rpoB sequence-based identification of *Mycobacterium avium* complex species

Iskandar Ben Salah, Toidi Adékambi,† Didier Raoult and Michel Drancourt

Correspondence

Michel Drancourt
Michel.Drancourt@
medecine.univ-mrs.fr

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UMR, CNRS-IRD 6236, IFR 48 Faculté de Médecine, Université de la Méditerranée, Marseille, France

Received 30 April 2008
Revised 8 September 2008
Accepted 10 September 2008

The *Mycobacterium avium* complex (MAC) comprises slowly growing mycobacteria responsible for opportunistic infections and zoonoses. The ability to speciate MAC isolates in the clinical microbiology laboratory is critical for determining the organism implicated in clinical disease and for epidemiological investigation of the source of infection. Investigation of a 711 bp variable fragment of *rpoB* flanked by the Myco-F/Myco-R primers found a 0.7–5.1 % divergence among MAC reference strains, with *Mycobacterium chimaera* and *Mycobacterium intracellulare* being the most closely related. Using a 0.7 % divergence cut-off, 83 % of 100 clinical isolates, which had been previously identified by phenotypic characteristics and 16S–23S rDNA intergenic spacer (ITS) probing, were identified as *M. avium*, 8 % as *M. intracellulare* and 2 % as *M. chimaera*. The uniqueness of seven isolates, exhibiting <99.3 % *rpoB* sequence similarity with MAC reference strains, was confirmed by 16S rDNA, ITS and *hsp65* sequencing and phylogenetic analyses. Partial *rpoB* gene sequencing using the Myco-F/Myco-R primers permits one-step identification of MAC isolates at the species level and the detection of potentially novel MAC species.

INTRODUCTION

The *Mycobacterium avium* complex (MAC) comprises slowly growing mycobacteria (Tortoli *et al.*, 1999; Turenne *et al.*, 2007; Yajko *et al.*, 1995) responsible for opportunistic infections and zoonoses (Karakousis *et al.*, 2004; Lauzi *et al.*, 2000). *M. avium* has been further divided into *M. avium* subsp. *avium*, *M. avium* subsp. *paratuberculosis* (Lauzi *et al.*, 2000), *M. avium* subsp. *silvaticum* (Thorel *et al.*, 1990) and *M. avium* subsp. *hominissuis* (Mijls *et al.*, 2002). It has recently been shown that *M. avium* subsp. *hominissuis* comprises a heterogeneous group of organisms, with *M. avium* subsp. *avium/silvaticum* and *M. avium* subsp. *paratuberculosis* being two independent, specialized clones (Turenne *et al.*, 2007). *Mycobacterium intracellulare* has been isolated more frequently in HIV-negative patients (Devallois *et al.*, 1996). Two novel species, *Mycobacterium chimaera* (formerly sequevar MAC-A) (Tortoli *et al.*, 2004) and *Mycobacterium colombiense* (formerly sequevar MAC-X) (Murcia *et al.*, 2006), have recently been described (Table 1).

†Present address: Mycobacteriology Laboratory Branch, Division of Tuberculosis Elimination, Centers for Disease Control and Prevention, Atlanta, GA, USA.

Abbreviations: ITS, intergenic spacer; MAC, *Mycobacterium avium* complex; RGM, rapidly growing mycobacteria.

The GenBank/EMBL/DDBJ accession numbers for the *Mycobacterium avium* complex sequences determined in this study are listed in Table 3.

Since phenotypic tests do not easily discriminate between these closely related species and subspecies, molecular identification has been developed. 16S rDNA sequencing has demonstrated low variability when compared with the Lipav2 assay (Lebrun *et al.*, 2005). *M. avium* serotype 2 *SalI*-*EcoRI* DT6 and *SalI*-*BamHI* DT1 fragment probing (Thierry *et al.*, 1993) and sequencing of the *sodA* (Liu *et al.*, 2001), *dnaJ* (Morita *et al.*, 2004), *gyrB* (Kasai *et al.*, 2000), *recA* (Blackwood *et al.*, 2000), the 32 kDa protein gene (Soini *et al.*, 1996) or a combination thereof (Devulder *et al.*, 2005) have been further used to identify MAC species. These studies, however, did not include all of the four MAC species and used restricted panels of reference strains, thus limiting the significance of the data. The *hsp65* gene has been thoroughly investigated (Devallois *et al.*, 1996; Leao *et al.*, 1999; Swanson *et al.*, 1997; Telenti *et al.*, 1993) and sequence analysis of this gene of almost 1600 bp has been used to distinguish between *M. avium* subsp. *avium* and *M. avium* subsp. *paratuberculosis* strains and disclosed six *M. avium* subsp. *hominissuis* sequavars (Turenne *et al.*, 2006). Intergenic spacer (ITS) sequencing (Abed *et al.*, 1995; Barry *et al.*, 1991; Glennon *et al.*, 1994; Roth *et al.*, 1998) further subdivided MAC into 32 sequevars and recently delineated *M. chimaera* (Tortoli *et al.*, 2004) and *M. colombiense* (Murcia *et al.*, 2006).

