGTA 1 CCACCATCGAGTA 1 CCACCATCGAGTA 2 CCACCATCGAGTA	130 CCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGTC ICCTGGTGCGCC ICCTGGTGCGGC ICCTGGTCCGGC	140 TCCAC TGCAC TGCAC TGCAC TGCAC TGCAC	150 GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGT GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT	80 1 CCGGGCGGGA CCGGGCGGCA CCCGGCGGCA CCCGGCGGCA CCCGGCGGCA CCCGGCGGCA	90 2 CCGAGGTTCC CCGAGGTGCC CCGAGGTGCC TCGAGGTGCC CCGAGGTGCC CCGAGGTGCC	:00 GGTTGGAG GGTTGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG	210 ACCGACGAI ACCGACGACGAC ACCGACGACGACGACGACGACGACGACGACGACGACGACG	216 I CAT CAT CAT CAT CAT CAT CAT
M.gordonael\ M.gordonael M.gordonaell M.aviun M.intracellulare M.gordonaell M.szulgai M.ulcerams M.ulcerams	111 120 I	130 CCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC	150 GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGTCAGT GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGCT	80 1 CCGGGCGGGA CCGGCGGCGCA CCGGCGGCGCA CCCGGCGGCGCA CCCGGCGGCGCA CCCGGCGGCGCA CCCGGCGGCGCA	90 2 CCGAGGTTCC CCGAGGTCCC CCGAGGTCCC TCGAGGTGCC CCGAGGTGCC CCGAGGTGCC TCGAGGTGCC TCGAGGTGCC	COO CGTTGAA CGTTGAA CGTGGAA CGTGGAA CGTGGAA CGTGGAA CGTCGAA CGTCGAA CGTCGAA	210 A ACCGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC	216 1 CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonael M.aviu M.intracellular M.gordonael M.szulga M.ulcerans M.marinu M.pereginu M.gereginu	111 120 I I CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCCCCCC	140 TCCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC	150 GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCGCGAT CCCCGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTT GACCGCT GACCGCT GACCGCT	80 1 CC666C666A CC666C66CA CC666C66CA CCC66C66CA CCC66C66C6 CCC66C66C6 CCC66C66T6 CCC66C66T6	90 2 CCGAGGTTCC CCGAGGTGCC CCGAGGTGCC TCGAGGTGCC CCGAGGTGCC CCGAGGTGCC TCGAGGTGCC TCGAGGTGCC	COO CGTTGAA CGTTGAA CGTTGAA CGTGGAA CGTGGAA CGTGGAA CGTCGAA CGTCGAA CGTCGAA CGTCGAA	210 ACCGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonaell M.aviu M.intracellular M.gordonaell M.szulgai H.ulcerans M.marinu M.peregrinu M.fortuitu	111 120 I I CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCC ICCTGGTGCGGC ICCTGGTGCGGC ICCTGGTCCGCT ICCTGGTCCGCT ICCTGGTGCCGCT	140 TCCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC	150 GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCGCGAT CCGCGAT CCCCGAT	A GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGCT GACCGCT GACCGTC GACCGTC	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC	:00 :GGTGGAG :GGTGGAG :GGTGGAG :GGTGGAG :GGTGGAG :GGTCGAG :GGTCGAG :GGTCGAG :GGTCGAG	210 A ACCGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGACGACGACGACGACGACGACGACGACGACG	216 I CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonaell M.aviu M.intracellulare M.gordonaell M.gordonaell M.szulgai M.ulcerans M.marinu M.fortuitun M.fortuitunII	111 120 I I CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGGC ICCTGGTGCGGC ICCTGGTCCGCI ICCTGGTCCGCI ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGCCCC	140 TCCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCGCGAT CCGCGAT CCCCGAT CCACGAT	A GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGCT GACCGCT GACCGCC GACCGTC GACCGTC	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC	GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG	210 ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGACGACGACGACGACGACGACGACGACGACG	216 I CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonaell M.aviur M.intracellulard M.gordonaell M.gordonaell M.szulgai M.ulcerans M.warinur M.peregrinur M.fortuitur M.fortuitur M.fortuitur	111 120 I I CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCC ICCTGGTGCGGC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC	140 TCCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACRCGAT ACRCGAT ACRCGAT CCRCGAT CCRCGAT CCGCGAT CCGCGAT CCGCGAT CCCCGAT CCCCGAT CCCCGAT CCCCGAT	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC	COO GGTGGAC GGTGAC GGTGAC GGTGAC GGTGAC GGTCAC GGTCAC GGTCAC GGTCAC GGTCAC GGTCAC	210 ; CCCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGACGACGACGACGACGACGACGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonael M.aviur M.intracellularc M.gordonael M.szulgai M.ulcerans M.ulcerans M.harinur M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.genavensc M.simiaa	111 120 I I I CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGACTA CCACCATCGAGTA CCACCATCGACTA CCACCATCGAGTA CCACCATCGAGTA CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCCCCCCC	140 TCCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC TGCRC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGTCAGC GGTCAGC GGTCAGC GGTCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGC GGCCAGC	170 ACRCGAT ACRCGAT ACRCGAT CCACGAT CCACGAT CCACGAT CCCCGAT CCCCGAT CCCCGAT CCACGAT CCACGAT CGACGAT CGACGAT	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGCC GACCGTC GACCGTC GACCGTC GACCGTT	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC	COO GGTGGAC GGTGGAC GGTGGAC GGTGGAC GGTGGAC GGTCGAC GGTCGAC GGTCGAC GGTCGAC GGTCGAC GGTCGAC GGTCGAC	210 ; accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac accgacgac	216 CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonael M.aviu M.intracellulare M.gordonael M.szulgai M.ulcerans M.narinu M.peregrinu M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.siniae M.siniae	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCCCCCCC	140 TCCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC	150 	160 GGTCAGC