# Dissection of phylogenetic relationships among 19 rapidly growing *Mycobacterium* species by 16S rRNA, *hsp65*, *sodA*, *recA* and *rpoB* gene sequencing

Toïdi Adékambi and Michel Drancourt

Unité des Rickettsies, CNRS UMR 6020 IFR 48, Faculté de Médecine, Université de la Méditerranée, 27, Boulevard Jean Moulin, 13385 Marseille Cedex 05, France

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

The current classification of non-pigmented and late-pigmenting rapidly growing mycobacteria (RGM) capable of producing disease in humans and animals consists primarily of three groups, the *Mycobacterium fortuitum* group, the *Mycobacterium chelonae–abscessus* group and the *Mycobacterium smegmatis* group. Since 1995, eight emerging species have been tentatively assigned to these groups on the basis of their phenotypic characters and 16S rRNA gene sequence, resulting in confusing taxonomy. In order to assess further taxonomic relationships among RGM, complete sequences of the 16S rRNA gene (1483–1489 bp), *rpoB* (3486–3495 bp) and *recA* (1041–1056 bp) and partial sequences of *hsp65* (420 bp) and *sodA* (441 bp) were determined in 19 species of RGM. Phylogenetic trees based upon each gene sequence, those based on the combined dataset of the five gene sequences and one based on the combined dataset of the *rpoB* and *recA* gene sequences were then compared using the neighbour-joining, maximum-parsimony and maximum-likelihood methods after using the incongruence length difference test. Combined datasets of the five gene sequences comprising nearly 7000 bp and of the *rpoB*+*recA* gene sequences comprising nearly 4600 bp distinguished six phylogenetic groups, the *M. chelonae–abscessus* group, the *Mycobacterium mucogenicum* group, the *M. fortuitum* group, the *Mycobacterium mageritense* group, the *Mycobacterium wolinskyi* group and the *M. smegmatis* group, respectively comprising four, three, eight, one, one and two species. The two protein-encoding genes *rpoB* and *recA* improved meaningfully the bootstrap values at the nodes of the different groups. The species *M. mucogenicum*, *M. mageritense* and *M. wolinskyi* formed new groups separated from the *M. chelonae–abscessus*, *M. fortuitum* and *M. smegmatis* groups, respectively. The *M. mucogenicum* group was well delineated, in contrast to the *M. mageritense* and *M. wolinskyi* groups. For phylogenetic organizations derived from the *hsp65* and *sodA* gene sequences, the bootstrap values at the nodes of a few clusters were $<70\,\%$. In contrast, phylogenetic organizations obtained from the 16S rRNA, *rpoB* and *recA* genes were globally similar to that inferred from combined datasets, indicating that the *rpoB* and *recA* genes appeared to be useful tools in addition to the 16S rRNA gene for the investigation of evolutionary relationships among RGM species. Moreover, *rpoB* gene sequence analysis yielded bootstrap values higher than those observed with *recA* and 16S rRNA genes. Also, molecular signatures in the *rpoB* and 16S rRNA genes of the *M. mucogenicum* group showed that it was a sister group of the *M. chelonae–abscessus* group. In this group, *M. mucogenicum* ATCC 49650$^T$ was clearly distinguished from *M. mucogenicum* ATCC 49649 with regard to analysis of the five gene sequences. This was in agreement with phenotypic and biochemical characteristics and suggested that these strains are representatives of two closely related, albeit distinct species.

Abbreviations: ILD, incongruence length difference; K2P, Kimura's two-parameter; PRA, PCR-RFLP analysis; RGM, rapidly growing mycobacterium.

The GenBank/EMBL/DDBJ accession numbers of the sequences reported in this study are given in Figs 1–5.

Details of the primers used for sequencing are available as supplementary material in IJSEM Online.

## INTRODUCTION

Mycobacteria are aerobic, non-mobile, mycolic acid-containing bacteria that are characteristically acid-fast and contain DNA with a G + C content of 62–72 mol%. They all lie within the genus *Mycobacterium*, the only genus in the family *Mycobacteriaceae*. They are classified into slowly growing and rapidly growing, and pigmented and non-pigmented species. The rapidly growing mycobacteria (RGM), which require less than 7 days to produce easily seen colonies on solid media, comprise 56 environmental species (Pfyffer *et al.*, 2003). Fifteen species commonly encountered in human and animals belong primarily to the *Mycobacterium chelonae–abscessus*, *Mycobacterium fortuitum* and *Mycobacterium smegmatis* groups (Brown-Elliott & Wallace, 2002; Lopez-Marin *et al.*, 1993). Indeed, *M. fortuitum*, *M. abscessus*, *M. chelonae*, *M. smegmatis*, *Mycobacterium immunogenum*, *Mycobacterium peregrinum*, *Mycobacterium houstonense* (formerly *M. fortuitum* third biovariant sorbitol-positive), *Mycobacterium neworleansense* (formerly *M. fortuitum* third biovariant sorbitol-negative), *Mycobacterium mageritense*, *Mycobacterium septicum*, *Mycobacterium mucogenicum* (formerly the *M. chelonae*-like organism), *Mycobacterium wolinskyi* and *Mycobacterium goodii* were isolated from patients (Brown-Elliott & Wallace, 2002) whereas *Mycobacterium senegalense*, *Mycobacterium farcinogenes* and *Mycobacterium porcinum* are pathogens of veterinary importance (Chamoiseau, 1979; Lopez-Marin *et al.*, 1993). The latter species has been recently shown to be indistinguishable from *M. fortuitum* third biovariant sorbitol-negative human isolates on the basis of PCR-restriction fragment length polymorphism analysis (PRA), partial 16S rRNA gene sequence and biochemical profile (Schinsky *et al.*, 2004).