Herein, we studied partial *rpoB* gene sequencing as a new method for MAC species identification. This single-copy chromosomal gene encoding the bacterial RNA polymerase

Table 1. Specific host and infectious diseases related to seven MAC reference strains

Species	Strain number	Host	Diseases	Year of description
<i>M. avium</i> subsp. <i>avium</i>	ATCC25291 ^T	Bird	Avian tuberculosis (or the like)	1991
<i>M. avium</i> subsp. <i>paratuberculosis</i>	ATCC19698 ^T	Ruminants	Johne's disease	1991
<i>M. avium</i> subsp. <i>silvaticum</i>	ATCC49884 ^T	Bird	Chronic enteritis	1991
<i>M. avium</i> subsp. <i>hominissuis</i>	IWGMT49	Human	Pulmonary abscess, Lady Windermere syndrome	2002
<i>M. intracellulare</i>	ATCC15985	Human	Pulmonary disease	1965
<i>M. chimaera</i>	DSM44623 ^T	Human	Pulmonary disease	2004
<i>M. colombiense</i>	CIP 108962 ^T	Human	Bacteraemia, lymphadenopathy	2006

β -subunit has been previously used for the identification of several bacterial groups, including rapidly growing mycobacteria (RGM) and MAC (Kim *et al.*, 1999). However, only *M. avium* and *M. intracellulare* were included in earlier studies, which were based on the analysis of a 342 bp (Kim *et al.*, 2001) or 705 bp (Gingeras *et al.*, 1998) region of *rpoB* comprising only 8.5–17 % of the entire *rpoB* gene length and did not ensure that the most variable *rpoB* region for identification was targeted.

Herein, we gained the opportunity afforded by genome sequence data to further investigate *rpoB* as a tool for molecular identification of MAC species, i.e. *M. avium*, *M. intracellulare*, *M. chimaera* and *M. colombiense*.

METHODS

Mycobacterial strains. *M. avium* subsp. *avium* ATCC25291^T, *M. avium* subsp. *silvaticum* ATCC49884^T, *M. avium* subsp. *paratuberculosis* ATCC19698^T and *M. intracellulare* ATCC15985 were purchased from ATCC. *M. chimaera* DSM44623^T was purchased from DSMZ. *M. avium* subsp. *hominissuis* IWGMT49 (originally isolated from a pig) was provided by Dick van Soolingen (National Institute of Public Health and the Environment, Bilthoven, The Netherlands). *M. colombiense* CIP 108962^T was purchased from the Institut Pasteur Collection. In this study, 100 MAC clinical isolates collected in our clinical microbiology laboratory in 1997–2007 were also studied (Table 2). Every isolate was inoculated into Middlebrook 7H9 liquid medium and subcultured onto Middlebrook and Cohn 7H10 agar (Becton Dickinson) at 30 °C. Clinical isolates coated on beads were inactivated as previously described (Djelouagji & Drancourt, 2006), and the DNA extracted using a Qiagen kit was used as the template for PCR amplification of the *rpoB*, 16S rDNA, ITS and *hsp65* genes.

Construction of a MAC partial *rpoB* gene database. Consensus PCR primers were designed after alignment of *rpoB* gene sequences of *M. avium* subsp. *paratuberculosis* K-10 (GenBank accession no. NC-002944), *M. avium* subsp. *hominissuis* strain 104 (GenBank accession no. NC-008595) and *Mycobacterium tuberculosis* strain H37Rv (GenBank accession number L27989). Interspecies and intraspecies *rpoB* gene sequence variability was analysed using the in-house VARIability Analysis Program (SVARAP) (Colson *et al.*, 2006). The SVARAP analysis allowed us to design consensus PCR primers Myco-F (5' GGCAAGGTCACCCCGAAGGG 3'; base positions 2479–2498 with reference to the *M. paratuberculosis* K-10 *rpoB* sequence) and Myco-R (5' AGCGGCTGCTGGGTGATCATC 3'; base positions 3219–3239) in two conserved regions flanking the most variable *rpoB* region

(Adékambi *et al.*, 2003). This primer pair was used for partial amplification of the *rpoB* gene in the seven MAC reference strains. PCR was carried out in a 2720 Thermal Cycler (Applied Biosystems) in a 50 µl final volume containing 25 µl H₂O, 5 µl 10× buffer (Qiagen), 2.5 µl 25 µM MgCl₂, 5 µl 100 µM dNTP, 5 µM of each primer (Eurogentec), 2.5 U *Taq* DNA polymerase (Invitrogen) and 10 µl mycobacterial DNA using the following program: 5 min at 95 °C, followed by 35 cycles consisting of 94 °C for 30 s, 64 °C for 30 s and 72 °C for 90 s, and a 10 min elongation step at 72 °C. PCR mix was used as the negative control. Amplicons were purified by adding 50 µl distilled H₂O to a purification plate (Millipore) and agitated for 10 min. Forward and reverse sequencing mixtures contained 3 µl buffer (Big Dye V1, Applied Biosystems), 10 µl distilled H₂O and 1 µl 3.2 pmol µl⁻¹ primer in a final volume of 16 µl. The sequencing reaction comprised an initial denaturation step of 1 min at 95 °C, followed by 25 cycles of denaturation at 96 °C for 10s, annealing at 50 °C for 5 s and elongation at 60 °C for 3 min. Sequencing products were purified using a Sephadex plate (Amersham Biosciences) that was centrifuged at 720 g for 3 min and deposited on a MicroAmp Optical 96-well reaction plate (Applied Biosystems). Sequencing electrophoresis was performed on a 3100 Genetic Analyzer (Applied Biosystems). Accurate correction of the *rpoB* nucleotide sequence was assessed by its translation into the amino acid sequence (<http://www.expasy.org/tools/dna.html>). The sequences have been deposited in GenBank (Table 3) and the percentage of similarity between nucleotide sequences was determined by BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) (Tables 4 and 5).

rpoB gene sequence-based identification of clinical isolates.