GGCCAGC GGTCAGC GGCCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACRCGAT ACRCGAT ACRCGAT CCACGAT CCACGAT CCACGAT CCCCGAT CCCCGAT CCACGAT CCACGAT	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGCT GACCGCT GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT	80 1 CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCG CCCGGCGCGCG CCCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCCG CCGGCGCGCCG	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC	COO CGGTGGAG CGTGGAG CGTGGAG CGTGGAG CGTGGAG CGTCGAG CGTCGAG CGTCGAG CGTCGAG CGTCGAG CGTCGAG CGTCGAG CGTGGAF CGTGGAF CGTGGAF	210 ACCGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGACGACGACGACGACGACGACGACGACGACG	216 CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.szulgai M.szulgai M.szulgai M.szulgai M.srainu M.fortuitun	111 120 I	130 CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTGCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTGCGCCCC CCTGGTGCGCCCCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC	140 TCCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT	1 GACCGTC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCC GACCGCT GACCGTT GACCGTT GACCGTT GACCGTT	80 1 CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGCGCGCA CCCGCGCGCA CCCGCGCGCA CCCGGCGCGCA CCCGGCGCGCG CCCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG	90 2 CCGAGGTTCC CCGAGGTCCC CCGAGGTCCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTCCC TCGAGGTCCC TCGAGGTCCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC	COO CGGTGGAG CGGTGAG CGGTGGAG CGGTGGAG CGGTGGAG CGGTCGAG CGGTCGAG CGGTCGAG CGGTCGAG CGGTGGAG CGGTGGAG CGGTGGAG CGGTGGAG CGGTGGAG	210 CACCGACGACGAC CACCGACGACGACGACGACGACGACGACGACGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.szulga M.ulcerans M.marinu M.geregrinur M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.genavensc M.siniaa M.tuberculosis M.hovis M.bovis	111 120 I I CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGCCCCC ICCTGGTGCCGCC ITCTGGTCCGCC ITCTGGTCCGCC ITCTGGTCCGCC ITCTGGTCCGCC	140 TCCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGTCAGA GGTCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT	A GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT	80 1 CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCG CCCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCCG CCGGCGCGCCG CCGGCCGCCG CCGGCCGCCG CCGGCCGCCG	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRCC	200 	210 CACCGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonael M.aviu M.intracellular M.gordonael M.szulgai M.szulgai M.szulgai M.szulgai M.szulgai M.szulgai M.szulgai M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.sulgai	111 120 I I CCACCATCGAGTA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ITCTGGTCCGCC ITCTGGTCCGCC ITCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGTCAGA GGTCAGA GGCCACA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTC GACCGTC GACCGTC	80 1 CCGGCGCGCA CCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCG CCCGGCGCGCG CCCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCGC	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC	COO GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG	210 CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonaell M.gordonaell M.aviu M.intracellulard M.gordonaell M.szulgai M.szulgai M.sulgai M.sulgai M.seregrinu M.fortuitunII M.genavensc M.siniae M.tuberculosis M.microti M.tuberculosis M.microti M.tuberculosis M.africanu M.celatunII M.celatunII M.kansasii	111 120 I I I CCACCATCGAGTA <	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGCCCCC ITCTGGTCCGCC ITCTGGTCCGCC ITCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGCCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGTCAGA GGTCAGA GGTCAGA GGCCACG GGCCACAG	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTG GACCGTG	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC	200 	210 CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC CACCGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonael M.gordonael M.gordonaell M.aviun M.intracellulard M.gordonaell M.gordonaell M.szulgai M.ulcerans M.sulgai M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.genavensc M.siniad M.siniad M.kansasii M.kansasii M.kansasii	111 120 I I I CCACCATCGAGTA <	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGCCGCC ICCTGGTGCGGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCC ICCTGGTGCGCCCCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCACA GGCCACA GGCCACA	170 	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTC GACCGTC GACCGTA GACCGTA GACCGTA GACCGTA	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTGCC TCGRGGTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRGCTCC TCGRC	200 	210 ACCGACGACGAC	216 I CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT
M.gordonaeI M.gordonaeI M.gordonaeI M.gordonaeI M.gordonaeI M.szulgai M.szulgai M.szulgai M.narinu M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.siniae M.siniae M.siniae M.siniae M.siniae M.siniae M.siniae M.siniae M.sassii M.haenophilum	111 120 I I CCACCATCGAGTA CCA	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC	140 TCCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 GGTCAGC GGCCAGC GGCCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGTCAGA GGTCAGA GGTCAGA GGCCACA GGCCACA GGCCACA GGCCACA GGCCACA GGCCACA GGCCACA GGCCACA	170 	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTG GACCGTA GACCGTA GACCGTA GACCGTA	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGTGCC TCGRG TCGRG TCGRG TCGR	200 GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG GGTCGAG	210 ; accgacgac accg	216