As for other bacterial genera, 16S rRNA gene-based phylogeny has had a major influence on our perception of taxonomic relationships among the pathogenic RGM (Tortoli, 2003) in addition to updates brought by HPLC of mycolic acid esters, fluorescence-HPLC and PRA of a 439-bp fragment of the 65-kDa heat-shock protein-encoding gene (*hsp65*) (Brown-Elliott & Wallace, 2002).

Based on nucleotide differences in the 16S rRNA gene sequence, RGM were classified into three groups: the *M. fortuitum* group included *M. fortuitum*, *M. peregrinum*, *M. houstonense*, *M. neworleansense*, *M. septicum*, *M. mageritense*, *M. mucogenicum* and *M. senegalense*; the *M. chelonae–abscessus* group comprised *M. abscessus*, *M. chelonae* and *M. immunogenum*; and the *M. smegmatis* group comprised *M. smegmatis*, *M. wolinskyi* and *M. goodii* (Brown-Elliott & Wallace, 2002). However, there has been controversy regarding the taxonomic position of *M. mucogenicum*, since its biochemical profile and antibiotic-susceptibility pattern were more closely related to those of members of the *M. chelonae–abscessus* group (Wallace *et al.*, 1993). A 16S rRNA gene-based phylogenetic tree revealed that *M. mageritense* was more closely related to species of the *M. smegmatis* group (Brown *et al.*, 1999) than to the *M. fortuitum* group, despite its antibiotic susceptibility and biochemical patterns (Brown-Elliott & Wallace, 2002; Wallace *et al.*, 2002). Furthermore, little is known about the phylogenetic relationships of the three RGM of veterinary importance (Hamid *et al.*, 2002; Tortoli, 2003). Indeed, variations in the datasets, in terms of the strains and the number of species used and analytical methods, have made it difficult to compare and evaluate the proposed phylogenetic relationships.

Taxonomic and phylogenetic studies of mycobacterial species were for many years based on 16S rRNA gene analysis (Tortoli, 2003). The suggestion that bacterial strains belong to the same species if they have fewer than 5 to 15 base differences in their 16S rRNA gene sequences, proposed for other micro-organisms (Fox *et al.*, 1992), is not applicable to the RGM, whose members are more closely related to each other. Since it is doubtful whether phylogenetic relationships should be based solely on the 16S rRNA in cases where sequence identities are $\geqslant 99\,\%$ (Drancourt *et al.*, 2000), it was desirable to analyse genes other than the 16S rRNA gene in order to assess phylogenetic relationships further among the pathogenic RGM and to help to clarify their controversial taxonomy. In RGM, *hsp65* gene analysis was previously limited to PRA of a 439-bp fragment (Brunello *et al.*, 2001; Steingrube *et al.*, 1995). The partial sequence of the *sodA* gene encoding the superoxide dismutase was used to define the taxonomic status of *M. mageritense* within nine RGM under study (Domenech *et al.*, 1997). Recently, PRA of the 16S–23S internal transcribed spacer (ITS) was shown to be a suitable tool for identifying most mycobacterial species and to segregate them into five clusters (Roth *et al.*, 2000).

In this study, we examined the sequences of four genes in addition to the 16S rRNA gene in 19 pathogenic RGM species and we utilized a combined approach to summarize common support and to reconcile discrepancies among datasets: *sodA*, *hsp65*, the *recA* gene, encoding part of the DNA recombination and repair system, and the *rpoB* gene, encoding the $\beta$ subunit of the RNA polymerase. These genes were previously proposed as suitable phylogenetic markers for the classification of mycobacteria (Adékambi *et al.*, 2003; Blackwood *et al.*, 2000; Domenech *et al.*, 1997; Kim *et al.*, 1999; Rinquet *et al.*, 1999).

## METHODS

**Bacterial strains and molecular methods.** The type strains used in this study are listed in Figs 1–7. They were preserved at −20 °C in skimmed milk until used. Each strain was inoculated into Middlebrook 7H9 liquid medium and subcultured onto Middlebrook and Cohn 7H10 agar (Becton Dickinson) at 30 °C. Purity was confirmed by naked-eye examination of colonies and microscopic examination after Ziehl–Nielsen staining. Colonies were scraped and genomic DNA was extracted using the FAST DNA kit according to the instructions of the manufacturer (Q.biogene). PCRs were carried out in a Biometra thermocycler (BIOLABO). PCR mixtures (50 µl) contained 5 µl 10× *Taq* buffer, 200 µM of each dNTP, 2·5 mM $MgCl_2$, 1 U *Taq* DNA polymerase (Invitrogen), 10 mmol of each

appropriate pair of primers (Eurogentec), 33 µl sterile water and 2 µl of the purified DNA. PCR conditions are described below. Purified amplicons (QIAquick PCR purification kit; Qiagen) were sequenced using the ABI Prism d-Rhodamine dye terminator cycle sequencing ready reaction kit according to the manufacturer's instructions (Perkin Elmer Applied Biosystems) and the following program: 30 cycles of denaturation at 94 °C for 10 s, primer annealing at 50 °C for 15 s and extension at 60 °C for 1 min. Products of sequencing reactions were recorded with an ABI Prism 3100 DNA sequencer following the standard protocol of the supplier (Perkin Elmer Applied Biosystems).

### PCR primers and conditions

**16S rRNA gene.** Several RGM 16S rRNA gene sequences were available in GenBank at the initiation of this work. However, they were incomplete sequences (about 1200 bp) containing many ambiguities and large gaps, preventing informative comparison of the complete sequences. For this reason, all the type strains were reidentified by 16S rRNA gene sequence analysis using primer pair fD1–rP2 described by Woese (1987) (details of sequencing primers are given in Supplementary Table A in IJSEM Online). Conditions for 16S rRNA gene amplification were 2 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.