The partial *rpoB* gene sequence was determined for every one of the 100 clinical isolates studied using the primer pair Myco-F/Myco-R and the PCR/sequencing protocol presented above. *rpoB* sequence similarity for the clinical isolates was determined in comparison with the reference sequences in the *rpoB* database by using the CLUSTAL_X program with a weighted residue table in the MEGALIGN package (Windows version 4.10e; DNASTAR). Isolates exhibiting ≥99.3 % sequence similarity with a reference strain were identified at the species level. Isolates exhibiting <99.3 % similarity with a reference strain were not identified at the species level (see below).

Additional tests for molecular identification. All reference strains, clinical isolates identified as *M. intracellulare*, *M. chimaera* and *M. avium* subsp. *paratuberculosis*, as well as isolates exhibiting an original, non-identified *rpoB* gene sequence, were submitted to further 16S rDNA, ITS and *hsp65* amplification and sequencing. 16S rDNA amplification and sequencing were done using the primer pair fd1–Rp2 (Woese, 1987). The conditions of amplification were 5 min at 95 °C, followed by 35 cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 1 min. For ITS amplification and sequencing, we used primers Ec16S.1390p and Mb23S.44n (Frothingham & Wilson, 1993) and 38 cycles of an initial 5 min

Table 2. List of clinical MAC and seven atypical clinical isolates investigated in this study

Species	No. of clinical isolates										Total (n=100)
	Sputum	Alveolar washing	Lymph node	Stomach aspirate	Cerebro- spinal fluid	Abscess	Biopsy	Blood	Stool	Unknown	
<i>M. avium</i> subsp. <i>paratuberculosis</i>								2			2
<i>M. avium</i> subsp. <i>hominissuis</i>	8	6	7	1	4		4	30	2	19	81
<i>M. intracellulare</i>	2	1				1	1	3			8
<i>M. chimaera</i>								2			2
68257	1										1
4355387	1										1
27497	1										1
5356591	1										1
62863		1									1
5351974	1										1
3256799	1										1
Total	16	8	7	1	4	1	5	37	2	19	100

denaturation at 95 °C, followed by 30 s denaturation at 94 °C, 30 s annealing at 62 °C and 90 s extension at 72 °C, with final extension for 7 min. For amplification and sequencing of the 3' end of *hsp65*, we initially used the primer pair *hsp65*-574F and *hsp65*-R (Turenne *et al.*, 2006); since these primers failed to amplify the homologous sequence in *M. colombiense*, we further designated primers *hsp65*-F106 (5' AACGTCGTCCTGGAGAAGAA 3') and *hsp65*-R1558 (5' GC CTTCTCCGGCTTGTC 3'), which covered almost the entire *hsp65* gene in this species with *hsp65*-574F (5' GGTTTCGACAAGGGYTACATC 3') as an additional primer for sequencing the 3' region of the gene. The conditions of PCR were 5 min at 95 °C, followed by 35 cycles of 95 °C for 45 s, 60 °C for 45 s and 72 °C for 90 s, with a final extension for 7 min. PCR mixtures (50 µl) contained 5 µl 10 × *Taq* buffer, 200 µM

each dNTP, 2.5 µM MgCl₂, 2.5 U *Taq* DNA polymerase, 10 mM of each primer, 25 µl sterile water and 10 µl mycobacterial DNA. Sequencing reactions were performed as described above.

Phylogenetic analyses. *M. tuberculosis* was used as an outgroup in all phylogenetic analyses. Sequences were trimmed to start and finish at the same nucleotide position for all reference strains and clinical isolates. Multisequence alignment was performed with the CLUSTAL_X program, version 1.81, in the PHYLIP software package. Five different phylogenetic trees were obtained from DNA sequences using the neighbour-joining method with the Jukes–Cantor parameter in the MEGA.3 program (Kumar *et al.*, 2004). Trees were based on the 16S rDNA, ITS, *hsp65* and *rpoB* sequences and concatenation of these four

Table 3. GenBank accession numbers of original sequences determined in this study