M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.szulgai M.ulcerans M.sazilgai M.ulcerans M.szulgai M.sulgai M.sortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.siniae M.tuberculosis M.siniae M.tuberculosis M.siniae M.tuberculosis M.siniae M.tuberculosis M.ficanu M.selatum M.kansasii M.kansasii M.kansasii M.kansasii	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCC ICCTGGTGCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC	140 TCCAC TGCAC	150 GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG	160 	170 ACACGAT ACACGAT ACACGAT CCACGAT	GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTG GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC	80 1 CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGCGCGCA CCCGCGCGCA CCCGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCA CCGGCGCA CCGGCGCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCCA CCGGCGCA CCGGCGCCA CCGGCCGCCA CCGGCGCCA CCGGCCGCCA CCGGCGCCA CCGGCCGCCA CCGGCCGCCA CCGGCCGCA CCGGCCGCA CCGGCGCCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGCCA CCGGCCA CCGCA CCGCA CCGCA CCGCA CCGCCA CCGC	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC	200 GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG GGTCGAG	210 ACCGACGACGAC ACCGACGAC ACCGACGACGAC ACCGACGACGAC ACC	216
M.gordonael M.gordonael M.gordonael M.gordonaell M.aviu M.intracellulare M.gordonael M.szulgai M.ulcerans M.szulgai M.sulcerans M.sarinu M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.selatum M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii	111 120 I	130 CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTGCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC	140 TCCAC TGCAC	150 GAG 	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGTCAGA GGCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTC GACCGTC GACCGTC	80 1 CCGGGCGGGA CCGGGCGGCA CCGGGCGGCA CCGGCGCGCA CCCGGCGGCA CCGGGCGGCA CCGGGCGGCA CCGGCGGGGG CCGGGCGGCA CCGGGCGGCA CCGGGCGGCA CCGGGCGGCA CCGGGCGGCA CCGGCGGCA CCGGGCGGCA CCGGCA CCGGCA	90 2 CCGAGGTCCC CCGAGGTCCC CCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTCCC TCGAGGTCCC TCGAGGTGCC TCGAGGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAC TCGAC TCGAGTCC TCGAC	200 	210 ACCGACGACGAC	216
M.gordonael M.gordonael M.gordonaell M.gordonaell M.aviu M.intracellulare M.gordonaell M.szulgai M.szulgai M.sulcerans M.marinu M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.selatun M.kansasiil M.haenophilu M.senegnatis M.malmoenss M.schelonae	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ITCTGGTCCGCC ITCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC TGCAC	150 GAG	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGCT GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT	80 1 CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCGCA CCCGCCGCGCGCA CCGGCCGCGCCA CCGGCCGCGCCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGCCGCCA CCCGCGCGCCA CCCGCGCGCCA CCCGCGCGCCA CCCGCGCGCCA	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGCTCCC TCGRC	200 	210 ;; CACCGACGACGACGACGACGACGACGACGACGACGACGAC	216
M.gordonael M.gordonael M.gordonael M.gordonael M.aviu M.intracellular M.gordonael M.szulga M.szulga M.szulga M.szulga M.szulga M.szulga M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.fortuitu M.ganavessi M.africanu M.celatum M.kansasii M.haenophilu M.snegmatis M.nalnoensa M.chelonaa M.abscessu	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCC ICCTGGTGCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC	140 TCCAC	150 GAG 	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGTT GACCGCT GACCGCT	80 1 CCGGCGCGCA CCGGCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCGCA CCCGCCGCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCCA CCCGCCCCCCA CCCGCCCCCCA CCCGCCCCCCA CCCGCCCCCCA CCCGCCCCCCA CCCGCCCCCA CCCGCCCCCCA CCCGCCCCCCA CCCGCCCCCCCCCC	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGCTCCC TCGRCCCC TCGRCCCC TCGRCCCC TCGRCCCC TCGRCCCC TCGRCCCC TCGRCCCC TCGRCCCCC TCGRCCCC TCGRCCCC TCGRCCCCC TCGRCCCCC TCGRCCCCC TCGRCCCCC TCGRCCCCCC TCGRCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	COO GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG	210 ACCGACGACGACGAC ACCGACGACGACGAC ACCGACGACGACGACGAC ACCGACGACGACGACGACGACGACGACGACGACGACGACG	216
M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.aviu M.intracellular M.szulgai M.szulgai M.szulgai M.szulgai M.szulgai M.szulgai M.fortuitun M.fortu	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGCC ICCTGGTGCGGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCGCCCCCC ITCTGGTCCGCCCCCCCCCC	140 TCCAC	150 GAG 	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACGAT	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTG GACCGTG GACCGTG GACCGTG GACCGTC GACCGCT GACCGCC GACCGTC GACCGCC GACCGTC GACCGCC GACCGTC GACCGCCC GACCGCCC GACCGTC	80 1 CCGGCGCGCA CCGGCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCGCA CCCGCCGCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCCA CCCGCCCCCCA CCCGCCCCCA CCCGCCCCA CCCGCCCCCA CCCGCCCCCCCCCA CCCGCCCCCA CCCGCCCCCCCCCC	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGTCCC TCGRC	200 	210 ACCGACGACGAC ACCGACGACCACAC ACCGACGACCACACACACACACACACACACACACACACA	216