***hsp65.*** The *hsp65* gene sequence was available in GenBank for 15/19 RGM under study at the initiation of this work. All the type strains were reanalysed by partial 477-bp *hsp65* gene sequencing using primers −21M13F and M13R described by Rinquet *et al.* (1999) (Supplementary Table A). Additional primers Tb11 and Tb12 described by Telenti *et al.* (1993) (Supplementary Table A) were used for sequencing reactions. Conditions for *hsp65* gene amplification were 2 min at 95 °C followed by 35 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min, with a final extension step at 72 °C for 5 min.

***sodA.*** Nine *sodA* sequences were available in GenBank at the initiation of this work. They were partial 388-bp sequences including identical sequence for certain type strains (Domenech *et al.*, 1997). The only complete sequence available was that of *M. fortuitum*. The type strains were reanalysed by partial 541-bp *sodA* gene sequencing using primer pair SodlgF–SodlgR (Supplementary Table A) designed after alignment of the complete *sodA* sequences of *M. fortuitum*, *Mycobacterium tuberculosis* and *Mycobacterium leprae* (GenBank/Entrez accession nos X70914, NC_000962 and AL450380). Additional primers SodF–SodR described by Domenech *et al.* (1997) (Supplementary Table A) were used as internal sequencing primers. Conditions for *sodA* gene amplification were 2 min at 95 °C followed by 35 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 2 min, with a final extension step at 72 °C for 5 min.

***recA.*** Partial *recA* sequences were available in GenBank for 8/19 RGM under study at the initiation of this work, as well as the complete *recA* gene sequence of *M. smegmatis*. The type strains were characterized by complete *recA* gene sequencing using the primers described by Blackwood *et al.* (2000) and additional primers recF2b, rec755R, rec3288F, rec3335R and rec3575R designed after alignment of the complete *recA* sequences of *M. smegmatis*, *M. tuberculosis* and *M. leprae* (GenBank/Entrez accession nos X99208, NC_000962 and AL450380). Primer pairs recF2–recR1 (or recF2b–recR1), recG1–recR2 and rec3288F–rec3575R (Supplementary Table A) were used to amplify three overlapping fragments of the complete *recA* gene. Conditions for these PCRs were 2 min at 95 °C followed by 35 cycles of 94 °C for 30 s, 54 °C for 30 s and 72 °C for 2 min, with a final extension step at 72 °C for 5 min. Products were gel-purified and extracted with the QIAquick gel extraction kit (Qiagen) when necessary.

***rpoB.*** The type strains were identified by complete *rpoB* gene sequence analysis into seven overlapping fragments using primer pairs Smeg7F–Smeg601R, Smeg529F–Smeg1485R, MF–Smeg2333R, Fort623F–Smeg2649R, Smeg2426F–Smeg3288R and Smeg2835F–Smeg3668R and Smeg2885F–Fort4260R as previously described (Adékambi *et al.*, 2003; Kim *et al.*, 1999). Additional primers allowed completion of *rpoB* gene sequencing (Supplementary Table B). Conditions for these PCRs were 2 min at 95 °C followed by 35 cycles of 94 °C for 30 s, 64 °C for 30 s and 72 °C for 2 min, with a final extension step at 72 °C for 5 min.

**Sequence analyses and phylogenetic comparisons.** Percentages of similarity between sequences were determined using the CLUSTAL program with weighted residue weight table in the MegAlign package (Windows version 4.10e; DNASTAR). For phylogenetic analyses, sequences were trimmed in order to start and finish at the same nucleotide position for all the strains under study. Multisequence alignment was performed by using the CLUSTAL X program, version 1.81, in the PHYLIP software package (Thompson *et al.*, 1997). Phylogenetic trees were obtained from DNA sequences by using the neighbour-joining method with Kimura's two-parameter (K2P) distance correction model with 1000 bootstrap replications (in MEGA version 2.1; Kumar *et al.*, 2001), maximum-parsimony method with min-mini heuristic search option (MEGA version 2.1) and maximum-likelihood method (DNAml software in PHYLIP). The incongruence length difference (ILD) test or partition homogeneity test (1000 randomizations of datasets) was performed using PAUP4.0b program (Swofford, 1998). A bootstrap analysis (1000 repeats) using *M. tuberculosis* and *M. leprae* as the outgroups was performed to evaluate the robustness of the phylogenetic trees and bootstrap values above 90 % were considered as significant. These trees were rooted by using *M. tuberculosis* and *M. leprae*, which are the most closely related species to RGM for which the five genes under study were available in GenBank/Entrez (accession nos NC_000962 and AL450380).

## RESULTS AND DISCUSSION

In our approach, we sequenced a battery of five genes previously proposed as suitable tools for studying RGM phylogeny. The 16S rRNA (1483–1489 bp), *hsp65* (420 bp), *sodA* (441 bp), *recA* (1041–1056 bp) and *rpoB* (3486–3495 bp) genes were sequenced in 19 species and showed interspecies sequence similarity of 95·5–100 %, 87·8–100 %, 86·7–100 %, 84·4–100 % and 87·7–99·5 %, respectively. We first analysed the five genes separately and then used different combined databases to create a single alignment dataset. Identical phylogenetic trees were inferred from multiple sequence alignments of the nucleotide sequences using the neighbour-joining, maximum-parsimony and maximum-likelihood methods. The three methods gave almost similar bootstrap values. The bootstrap values given below were derived from the neighbour-joining method only using the K2P distance correction model.