Species	Strain	GenBank accession number			
		16S rDNA	<i>rpoB</i>	ITS	<i>hsp65</i>
<i>M. avium</i> subsp. <i>avium</i>	ATCC25291 ^T	EF521895	EF521907	EF521901	EU239779
<i>M. avium</i> subsp. <i>paratuberculosis</i>	ATCC19698 ^T	EF521896	EF521906	EF521900	EU239780
<i>M. avium</i> subsp. <i>silvaticum</i>	ATCC49884 ^T	EF521891	EF521905	EF521899	EU239781
<i>M. avium</i> subsp. <i>hominissuis</i>	IWGMT49	EF521892	EF521911	EF521898	EU239782
<i>M. intracellulare</i>	ATCC15985	X52927	EF521900	EF521903	EU239784
<i>M. chimaera</i>	DSM44623 ^T	EF521894	EF521908	EF521902	EU239783
<i>M. colombiense</i>	CIP 108962 ^T	AM062764	EF521910	AM062764	EU239785
Atypical clinical isolate	68257	AM062764	EF584433	EU266629	EU239786
Atypical clinical isolate	4355387	EF591053	EF584445	EF591051	EU239787
Atypical clinical isolate	27497	EF611344	EF584432	EF591052	EU239788
Atypical clinical isolate	5356591	EU266632	EF584432	EU266631	EU239789
Atypical clinical isolate	62863	EU266632	EF584444	EU266630	EU239790
Atypical clinical isolate	3256799	EF591054	EF584435	EF591050	EU239791
Atypical clinical isolate	5351974	EF591054	EF584434	EU266628	EU239792

Table 4. Sequence similarity values for *rpoB* sequences (lower left) and *hsp65* sequences (upper right) of MAC reference strains and seven atypical clinical isolates

M. a. a, *M. avium* subsp. *avium*; *M. a. p*, *M. avium* subsp. *paratuberculosis*; *M. a. s*, *M. avium* subsp. *silvaticum*; *M. a. h*, *M. avium* subsp. *hominissuis*.

MAC species	Sequence similarity (%)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. <i>M. a. a</i>		99.0	100	95.7	94.7	94.8	95.0	94.3	95.0	99.5	94.2	95.7	94.8	94.7
2. <i>M. a. p</i>	99.7		99.2	95.3	94.5	94.6	94.6	94.1	94.6	99.1	94.0	95.3	95.0	94.5
3. <i>M. a. s</i>	99.8	99.5		95.7	94.7	94.8	95.0	94.3	95.0	99.5	94.2	95.7	94.8	94.7
4. <i>M. a. h</i>	99.7	99.7	99.5		91.0	91.1	91.5	90.9	91.5	96.1	90.8	92.0	91.1	91.0
5. <i>M. chimaera</i>	94.7	94.7	94.6	94.9		99.8	97.5	97.6	97.5	94.7	97.4	97.3	98.0	98.1
6. <i>M. intracellulare</i>	94.3	94.3	94.2	94.5	99.2		97.6	97.6	97.6	94.8	97.5	97.0	98.1	98.2
7. <i>M. colombiense</i>	92.9	92.9	92.8	92.9	93.9	93.5		97.5	99.6	95.0	97.5	97.8	97.3	97.0
8. 62863	93.5	93.5	93.6	93.6	96.7	96.3	94.2		98.0	94.3	99.8	96.7	98.0	97.2
9. 68257	93.6	93.6	93.5	93.6	94.9	94.5	97.7	93.2		95.0	97.4	97.8	97.3	97.0
10. 27497	95.0	95.0	94.9	95.0	95.7	95.3	96.2	94.6	97.1		94.0	96	94.8	94.7
11. 5356591	94.6	94.6	94.7	94.7	97.8	97.4	93.3	98.8	94.3	95.7		96.6	97.6	97.1
12. 4355387	95.2	95.2	95.0	95.2	96.2	95.7	97.1	94.6	98.1	98.8	95.7		97.4	96.9
13. 3256799	94.6	94.6	94.5	94.7	97.8	97.4	93.9	97.1	94.9	95.6	98.1	96.3		99.0
14. 5351974	94.6	94.6	94.5	94.7	97.8	97.4	93.9	97.1	94.9	95.6	98.1	96.3	100	

sequences. A bootstrap analysis (1000 repeats) was performed to evaluate the topology of each phylogenetic tree; bootstrap values above 75 % were considered significant.

RESULTS

Construction of a partial *rpoB* gene sequence database of MAC species

The primer pair Myco-F/Myco-R amplified a 711 bp fragment in every MAC reference strain studied, whereas

negative controls remained negative. Each amplicon yielded a deduced amino acid sequence that comprised 237 residues. Further analysis revealed the presence of nucleotide substitutions among the species tested with interspecies nucleotide sequence similarity ranging from 94.24 % between *M. intracellulare* and *M. avium* subsp. *silvaticum* to 99.3 % between *M. intracellulare* and *M. chimaera* (Table 4). Intraspecies similarity among *M. avium* subspecies ranged from 99.58 % among *M. avium* subsp. *silvaticum*, *M. avium* subsp. *hominissuis* and *M. avium* subsp. *paratuberculosis* to 99.86 % among *M. avium*

Table 5. Sequence similarity values for 16S rDNA sequences (lower left) and 16S–23S rDNA spacer (ITS) sequences (upper right) of MAC reference strains and seven atypical clinical isolates

M. a. a, *M. avium* subsp. *avium*; *M. a. p*, *M. avium* subsp. *paratuberculosis*; *M. a. s*, *M. avium* subsp. *silvaticum*; *M. a. h*, *M. avium* subsp. *hominissuis*.