M.gordonaeli M.gordonaeli M.gordonaeli M.aviu M.intracellulare M.gordonaeli M.szulgai M.ucerans M.szulgai M.ucerans M.striat M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.sinias M.africanu M.kansasii M.haenophilum M.senegmatis M.malmoenss M.kenopi M.chelonaa M.ternaa M.ternaa	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGGC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCC ITCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCCCCC ICCTGGTGCCGCCCCCCCCCC	140 TCCAC	150 GAG 	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGC GGCCAGA	170 	1 GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC	80 1 	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTGCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGGTCC TCGRGCTC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRGTCC TCGRC	200 	210 ACCGACGACGACGAC ACCGACGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGAC ACCGACGACGACCGACGAC ACCGACGACGACGAC ACCGACGACGACGAC ACCGACGACGACCACGACGA	216
M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.szulgai M.ulcerans M.szulgai M.ulcerans M.szulgai M.sulgai M.sulcerans M.sulgai M.sulcerans M.siniae M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.siniae M.siniae M.sengai M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansasii M.kansenji M.sengmatis M.sengmatis M.sengmatis M.sengalae M.serofulaceu M.serofulaceu	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCC ICCTGGTGCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCC ICCTGGTGCCGCCCCCC ICCTGGTGCCGCCCCCCCCCC	140 TCCAC TGCAC TT	150 GAG	160 	170 ACACGAT ACACGAT ACACGAT CCACG-	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTG GACCGTG GACCGTG GACCGTG GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTGTC GACCGTGTC GACCGTGTC GACCGTGTC	80 1 CCGGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGCGCGCA CCCGCGCGCA CCCGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGCCGCA CCCGCCA CCCGCCA CCCGCCCA CCCGCCCA CCCGCCA CCCGCCA CCCGCCCA CCCGCCCA CCCGCCA CCCCCCCA CCCGCCA CCCGCCA CCCGCCA CCCCCCCA CCCGCCCA CCCGCCA C	90 2 CCGAGGTCC CCGAGGTCC CCGAGGTCC CCGAGGTCC CCGAGGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGTCC TCGAGGTCC TCGAGGTCC	200 GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG	210 CACCGACGACGAC CACCGACGACGACGACGACGACGACGACGACGACGACGAC	216
M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.gordonael M.szulgai M.ulcerans M.szulgai M.sulgai M.sulgai M.sulgai M.siniae M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.fortuitum M.senavenss M.siniae M.tuberculosis M.siniae M.sulgai M.sovis M.africanum M.celatum M.kansasii M.kansasii M.kansephilu M.haenophilu M.sheephilu M.sheephilu M.sheescens M.flavescens M.gastri M.kansasiiT	111 120 I	130 CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTCCGCC CCTGGTGCCGCC CCTGGTGCCGCC CCTGGTCCGCCCCCC CCTGGTCCGCCCCCCCC	140 TCCAC TGCAC TGCACCACC	150 GAG	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACG-	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGCC GACCGTC	80 1 CCGGGCGCGCA CCGGGCGCGCA CCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCA CCCGGCGCGCG CCCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCGGCGCGCG CCCGGCGCGC CCCGGCGCGC CCCGCGCGCG	90 2 CCGAGGTCCC CCGAGGTCCC CCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTGCC TCGAGGTCCC TCGAGGTCCC TCGAGGTCCC TCGAGGTGCC TCGAGGTCC TCGAGCTC TCGAGTCC TCGAGTCC TCGAGCTC TCGAGTCC TCGAGTCC TCGAC TCGAC TCGAGTCC TCGAGTCC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCGAC TCC TCGAC TCC TCGAC TCC TCGAC TCC TCC TCC TCC TCC TCC TCC T	200 	210 ACCGACGACGAC ACCGACGAC ACCGACGACGAC	216
M.gordonael M.gordonael M.gordonaell M.gordonaell M.gordonaell M.gordonaell M.gordonaell M.gordonaell M.gordonael M.gordonael M.gordonael M.gordonael M.genavenso	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTGCGCCCCC ICCTGGTGCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC TGCAC	150 GAG 	160 GGTCAGC GGCCAGC GGTCAGC GGTCAGC GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACG-	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTT GACCGTT GACCGTT GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC	80 1 CCGGCGCGCA CCGGCGCGCA CCGGCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCGCA CCCGCCGCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCGCGCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGGCCA CCGCA CCGGCCA CCGCA CCGCA CCGCA CCGCA CCGCA CCGCA CCGGCCA CCGCA	90 2 CCGRGGTCCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTCC TCGRGCTCC TCGRC	000 GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG	210 ;; CACCGACGACGACGACGACGACGACGACGACGACGACGAC	216
M.gordonaell M.gordonaell M.gordonaell M.gordonaell M.aviuu M.intracellulare M.gordonaell M.gordonaell M.gordonaell M.gordonaell M.szulgai M.szulgai M.sucerans M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.fortuitun M.selatun M.kansasiil M.kansasiil M.serofulaceu M.gastri M.kansasiill M.kansasiill M.kansasiill M.kansasiill	111 120 I	130 ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ITCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCC ICCTGGTCCGCCCCCCCCCC	140 TCCAC	150 GAG 	160 GGCCAGC GGCCAGC GGCCAGC GGCCAGA	170 ACACGAT ACACGAT ACACGAT CCACG-	1 GACCGTT GACCGTC GACCGTC GACCGTC GACCGTC GACCGTC GACCGCT GACCGCT GACCGCT GACCGTT GACC	80 1 CCGGCGCGCA CCGGCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCCGCCGCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCGGCCGCA CCCGCCCGCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCA CCCGCCCCCCCCCC	90 2 CCGRGGTGCC CCGRGGTGCC CCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTGCC TCGRGGTCCC TCGRGGTCCC TCGRGGTCCC TCGRGGTGCC TCGRGGTCC TCGRGTCC TCGRC TC	000 GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTGGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTCGAG GGTGGAG	210 ;; accgacgacgac accgacgac accgacgacgac accgacga	216