### Separate analyses

**16S rRNA gene phylogenetic trees (Fig. 1).** The phylogenetic trees were based on 1490 unambiguously aligned positions, 71 of which were informative under the parsimony criterion. Five separate clusters were characterized. *M. chelonae* CIP 194535$^T$, *M. chelonae* ATCC 19237, *M. abscessus* and *M. immunogenum* formed a monophyletic

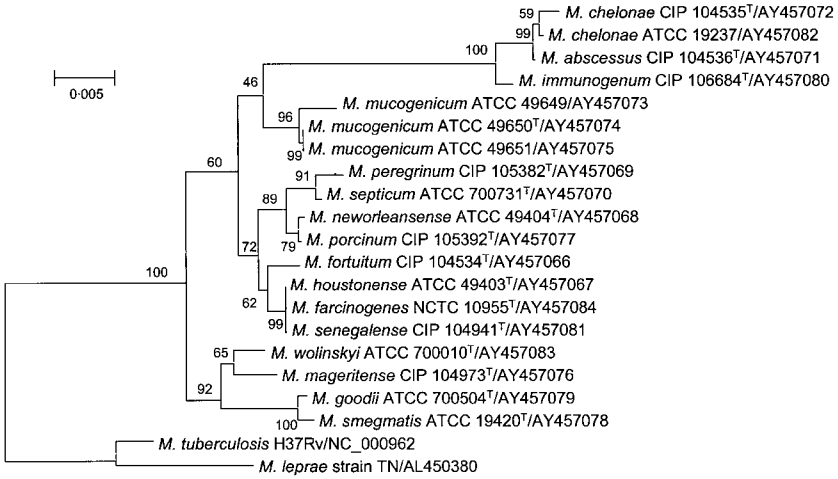


**Fig. 1.** Phylogenetic tree of 16S rRNA gene sequences of 19 RGM using the neighbour-joining method with K2P distance correction. The support of each branch, as determined from 1000 bootstrap samples, is indicated by percentages at each node. *M. tuberculosis* and *M. leprae* were used as the outgroups. Bar, 0·05 % nucleotide sequence difference.

cluster I with a 100 % bootstrap value. The three strains of *M. mucogenicum* formed a second monophyletic cluster II with a 96 % bootstrap value. *M. peregrinum*, *M. septicum*, *M. neworleansense*, *M. porcinum*, *M. farcinogenes*, *M. senegalense*, *M. houstonense* and *M. fortuitum* formed cluster III, with a 72 % bootstrap value. *M. mageritense* and *M. wolinskyi* formed cluster IV, albeit with a low bootstrap value (65 %). *M. smegmatis* and *M. goodii* formed a monophyletic cluster V with a 100 % bootstrap value, a sister group of cluster IV. Sister clusters I, II and III were supported by non-significant bootstrap values of 46–60 %. Two subclusters, IIIa (*M. peregrinum*, *M. septicum*, *M. neworleansense* and *M. porcinum*) and IIIb (*M. farcinogenes*, *M. senegalense*, *M. houstonense* and *M. fortuitum*), were supported by bootstrap values of 89 and 62 %, respectively. The bootstrap values at the nodes of a few clusters obtained were too low to induce much confidence. The 16S rRNA gene sequence did not discriminate *M. houstonense*, *M. senegalense* and *M. farcinogenes* or *M. mucogenicum* ATCC 49650$^T$ from *M. mucogenicum* ATCC 49651. These data indicate that the 16S rRNA gene sequence alone did not resolve the phylogenetic relationships between all currently recognized RGM species as indicated by the low bootstrap values obtained at the nodes of different sister groups and subgroups ($\leqslant$72 %). The percentage similarity of the 16S rRNA gene sequence in 19 RGM examined was 95·5–100 %, explaining the poor phylogenetic information obtained from this sequence.

***hsp65* phylogenetic trees (Fig. 2).** The phylogenetic trees were based on 420 unambiguously aligned positions (omitting the primer sequences used for the amplification), 59 of which were informative under the parsimony criterion. The deduced amino acid sequences comprised 140 residues [$G_{52}$ to $M_{191}$ (*M. tuberculosis* numbering)]. *M. chelonae* CIP 194535$^T$, *M. chelonae* ATCC 19237, *M. abscessus* and *M. immunogenum* grouped together with a non-significant bootstrap value of 51 %. The three *M. mucogenicum* strains formed a monophyletic cluster II with a 99 % bootstrap value. Subclusters IIIa (*M. peregrinum*, *M. septicum*, *M. porcinum* and *M. neworleansense*) and IIIb (*M. fortuitum*, *M. houstonense*, *M. senegalense* and *M. farcinogenes*) were supported by bootstrap values of 61 and 84 %, respectively. Subcluster IIIb was found to be a sister group of cluster IV (*M. mageritense* and *M. wolinskyi*), which was found to be a poorly monophyletic

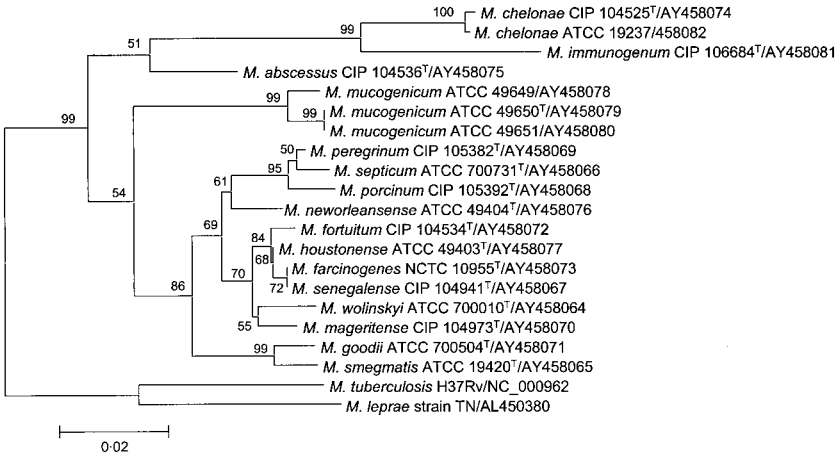


**Fig. 2.** Phylogenetic tree of *hsp65* gene sequences of 19 RGM using the neighbour-joining method with K2P distance correction model. See Fig. 1 for further details. Bar, 2 % nucleotide sequence difference.