MAC species	Sequence similarity (%)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. <i>M. a. a</i>		100	100	99.7	45.4	44.2	62.5	35.5	44.2	35.0	35.5	43.3	38.8	38.3
2. <i>M. a. p</i>	100		100	99.7	45.4	44.2	62.5	35.5	44.2	35.0	35.5	43.3	38.8	38.3
3. <i>M. a. s</i>	100	100		99.7	45.4	44.2	62.5	35.5	44.2	35.0	35.5	43.3	38.8	38.3
4. <i>M. a. h</i>	100	100	100		45.4	44.5	62.2	35.7	44.2	35.3	35.7	43.3	39.0	38.5
5. <i>M. chimaera</i>	96.3	96.3	96.3	96.3		95.6	51.1	86.3	95.6	84.5	86.1	95.4	90.1	90.4
6. <i>M. intracellulare</i>	96.7	96.7	96.7	96.7	100		50.6	87.4	94.0	84.7	87.2	95.6	90.8	90.6
7. <i>M. colombiense</i>	96.7	96.7	96.7	96.7	99.4	99.4		50.4	51.8	51.3	50.4	51.8	50.2	49.7
8. 62863	96.4	96.4	96.4	96.4	99.8	99.5	99.5		85.4	95.6	99.7	86.3	86.5	86.5
9. 68257	96.7	96.7	96.7	96.7	100	100	99.4	99.5		87.2	85.1	97.4	90.6	90.4
10. 27497	96.7	96.7	96.7	96.7	100	100	99.4	99.5	100		95.4	87.2	86.3	86.3
11. 5356591	96.4	96.4	96.4	96.4	99.8	99.5	99.5	100	99.5	99.5		86.1	86.3	86.3
12. 4355387	96.6	96.6	96.6	96.6	99.9	99.9	99.3	99.5	99.9	99.9	99.5		90.4	90.6
13. 3256799	96.3	96.3	96.3	96.3	99.6	99.6	99.6	99.7	99.6	99.6	99.7	99.5		99.3
14. 5351974	96.3	96.3	96.3	96.3	99.6	99.6	99.6	99.7	99.6	99.6	99.7	99.5	100	

subsp. *avium* and *M. avium* subsp. *silvaticum* (Table 4). We found no deletion or insertion within this *rpoB* gene region in the MAC reference strains.

***rpoB* sequence-based identification of clinical isolates**

All of the clinical isolates yielded a 711 bp amplicon when amplified using the Myco-F/Myco-R primer pair, whereas negative controls remained negative. Ninety-three clinical isolates (93 %) exhibited a partial *rpoB* gene sequence with ≥ 99.3 % similarity to one of the four MAC reference strains, and these isolates were regarded as identified at the species level (Table 4). They comprised 81 isolates exhibiting 100 % partial *rpoB* sequence similarity to *M. avium* subsp. *hominissuis* IWGMT49^T and two isolates with 99.7 % partial *rpoB* sequence similarity to *M. avium* subsp. *paratuberculosis*, eight isolates (8 %) exhibiting 99.2–99.7 % sequence similarity to *M. intracellulare*, and two blood isolates (2 %) exhibiting 100 % *rpoB* sequence similarity to *M. chimaera*. Seven isolates (7 %) exhibiting partial *rpoB* gene sequence <99.3 % similarity with any one of the reference strains were regarded as non-identified (Table 2). They comprised two sputum isolates (nos 3256799 and 5351974) exhibiting 97.89 % *rpoB* sequence similarity with *M. chimaera*, two isolates recovered from sputum (no. 62863) and alveolar washing (no. 5356591) exhibiting 96.7–97.8 % *rpoB* sequence similarity with *M. chimaera*, one sputum isolate (no. 68257) exhibiting 97.7 % *rpoB* gene sequence similarity with *M. colombiense*, one sputum isolate (no. 4355387) exhibiting 97.1 % *rpoB* sequence similarity with *M. colombiense*, and one sputum isolate (no. 27497) exhibiting 96.2 % *rpoB* sequence similarity with *M. colombiense*.

Reference molecular test results

The four *M. avium* subspecies yielded an identical 16S rDNA sequence, as did *M. intracellulare* and *M. chimaera* (Table 5). In the reference strains, ITS sequence similarity varied from 99.7 to 100 % among the four *M. avium* subspecies; this value was 62.5 % between *M. avium* subsp. *avium* and *M. colombiense*, 45.4 % between *M. avium* subsp. *avium* and *M. chimaera*, and 44.2 % between *M. avium* subsp. *avium* and *M. intracellulare*. The primer pair *hsp65*-574F and *hsp65*-R (Turenne *et al.*, 2006) amplified *hsp65* in *M. avium*, *M. intracellulare* and *M. chimaera*, but not in *M. colombiense*. Instead, we were able to amplify the 3' extremity of the *hsp65* gene in all reference strains, including *M. colombiense*, using PCR primers *hsp65*F-106 and *hsp65*R-1558 and primer *hsp65*F-574 for sequencing. The analysis of the *M. colombiense* sequence disclosed a T/C mismatch at position 588 (*M. avium* subsp. *paratuberculosis* K-10 numbering). The 3' extremity of *hsp65* had sequence similarity values ranging from 95.3 to 100 % among *M. avium* subspecies, 94.8 % between *M. avium* and *M. intracellulare*, 97.6 % between *M. intracellulare* and *M. colombiense*, and 99.8 % between *M. intracellulare* and *M.*