FIG. 1. Multialignment sequences identified in the *rpoB* region of 35 strains of 26 mycobacterial species of clinical importance. The software used for the alignment of multiple sequences was obtained from reference 2. Sequences with high consensus are shown as red letters, low consensus as blue letters, and neutral consensus as black letters. Among the 360-bp sequences identified in this study, only the 216-bp region of the upstream sequences is shown, since the downstream sequences have been reported elsewhere. The sequences shown here have been submitted to GenBank.

Name of oligonucleotide	Sequence	Target organism(s)		
МҮС	GACGTCGTCGCCACCATCGA	All mycobacterial species		
MTB	CATGTCGGCGAGCCC	M. tuberculosis complex		
AVI	CGGTGAGCCGATCACC	M. avium		
INT	CCTGCACGCGGGCGA	M. intracellulare		
SCR	CGTACGGATGGCCAGC	M. scrofulaceum		
KAN-I	GGCCACGATGACCGTG	M. kansasii types I and V		
KAN-II	TCTCAGGATGGCCAGC	M. kansasii types II, III, and IV		
GAS	TCTCAGGGTGGCCAGG	M. gastri		
FOR-C	CCTGAACGCCGGCCAG	M. fortuitum complex		
PER	GTTCCGGTCGAGGTGG	M. peregrinum		
CHE	TGGTGACTGCCACCACG	M. chelonae		
ABS	GGTGACCACCACC	M. abscessus		
ULC	GGCCAGCCCATCACC	M. ulcerans		
GEN/SIM	CCAGCCGACGATGACG	M. genavense-M. simiae		
GOR-I	GTCGGCGATCCGATCA	<i>M. gordonae</i> types I, III, and IV		
GOR-II	CGTCGGCAAGCCGA	M. gordonae type II		
SZU	TCTGAACGTCGGCGAG	M. szulgai		