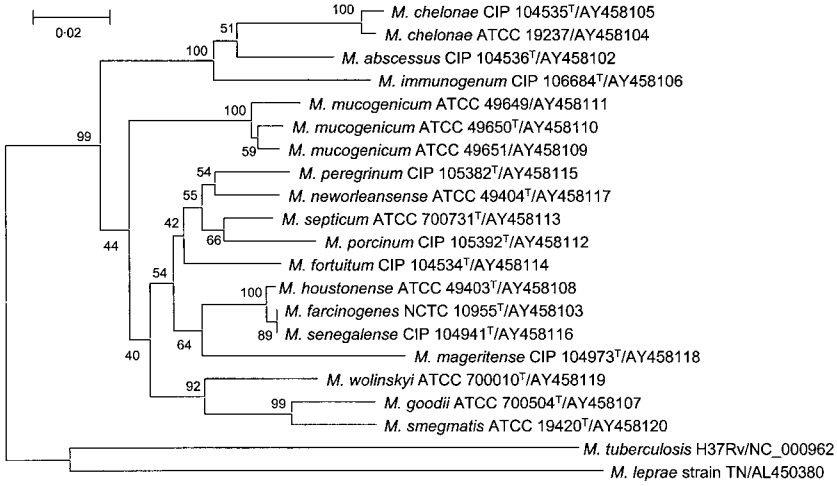


**Fig. 3.** Phylogenetic tree of *sodA* gene sequences of 19 RGM using the neighbour-joining method with K2P distance correction model. See Fig. 1 for further details. Bar, 2 % nucleotide sequence difference.

cluster (55 %). *M. smegmatis* and *M. goodii* formed a monophyletic cluster V with a bootstrap value of 99 %. However, the *hsp65* gene sequence did not discriminate either *M. senegalense* from *M. farcinogenes* or *M. mucogenicum* ATCC 49650$^T$ from *M. mucogenicum* ATCC 49651.

***sodA* phylogenetic trees (Fig. 3).** The phylogenetic trees were based on 441 unambiguously aligned positions, 86 of which were informative under the parsimony criterion. The deduced amino acid sequences comprised 147 residues [$Q_{23}$ to $Q_{169}$ (*M. tuberculosis* numbering)]. *M. chelonae* CIP 194535$^T$, *M. chelonae* ATCC 19237, *M. abscessus* and *M. immunogenum* formed a monophyletic cluster I with a 100 % bootstrap value. The three *M. mucogenicum* strains formed a monophyletic cluster II (100 %). *M. peregrinum*, *M. septicum*, *M. neworleansense*, *M. porcinum* and *M. fortuitum* grouped into subcluster IIIa (42 %), whereas *M. houstonense*, *M. senegalense* and *M. farcinogenes* formed a monophyletic subcluster IIIb with a bootstrap value of 100 %. *M. mageritense* appeared to be the only species of cluster IV, forming a sister group of subcluster IIIb with a 64 % bootstrap value at the node. *M. smegmatis* and *M. goodii* formed a monophyletic cluster V with a 99 % bootstrap value. *M. wolinskyi* seemed to be the only species of cluster VI, forming a sister group of cluster V with a 92 % bootstrap value at the node. The *sodA* gene sequence did not discriminate *M. senegalense* from *M. farcinogenes*.

***recA* phylogenetic trees (Fig. 4).** The phylogenetic trees were based on 1056 unambiguously aligned positions, including 236 informative positions under the parsimony criterion. *M. chelonae* CIP 194535$^T$, *M. chelonae* ATCC 19237, *M. abscessus* and *M. immunogenum* formed a monophyletic cluster I with a 100 % bootstrap value. The three *M. mucogenicum* strains formed a monophyletic cluster II with a 100 % bootstrap value. *M. peregrinum*, *M. septicum*, *M. neworleansense*, *M. porcinum* and *M. fortuitum* formed subcluster IIIa (88 %) and *M. houstonense*, *M. senegalense* and *M. farcinogenes* formed a monophyletic subcluster IIIb (100 %). *M. smegmatis* and *M. goodii* grouped into cluster V with a 100 % bootstrap value. *M. mageritense* and *M. wolinskyi* were the only species of clusters IV and VI, respectively; they grouped with 30 and

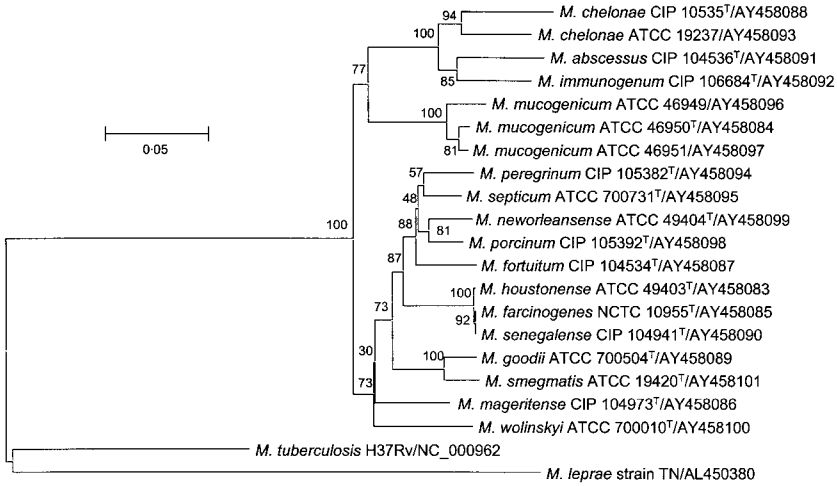


**Fig. 4.** Phylogenetic tree of *recA* gene sequences of 19 RGM using the neighbour-joining method with K2P distance correction model. See Fig. 1 for further details. Bar, 5 % nucleotide sequence difference.