chimaera. The same three genes, 16S rDNA, ITS and *hsp65*, were further amplified and sequenced in clinical isolates, including the seven non-identified isolates. Accordingly, PCR targeting the *hsp65* gene using primer set *hsp65*-574F/*hsp65*-R (Turenne *et al.*, 2006) failed to amplify this gene in reference strain *M. colombiense* CIP 108962^T and three out of seven non-identified isolates, numbers 68257, 5356591 and 62863, whereas primers *hsp65*-F106/*hsp65*-R1558 successfully amplified almost the entire 1462 bp fragment of the *hsp65* gene for further sequencing. Isolates 3256799 and 5351974 yielded identical sequence similarity values of 99.6 % for 16S rDNA, 90.1 % for ITS and 98 % for *hsp65* with *M. chimaera* (Tables 4 and 5); isolates 62863 and 5356591 yielded identical similarity values of 99.8 % for 16S rDNA, 86.3 % for ITS and 97.5 % for *hsp65* with *M. chimaera*; isolate number 68257 yielded 100 % for 16S rDNA sequence similarity, 51.8 % for ITS sequence similarity and 99.6 % for *hsp65* sequence similarity with *M. colombiense*; isolate number 4355387 yielded 99.9 % for 16S rDNA sequence similarity, 51.8 % for ITS sequence similarity and 97.8 % for *hsp65* sequence similarity with *M. colombiense*; and isolate number 27497 yielded 100 % for 16S rDNA sequence similarity, 51.3 % for ITS sequence similarity and 95 % for *hsp65* sequence similarity with *M. colombiense*.

Phylogenetic analyses

We constructed phylogenetic trees (Figs 1 and 2) based on the 16S rDNA, ITS, partial *hsp65* and *rpoB* sequences determined for the seven reference strains of the four MAC species, and for the seven original clinical isolates that were not assigned to a known species on the basis of *rpoB* sequencing. Trees derived from separated sequences and the tree derived from concatenated genes gave identical results (Fig. 2). In all trees, the four subspecies of *M. avium* grouped together with bootstrap values of 94–100 %, while the other reference strains were clearly separated, with the exception of the 16S rDNA sequence-based tree, which did not separate *M. intracellulare* and *M. colombiense* because of the 100 % sequence similarity (Fig. 1a). The *rpoB* sequence-based tree was most similar to that of the concatenated tree, as it indicated that isolates 68257 and 4355387 were most closely related to *M. colombiense* with a 100 % bootstrap value, and that isolates 5351974 and 3256799 on the one hand and 62863 and 5356591 on the other hand formed two clusters separated from known MAC species with 100 % bootstrap values in the concatenated tree. The phylogenetic position of isolate number 27497 remained uncertain because this isolate was more closely related to *M. avium* in the *hsp65* sequence-derived tree and the concatenated tree with a 99 % bootstrap value, and more closely related to *M. colombiense* in the other trees with a 93 % bootstrap value in the 16S rDNA sequence-based tree, an 89 % bootstrap value in the ITS sequence-based tree and a 99 % bootstrap value in the *rpoB* sequence-based tree.

Fig. 1. Phylogenetic tree traced using the neighbour-joining method based on the 1442 bp 16S rDNA (a), the 438 bp ITS (b); indicating the sequavar for unidentified isolates, with MAC-X designating a previously unknown sequavar, the partial 995 bp *hsp65* gene (c) and the partial 711 bp *rpoB* gene (d) sequences of 100 MAC clinical isolates and seven reference strains. The sequences were aligned using the neighbour-joining method and Kimura's two-parameter distance correction model. Bootstrap values $\geq 75\%$ are indicated by the value at each node (as a percentage). *M. tuberculosis* H37Rv was used as the outgroup. Bars: (a, b) 0.2% difference in nucleotide sequence; (c, d), 0.1%. T, type sequence.

DISCUSSION

In this study, we found a unique partial *rpoB* sequence in each member of seven strains representative of the four MAC species and *M. avium* subspecies. Sequences determined herein were interpreted as authentic, since negative controls remained negative in all PCR experiments and the amplified fragments were of the predicted 711 bp size for the reference strains and clinical isolates. Moreover, we found original nucleotide sequences that were translated *in silico* into a peptide without any internal stop codon, further suggesting that these sequences were authentic and did not merely result from sequencing errors. All original nucleotide *rpoB* sequences have been deposited into GenBank (Table 3). In this study, partial *rpoB* sequences have been determined

using the primer pair Myco-F/Myco-R, previously used for the identification of RGM (Adékambi *et al.*, 2003). The present data therefore indicated that the usefulness of the Myco-F/Myco-R primer pair could be extended to the identification of slowly growing mycobacteria. This point may warrant further exploration in additional slowly growing mycobacteria species in order to confirm that the Myco-F/Myco-R primer pair is a broad-spectrum primer pair for partial *rpoB* gene amplification and sequencing in *Mycobacterium* species.