TABLE 2. Mycobacterial species-specific oligonucleotide probes designed and confirmed for their specificity by dot blot hybridization in this study

Based on the sequence information, we designed 15- to 20-mer oligonucleotide molecules as species- and subtype-specific probes for 16 strains of 14 mycobacterial species of clinical importance and one for all the mycobacterial species (Table 2). Subsequently, in order to determine the specificity of the probes for the targeted mycobacterial species, dot blot hybridization was carried out using each probe molecule. For these experiments, the target region of the *rpoB* gene from 48 different mycobacterial strains, representing 39 species listed in Table 1, was amplified by PCR, and the PCR products were then dot blotted on membranes.

To prepare the DNA dot blot, precut membrane (Hybond-N⁺ [10 by 10 cm]; Amersham Pharmacia Biotech Korea Ltd., Seoul, Korea) was immersed into the denaturing solution (0.4 N NaOH, 25 mM EDTA; pH 8.0) for 1 min. After excess denaturing solution was allowed to drip from the membrane, it was placed on Whatman 3MM filter paper, and 1 to 2 µl of the PCR product was blotted onto the membrane. The membrane was then air dried for 5 min, rinsed with denaturing solution for another 1 min, placed between two sheets of 3MM filter paper, and baked for 2 h at 80°C. Oligonucleotide probes were labeled using an enhanced chemiluminescence kit for 3' oligolabeling and detection (Amersham Pharmacia Biotech Korea Ltd.). Subsequent processing, including hybridizing the oligonucleotide probes to the membrane (42°C for 1 h), membrane washing (52°C for 15 min), and signal detection, was carried out using the method recommended by the manufacturer. Each species-specific probe was tested separately using the membrane onto which all PCR products were blotted, and there was no difference in dot blot hybridization conditions for all 16 probes analyzed in this study.

Figure 2 showed dot blot hybridization results using the *M. tuberculosis* complex probe and *M. gastri* probe as examples. Each probe hybridized only to the corresponding mycobacterial species, indicating the specificity of molecular probes to each mycobacterial species. The rest of the probes in Table 2 also showed species-specific hybridization to the corresponding species (data not shown).

Finally, in order to determine if all clinical isolates that belong to each mycobacterial species can be detected using its specific probe, dot blot hybridization using the *M. avium*-probe molecule was carried out using 36 clinical isolates, including 6 isolates of *M. avium*. As shown in Fig. 3, all six *M. avium* isolates were identified correctly by the *M. avium*-specific probe in a dot blot hybridization assay, while the probe did not bind to any clinical isolates of other mycobacterial species.

This study demonstrates that the *rpoB* gene of *Mycobacterium* species contains a highly polymorphic region whose DNA sequences can be used for species identification. Sequence analysis of the region showed clearly the difference in nucleotide sequences among 26 *Mycobacterium* species and subtypes of four species examined in this study. The results also supported clearly our previous report on species identification of mycobacteria by RFLP of the polymorphic region (13). In addition, we showed that the oligonucleotide probes based on the sequences of the region were specific to each *Mycobacterium* species and useful for species identification of mycobacteria in a dot blot hybridization.