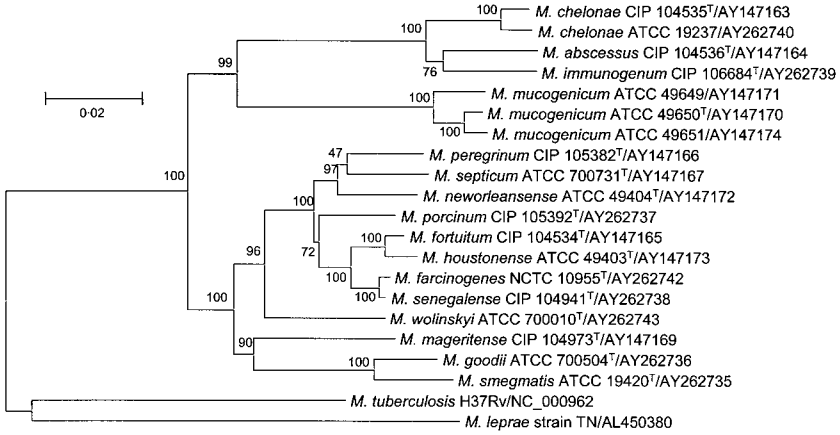


**Fig. 5.** Phylogenetic tree of *rpoB* gene sequences of 19 RGM using the neighbour-joining method with K2P distance correction model. See Fig. 1 for further details. Bar, 2 % nucleotide sequence difference.

73 % bootstrap values with the respective sister group at the node. The *rec*A gene sequence did not discriminate *M. senegalense* from *M. farcinogenes*.

***rpoB* phylogenetic trees (Fig. 5).** The phylogenetic trees were based on 3533 bp unambiguously aligned positions, comprising 650 informative positions under the parsimony criterion. *M. chelonae* CIP 194535$^T$, *M. chelonae* ATCC 19237, *M. abscessus* and *M. immunogenum* formed a monophyletic cluster I with a 100 % bootstrap value. The three *M. mucogenicum* strains formed a monophyletic cluster II with a 100 % bootstrap value. Subclusters IIIa (*M. peregrinum*, *M. septicum* and *M. neworleansense*) and IIIb (*M. porcinum*, *M. fortuitum*, *M. houstonense*, *M. senegalense* and *M. farcinogenes*) were supported by bootstrap values of 97 and 72 %, respectively, and formed cluster III, supported by a 100 % bootstrap value. *M. wolinskyi* was the only representative of cluster IV, with a 96 % bootstrap value at the node with cluster III. *M. smegmatis* and *M. goodii* formed a monophyletic cluster VI with a 100 % bootstrap value. *M. mageritense* was the only species in cluster V, with a 90 % bootstrap value at the node with cluster VI.

## Combination in a simultaneous analysis (Fig. 6)

The combined 16S rRNA, *hsp65*, *sodA*, *recA* and *rpoB* phylogenetic trees were compared with individual phylogenies. *M. chelonae* CIP 194535$^T$, *M. chelonae* ATCC 19237, *M. abscessus* and *M. immunogenum* formed a monophyletic cluster I with a 100 % bootstrap value. The three *M. mucogenicum* strains formed a monophyletic cluster II with a 100 % bootstrap value. Subclusters IIIa (*M. peregrinum*, *M. septicum*, *M. neworleansense* and *M. porcinum*) and IIIb (*M. fortuitum*, *M. houstonense*, *M. senegalense* and *M. farcinogenes*) were supported by bootstrap values of 95 and 100 %, respectively. *M. wolinskyi* and *M. mageritense* formed two independent branches. However, *M. wolinskyi* (cluster IV) and *M. mageritense* (cluster V) and their respective sister groups were supported by low bootstrap values of 44 and 35 %, respectively. *M. smegmatis* and *M. goodii* formed a monophyletic cluster VI with a 100 % bootstrap value.

## Conditional combination approach (Fig. 7)

In this approach, the ILD test was used as a preliminary step before choosing either to combine congruent data to

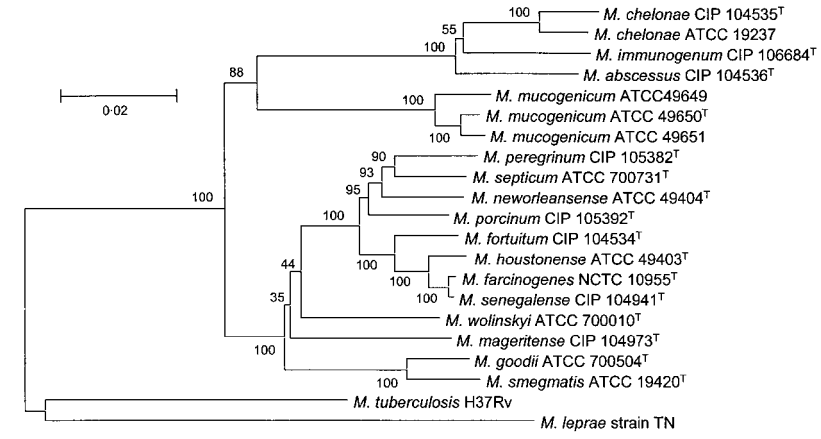


**Fig. 6.** Phylogenetic tree of the combined *rpoB*+*recA*+*sodA*+*hsp65*+16S rRNA gene sequences of 19 RGM using the neighbour-joining method with K2P distance correction model. See Fig. 1 for further details. Bar, 2 % nucleotide sequence difference.