Reference strains of the four MAC species studied exhibited a 4.5–5.7% partial *rpoB* sequence divergence with the notable exception of *M. intracellulare* and *M. chimaera*, which shared a high, 99.3% partial *rpoB* sequence

Fig. 1. Continued.

similarity. The latter value was of the same order of magnitude as that observed among the four *M. avium* subspecies ($\leq 0.4\%$ sequence divergence). The close molecular proximity of *M. intracellulare* and *M. chimaera* disclosed herein by the *rpoB* sequence analysis was in agreement with the analysis of 16S rDNA, ITS and *hsp65* gene sequences in this study and the description of *M. chimaera* (Tortoli *et al.*, 2004). 16S rDNA has been regarded as the gold standard for molecular identification in *Mycobacterium* (Tortoli *et al.*, 2003), but the present data indicated that this may not be the case for *M. intracellulare* and *M. chimaera*, in agreement with the description of the latter species, which found only one nucleotide mismatch between *M. chimaera* and *M. intracellulare* sequavar i (Tortoli *et al.*, 2004). Another 16S rDNA sequence-based phylogenetic tree incorporating the four MAC species also showed the almost identical 16S rDNA sequences in *M. intracellulare* and *M. chimaera* (Murcia *et al.*, 2006). This situation reminded us of what we previously observed when using the same *rpoB* region

for the identification of RGM (Adékambi *et al.*, 2003). For RGM, a 2 % cut-off value was suitable for all of the tested RGM species, except for distinguishing *Mycobacterium fortuitum* and *Mycobacterium houstonense*, which exhibited 98.2 % sequence homology. Based on results obtained with reference strains of the currently acknowledged four MAC species, a 0.7 % *rpoB* sequence divergence cut-off was used for the identification of MAC clinical isolates.

Using this cut-off value, the vast majority of the clinical isolates (93 %) were identified at the species level. All *rpoB* sequence-based identifications have been further confirmed by polyphasic molecular identification including 16S rDNA, ITS and *hsp65* sequencing. Whereas the prevalence of *M. intracellulare* was 8 %, the vast majority of these isolates were *M. avium* subsp. *hominissuis*, a subspecies recently delineated among *M. avium* (Mijs *et al.*, 2002) and regarded as the ancestor group for the highly adapted *M. avium* subsp. *avium/silvaticum* and *M. avium* subsp. *paratuberculosis* organisms (Turenne *et al.*, 2007).