The *rpoB* gene encodes the β subunit of RNA polymerase, which produces RNA molecules in cells. Thus, rpoB is one of the very critical housekeeping genes that are closely related to cellular vitality and thus becomes the target for rifampin, the major bactericidal drug for *M. tuberculosis* and *M. leprae*. It is, therefore, reasonable to assume that the genetic structure of the rpoB gene is highly conserved within the same species. However, unlike 16S rRNA or any other rRNA whose primary structure is functionally critical, the *rpoB* gene seems to tolerate a more diverse sequence alteration without causing any changes in protein function. In particular, the DNA region that is not involved in the active site of the protein seems to be more polymorphic and does not cause major functional defects. Based on these relationships, it is easily understood that there exist highly conserved DNA regions and relatively variable DNA regions in the *rpoB* gene. The tolerable sequence variation in the rpoB gene becomes a useful clue for species identification of mycobacteria as reported previously (9, 10, 13). However, sequence analysis of the 360-bp region of the rpoB that we reported herein clearly revealed more extensive variation than expected, leading to develop mycobacterial species-specific probe molecules.



M. tuberculosis complex probe

M. gastri probe

FIG. 2. The results of the dot blot hybridizations using each mycobacterial species-specific oligomer probe derived from the novel region of the *rpoB* genes of species. Dot blot hybridizations were conducted using probes specific for *M. tuberculosis* complex and *M. gastri*. The PCR-amplified products from 48 mycobacterial species were blotted on the membrane, and this was followed by hybridization with probes. The identification numbers in the membranes are matched with numbers and species names in Table 1.

The oligonucleotide probes designed on the basis of this polymorphic region were useful in *Mycobacterium* species identification by dot blot hybridization assay. No cross-reactive hybridization was found between the 16 mycobacterial species-



FIG. 3. Dot blot hybridizations using the *M. avium* specific oligomer probe. The membrane contained PCR products amplified from 36 clinical isolates of mycobacteria, including six isolates of *M. avium* species. The clinical isolates were identified by conventional culture and biochemical tests at the KIT.

specific probes (Table 2) and 48 strains of 39 Mycobacterium species (Fig. 2). For example, the M. gastri-specific probe did not hybridize with five subtypes of M. kansasii, although the two species could not be differentiated by 16 rRNA sequence analysis (18). One of our concerns was sequence variation in the 360-bp region among clinical isolates of each Mycobacterium species. In this study, however, there seems to be no variation in the nucleotides of the region among M. avium clinical isolates, because the probe hybridized with all M. avium clinical isolates, as shown in Fig. 3. This was also supported by our previous study in which no variation was found in RFLP enzyme restriction sites among 40 clinical isolates of M. tuberculosis, 40 clinical isolates of M. avium, 50 clinical isolates of M. intracellulare, and 25 clinical isolates of M. gordonae, etc. (13). Although other probes still need to be confirmed for their sensitivity and specificity using multiple clinical isolates of each Mycobacterium species, the probes will be useful in developing a reverse blot hybridization assay by which many isolates can be analyzed for their species identification at the same time. In addition, since the 360-bp region is located near the *rpoB* mutation sites, which are associated with resistance to rifampin, one can develop an assay in the future which can simultaneously provide information about mycobacterial species identity and rifampin resistance.

Nucleotide sequence accession numbers. The nucleotide sequences listed in Table 2 have been submitted to the EMBL database and have been given the following accession num-

bers: AY271315 for M. microti, AY271316 for M. terrae, AY271317 for M. scrofulaceum, AY271318 for M. marinum, AY271319 for M. szulgai, AY271320 for M. gastri, AY271321 for M. malmoense, AY271322 for M. avium, AY271323 for M. bovis, AY271324 for M. peregrinum, AY271325 for M. fortuitumI, AY271326 for M. celatum type II, AY271327 for M. flavescens, AY271328 for M. intracellulare, AY271329 for M. abscessus, AY271330 for M. africanum, AY271331 for M. haemophilum, AY271332 for M. xenopi, AY271333 for M. kansasii type I, AY271334 for M. kasasii type II, AY271335 for M. kansaii type IV, AY271336 for M. kansasii type IV, AY271336 for M kansaii AY271337 M. celatum type I, AY271338 for M. genavense, AY271339 for M. simiae, AY271340 for M. fortuitum type II, AY271341 for M. gordonae type IV, AY271342 for M. gordonae type I, AY271343 for M. gordonae type II, AY271344 for M. gordonae type III, and AY271345 for M. smegmatis.

This work was supported in part by National Research Laboratory Program M1-0001-00-0089, by the Ministry of Science and Technology National Research and Development Program, and by Brain Korea 21 Project for Medical Sciences in Yonsei University.