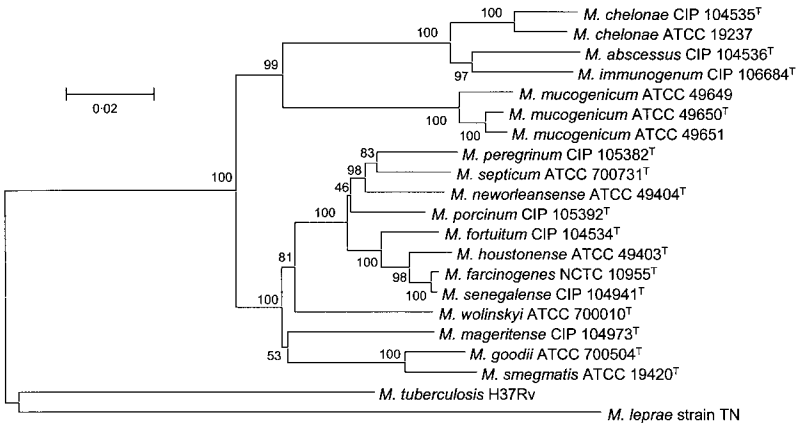


**Fig. 7.** Phylogenetic tree of the combined *rpoB*+*recA* gene sequences of 19 RGM using the neighbour-joining method and K2P distance correction model. See Fig. 1 for further details. Bar, 2 % nucleotide sequence difference.

increase the accuracy of the phylogenetic reconstruction or to analyse the data separately to discover the reasons for incongruence. Despite the incongruence observed within the five genes, *rpoB*, *recA* and 16S RNA gene sequencing trees had globally the same topology. The ILD test was applied to check whether trees for the different genes were sufficiently similar in rates of divergence and branching order for it to be legitimate to combine the data (Cunningham, 1997). The test calculates how much longer the combined parsimony tree is than the sum of the separate trees and compares this difference with the expected distribution estimated by allocating the data randomly into partitions of the same size as the gene. A three-partition test was conducted on the *rpoB*, *recA* and 16S rRNA gene sequences. Invariant characters were removed for all analyses (Gaunt *et al.*, 2001). The ILD was supported for *rpoB*+*recA* ($P=0{\cdot}73$) but rejected for *rpoB*+*recA*+16S rRNA ($P<0{\cdot}001$). These results allowed us to combine the *rpoB* and *recA* gene sequences, but suggested that the 16S rRNA gene should be treated separately, probably because of the small number of informative sites (Darlu & Lecointre, 2002). A neighbour-joining phylogenetic tree derived from the combined *rpoB*+*recA* gene sequences (Fig. 7) was compared with the corresponding five gene sequences and 16S rRNA gene sequence. The two protein-encoding genes *rpoB*+*recA* led to the same topology as observed with the simultaneous analyses but they improved the bootstrap values meaningfully. The bootstrap values at the nodes of *M. wolinskyi* (cluster IV) and *M. mageritense* (cluster V) with their respective sister groups increased from 44 to 81 % and from 35 to 53 %, respectively. The bootstrap values at the nodes of the *M. chelonae–abscessus* and *M. mucogenicum* groups increased from 88 to 99 % (Fig. 7). Furthermore, the phylogenetic tree derived from the combined *rpoB*+*recA* gene sequences supported the 16S rRNA gene classification of the 19 RGM (Figs 1 and 7). Thus, our *rpoB*+*recA* data reinforce the view that the 16S rRNA gene-based phylogeny holds for the majority of the genome.

## Evidence from insertions and deletions (molecular signatures)

Several molecular signatures due to insertions and deletions in the sequence alignment under study provided additional phylogenetic evidence. Relative to all the other *rpoB* gene sequences (including outgroups), the *M. chelonae–abscessus* and *M. mucogenicum* groups have 3- and 9-bp deletions at positions 37–39 and 2830–2838, respectively, and a 6-bp insertion at positions 967–972. Moreover, the *M. mucogenicum* group has an additional 3-bp deletion at positions 34–36 (*M. fortuitum* numbering). These deletions and insertions suggested that these two groups were sister groups. *M. smegmatis* and *M. goodii* have a 3-bp deletion at positions 28–30. Within the *M. fortuitum* group, subcluster IIIa with the exception of *M. neworleansense* exhibited a 3-bp deletion at positions 22–24, whereas *M. porcinum*, a member of subcluster IIIb, had the same deletion.

There were also length variants in the 16S rRNA gene sequence resulting from regions too variable to ascertain the exact alignment. In *M. smegmatis* and *M. goodii*, there were 2-bp insertions at positions 82–83 and 185–186, a 1-bp insertion at position 202 and a 1-bp deletion at position 178 (*Escherichia coli* numbering). A 1-bp deletion was found at positions 818 and 1267 in *M. mucogenicum* strains and the species of cluster I (*M. chelonae*, *M. abscessus*, *M. immunogenum*). Similarly, *M. mageritense* and *M. wolinskyi* shared two 1-bp insertions at positions 1433 and 1440 also noted in the *M. goodii* 16S rRNA gene sequence.

Several signatures due to small insertions and deletions in the *recA* sequence alignment provided additional phylogenetic evidence. Relative to all the other sequences (including outgroups), *M. fortuitum*, *M. mucogenicum*, *M. smegmatis*, *M. mageritense* and *M. wolinskyi* groups exhibited insertions of 3 and 6 bp at positions 7–9 and 1012–1017. Likewise, *M. tuberculosis*, *M. fortuitum* and *M. peregrinum* have an additional 6-bp insertion at positions

1018–1023. These related deletions and insertions in sequence provided good support for the monophyly of these trees as shown in Figs 1, 4 and 5. No deletions or insertions were found in the partial sequences of *hsp65* (420 bp) and *sodA* (441 bp).