- Devulder, G., Pêrouse de Montclos, M. & Flandrois, J. P. (2005). A multigene approach to phylogenetic analysis using the genus *Mycobacterium* as a model. *Int J Syst Evol Microbiol* 55, 293–302.
- Djelouagji, Z. & Drancourt, M. (2006). Inactivation of cultured *Mycobacterium tuberculosis* organisms prior to DNA extraction. *J Clin Microbiol* 44, 1594–1595.
- Esparcia, Ó., Navarro, F., Quer, M. & Coll, P. (2008). Lymphadenopathy caused by *Mycobacterium colombiense*. *J Clin Microbiol* 46, 1885–1887.
- Frothingham, R. & Wilson, K. H. (1993). Sequence-based differentiation of strains in the *Mycobacterium avium* complex. *J Bacteriol* 175, 2818–2825.
- Gingeras, T. R., Ghandour, G., Wang, E., Berno, A., Small, P. M., Drobniowski, F., Alland, D., Desmond, E., Holodniy, M. & Drenkow, J. (1998). Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res* 8, 435–448.
- Glennon, M., Smith, T., Cormican, M., Noone, D., Barry, T., Maher, M., Dawson, M., Gilmartin, J. J. & Gannon, F. (1994). The ribosomal intergenic spacer region: a target for the PCR based diagnosis of tuberculosis. *Tuber Lung Dis* 75, 353–360.
- Karakousis, P. C., Moore, R. D. & Chaisson, R. E. (2004). *Mycobacterium avium* complex in patients with HIV infection in the era of highly active antiretroviral therapy. *Lancet Infect Dis* 4, 557–565.
- Kasai, H., Ezaki, T. & Harayama, S. (2000). Differentiation of phylogenetically related slowly growing mycobacteria by their *gyrB* sequences. *J Clin Microbiol* 38, 301–308.
- Kim, B. J., Lee, S. H., Lyu, M. A., Kim, S. J., Bai, G. H., Chae, G. T., Kim, E. C., Cha, C. Y. & Kook, Y. H. (1999). Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). *J Clin Microbiol* 37, 1714–1720.
- Kim, B. J., Lee, K. H., Park, B. N., Kim, S. J., Bai, G. H., Kim, S. J. & Kook, Y. H. (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.
- Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5, 150–163.
- Lauzi, S., Pasotto, D., Amadori, M., Archetti, I. L., Poli, G. & Bonizzi, L. (2000). Evaluation of the specificity of the gamma-interferon test in Italian bovine tuberculosis-free herds. *Vet J* 160, 17–24.
- Leao, S. C., Briones, M. R., Sircili, M. P., Balian, S. C., Mores, N. & Ferreira-Neto, J. S. (1999). Identification of two novel *Mycobacterium avium* allelic variants in pig and human isolates from Brazil by PCR-restriction enzyme analysis. *J Clin Microbiol* 37, 2592–2597.
- Lebrun, L., Weill, F. X., Lafendi, L., Houriez, F., Casanova, F., Gutierrez, M. C., Ingrand, D., Lagrange, P., Vincent, V. & Herrmann, J. L. (2005). Use of the INNO-LiPA-MYCOBACTERIA assay (version 2) for identification of *Mycobacterium avium*–*Mycobacterium intracellulare*–*Mycobacterium scrofulaceum* complex isolates. *J Clin Microbiol* 43, 2567–2574.
- Liu, X., Feng, Z., Harris, N. B., Cirillo, J. D., Bercovier, H. & Barletta, R. G. (2001). Identification of a secreted superoxide dismutase in *Mycobacterium avium* ssp. *paratuberculosis*. *FEMS Microbiol Lett* 202, 233–238.
- Mijs, W., de Haas, P., Rossau, R., Van der Laan, T., Rigouts, L., Portaels, F. & van Soolingen, D. (2002). Molecular evidence to support a proposal to reserve the designation *Mycobacterium avium* subsp. *avium* for bird-type isolates and ‘*M. avium* subsp. *hominissuis*’ for the human/porcine type of *M. avium*. *Int J Syst Evol Microbiol* 52, 1505–1518.
- Morita, Y., Maruyama, S., Kabeya, H., Nagai, A., Kozawa, K., Kato, M., Nakajima, T., Mikami, T., Katsube, Y. & Kimura, H. (2004). Genetic diversity of the *dnaJ* gene in the *Mycobacterium avium* complex. *J Med Microbiol* 53, 813–817.
- Murcia, M. I., Tortoli, E., Menendez, M. C., Palenque, E. & Garcia, M. J. (2006). *Mycobacterium colombiense* sp. nov., a novel member of the *Mycobacterium avium* complex and description of MAC-X as a new ITS genetic variant. *Int J Syst Evol Microbiol* 56, 2049–2054.
- Roth, A., Fischer, M., Hamid, M. E., Michalke, S., Ludwig, W. & Mauch, H. (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., Agha, S. A., El-Fiky, A. & Viljanen, M. K. (1996). Comparison of amplicon and 32-kilodalton PCR for detection of *Mycobacterium tuberculosis* from sputum specimens. *J Clin Microbiol* 34, 1829–1830.
- Swanson, D. S., Kapur, V., Stockbauer, K., Pan, X., Frothingham, R. & Musser, J. M. (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.
- Telenti, A., Marchesi, F., Balz, M., Bally, F., Bottger, E. C. & Bodmer, T. (1993). Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *J Clin Microbiol* 31, 175–178.
- Thierry, D., Bauge, S., Poveda, J. D., Vincent, V. & Guesdon, J. L. (1993). Rapid identification of *Mycobacterium avium*–intracellular complex strains: clinical practice evaluation of DT6 and DT1 probes. *J Infect Dis* 168, 1337–1338.
- Thorel, M. F., Krichevsky, M. & Levy-Frebault, V. V. (1990). Numerical taxonomy of mycobactin-dependent mycobacteria, emended description of *Mycobacterium avium*, and description of *Mycobacterium avium* subsp. *avium* subsp. nov., *Mycobacterium avium* subsp. *paratuberculosis* subsp. nov., and *Mycobacterium avium* subsp. *silvaticum* subsp. nov. *Int J Syst Bacteriol* 40, 254–260.
- Tortoli, E., Cichero, P., Piersimoni, C., Simonetti, M. T., Gesu, G. & Nista, D. (1999). Use of BACTEC MGIT 960 for recovery of mycobacteria from clinical specimens: multicenter study. *J Clin Microbiol* 37, 3578–3582.
- Tortoli, E., Mariottini, A. & Mazzarelli, G. (2003). Evaluation of INNO-LiPA MYCOBACTERIA v2: improved reverse hybridization multiple DNA probe assay for mycobacterial identification. *J Clin Microbiol* 41, 4418–4420.
- Tortoli, E., Rindi, L., Garcia, M. J., Chiaradonna, P., Dei, R., Garzelli, C., Kroppenstedt, R. M., Lari, N., Mattei, R. & other authors (2004). Proposal to elevate the genetic variant MAC-A, included in the *Mycobacterium avium* complex, to species rank as *Mycobacterium chimaera* sp. nov. *Int J Syst Evol Microbiol* 54, 1277–1285.
- Turenne, C. Y., Semret, M., Cousins, D. V., Collins, D. M. & Behr, M. A. (2006). Sequencing of *hsp65* distinguishes among subsets of the *Mycobacterium avium* complex. *J Clin Microbiol* 44, 433–440.
- Turenne, C. Y., Wallace, R., Jr & Behr, M. A. (2007). *Mycobacterium avium* in the postgenomic era. *Clin Microbiol Rev* 20, 205–229.
- Woese, C. R. (1987). Bacterial evolution. *Microbiol Rev* 51, 221–271.
- Yajko, D. M., Chin, D. P., Gonzalez, P. C., Nassos, P. S., Hopewell, P. C., Reingold, A. L., Horsburgh, C. R., Jr, Yakus, M. A., Ostroff, S. M. & Hadley, W. K. (1995). *Mycobacterium avium* complex in water, food, and soil samples collected from the environment of HIV-infected individuals. *J Acquir Immune Defic Syndr Hum Retrovirol* 9, 176–182.