REFERENCES

- Butler, W. R., K. C. Jost, Jr., and J. O. Kilburn. 1991. Identification of mycobacteria by high-performance liquid chromatography. J. Clin. Microbiol. 29:2468–2472.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16:10881–10890.
- Evans, K. D., A. S. Nakasone, P. A. Sutherland, L. M. de la Maza, and E. M. Peterson. 1992. Identification of *Mycobacterium tuberculosis* and *Mycobacterium avium-M. intracellulare* directly from primary BACTEC cultures by using acridinium-ester-labeled DNA probes. J. Clin. Microbiol. 30:2427–2431.
- French, A. L., D. A. Benator, and F. M. Gordin. 1997. Nontuberculous mycobacterial infections. Med. Clin. N. Am. 81:361–379.
- Gingeras, T. R., G. Ghandour, E. Wang, A. Berno, P. M. Small, F. Drobniewski, et al. 1998. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. Genome Res. 8:435–448.
- Goto, M., S. Oka, K. Okuzumi, S. Kimura, and K. Shimada. 1991. Evaluation of acridinium-ester-labeled DNA probes for identification of *Mycobacterium tuberculosis* and *Mycobacterium avium-Mycobacterium intracellulare* complex in culture. J. Clin. Microbiol. 29:2473–2476.
- Horsburgh, C. R. 1991. Mycobacterium avium complex infection in the acquired immunodeficiency syndrome. N. Engl. J. Med. 324:1332–1338.
- 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.
- Kim, B.-J., K.-H. Lee, B.-N. Park, S.-J. Kim, G.-H. Bai, S.-J. Kim, and Y.-H. Kook. 2001. Differentiation of mycobacterial species by PCR-restriction analysis of DNA (342 base pairs) of the RNA polymerase gene (*rpoB*). J. Clin. Microbiol. 39:2102–2109.

- Kim, B.-J., S.-H. Lee, M.-A. Lyu, S.-J. Kim, G.-H. Bai, S.-J. Kim, G.-T. Chae, E.-C. Kim, C.-Y. Cha, and Y.-H. Kook. 1999. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). J. Clin. Microbiol. 37:1714–1720.
- Kirschner, P., B. Springer, U. Vogel, A. Meier, A. Wrede, M. Kiekenbeck, F.-C. Bange, and E. C. Böttger. 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.
- 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.
- Lee, H., H.-J. Park, S.-N. Cho, G.-H. Bai, and S.-J. Kim. 2000. Species identification of mycobacteria by PCR-restriction fragment length polymorphism of the *rpoB* gene. J. Clin. Microbiol. 38:2966–2971.
- Marks, J., and T. Szulga. 1965. Thin-layer chromatography of mycobacterial lipids as an aid to classification; technical procedures; *Mycobacterium fortuitum*. Tubercle 46:400–411.
- Pai, S., N. Esen, X. Pan, and J. M. Musser. 1997. Routine rapid *Mycobac*terium species assignment based on species-specific allelic variation in the 65-kilodalton heat shock protein gene (hsp65). Arch. Pathol. Lab. Med. 121:859–864.
- Plikaytis, B. B., B. D. Plikaytis, M. A. Yakrus, 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.
- Reisner, B. S., A. M. Gatson, and G. L. Woods. 1994. Use of Gen-Probe AccuProbes to identify Mycobacterium avium complex, Mycobacterium tuberculosis complex, Mycobacterium kansasii, and Mycobacterium gordonae directly from BACTEC TB broth cultures. J. Clin. Microbiol. 32:2995–2998.
- Rogall, T., T. Flohr, and E. Bottger. 1990. Differentiation of mycobacterial species by direct sequencing of amplified DNA. J. Gen. Microbiol. 136:1915– 1920.
- 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.
- 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.
- Springer, B., P. Kirschner, G. Rost-Mayer, K.-H. Schröder, R. M. Kroppenstedt, and E. C. Böttger. 1993. *Mycobacterium interjectum*, a new species isolated from a patient with chronic lymphadenitis. J. Clin. Microbiol. 31: 3083–3089.
- 22. Springer, B., E. Tortoli, I. Richter, R. Grünewald, S. Rüsch-Gerdes, K. Uschmann, F. Suter, M. D. Collins, R. M. Kroppenstedt, and E. C. Böttger. 1995. *Mycobacterium conspicuum* sp. nov., a new species isolated from patients with disseminated infections. J. Clin. Microbiol. 33:2805–2811.
- Telenti, A., F. Marchesi, M. Balz, F. Baly, 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.
- 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.
- 25. Tsang, A. Y., I. Drupa, M. Goldberg, J. K. McClatchy, and P. J. Brennan. 1983. Use of serology and thin-layer chromatography for the assembly of an authenticated collection of serovars within the *Mycobacterium avium-Mycobacterium intracellulare-Mycobacterium scrofulaceum* complex. Int. J. Syst. Bacteriol. 33:285–292.