### Comparison between *rpoB*, *recA*, *sodA*, *hsp65* and 16S rRNA gene phylogenetic trees

*rpoB*-, *recA*-, *sodA*-, *hsp65*- and 16S rRNA gene-based phylogeny inferred from neighbour-joining was compared. The topologies of the *rpoB*-based phylogenetic trees (Fig. 5) were almost the same as those derived from 16S rRNA and *recA* gene sequence analysis (Figs 1 and 4). However, the trees derived from *rpoB* gene sequence showed higher bootstrap values and more divergence than the *recA* and 16S rRNA gene-based trees. Bootstrap support for each cluster was higher than 90 % in the *rpoB* sequence analysis. The non-pigmented *M. mageritense* and *M. wolinskyi* groups were not well defined, with lower bootstrap values in the *recA* and 16S rRNA gene-based trees. All tested species showed good discrimination with regard to the *rpoB* gene. On the other hand, *M. abscessus* and *M. chelonae*, poorly discriminated by the 16S rRNA gene sequence (Adékambi *et al.*, 2003; Kirschner *et al.*, 1992), were clearly delineated as two closely related species, like *M. senegalense* and *M. farcinogenes*. Similarly, *M. mucogenicum* ATCC 49649 was clearly separated from *M. mucogenicum* ATCC 49650$^{T}$ and *M. mucogenicum* ATCC 49651 despite the fact that they belong to the same species (Springer *et al.*, 1995). As for the *hsp65* gene, *M. mageritense* and *M. wolinskyi* formed a monophyletic group with a non-significant bootstrap value (55 %). In the *sodA* gene sequence analysis, *M. mageritense* formed a sister cluster with that formed by *M. houstonense*, *M. senegalense* and *M. farcinogenes* with a non-significant bootstrap value (64 %). Indeed, the bootstrap values at the nodes of a few clusters were too low to induce much confidence with regard to the *hsp65* and *sodA* trees (Figs 2 and 3) and the suggestion that the relationships among the species are not the same in comparison to the 16S rRNA, *recA* and *rpoB* genes. These discrepancies can arise from inadequate sample sizes (Bull *et al.*, 1993), since it was recently suggested that long sequences of several kilobases should be used in studies that aim to use clock tests to select sequences that approximate rate constancy (Bromham *et al.*, 2000). These findings suggested that the *rpoB*- and *recA*-based phylogenies of RGM can be additional phylogenetic tools that globally support the 16S rRNA gene-based phylogeny of RGM.

However, separate analyses of individual datasets ignore hidden character support within datasets that emerges in combined analysis, and this support can be substantial (Gatesy & Arctander, 2000). Separate analyses of different character sets are not necessary to detect conflict among datasets and can distort interpretations of common character support (Gatesy *et al.*, 1999). Therefore, it was suggested that the distribution of conflicts/support among datasets in a comprehensive combined analysis should be used to assess dataset congruence and to question the strength of support for different clusters. Keeping account of potentially dependent characters in simultaneous analysis offers the researcher a better understanding of the global evidence. In our study, nearly 7000 bp were analysed, corresponding to roughly 0·05 % of the genomic sequence and 0·1 % of the predicted genes. To our knowledge, this is the first reported example of so many nucleotides incorporated into a combined dataset analysis of a bacterial genus. When using the ILD test to quantify the conflicts that can occur between sets of characters from different data sources, combining *rpoB* and *recA* genes (nearly 4600 bp) broadly supported the expected phylogeny and improved the bootstrap values. Also, removal of two not-well-delineated groups, *M. mageritense* and *M. wolinskyi*, increased the mean bootstrap value to $\geqslant$99 % at the node of each group. However, the lower bootstrap value of certain species within a group for combined genes relative to randomly resampled nucleotides is consistent with the hypothesis that nucleotides within genes have not evolved independently (Averof *et al.*, 2000; Cummings *et al.*, 1995).

According to these observations, the two combined datasets obtained allowed us to discriminate six groups, as described below (Figs 6 and 7).

The *M. chelonae–abscessus* group (cluster I) was created to accommodate species with distinctive properties including positive 3-day arylsulfatase, better growth at 30 than at 35 °C, negative nitrate reductase, negative iron uptake and resistance to polymyxin B (Brown-Elliott & Wallace, 2002). They formed a defined cluster in the 16S rRNA and *hsp65* gene phylogenies (Tortoli, 2003; Wilson *et al.*, 2001). The recently described *M. immunogenum* was added to this group according to its phenotypic characteristics, its unique *hsp65* sequence pattern and its 15–18 % DNA–DNA relatedness to other species of this group (Wilson *et al.*, 2001). Our analyses of the single and combined datasets fully supported both the existence of this group and its taxonomic relationships. Further evidence was provided by molecular signatures: a 3-bp insertion at positions 34–36 of the *rpoB* gene, 3- and 6-bp deletions at positions 7–9 and 1012–1017 of the *recA* gene and a 1-bp deletion at position 1267 of the 16S rRNA gene. A recent study of ITS sequences (Hamid *et al.*, 2002) confirmed that the *M. chelonae–abscessus* group was a separate group. With the data from the five combined genes, *M. immunogenum* appears to be a hybrid between *M. abscessus* and *M. chelonae*, with a 55 % bootstrap value. This phenomenon was previously described using PRA band patterns from a 439-bp segment of the *hsp65* gene digested with *Bst*EII and *Hae*III (Wilson *et al.*, 2001). Further studies are warranted to assess better the taxonomic relationships of this emerging taxon.

The *M. mucogenicum* group (cluster II) formed a distinct cluster in our study. This was discrepant with previous

suggestions of its incorporation into the *M. chelonae–abscessus* group (Wallace *et al.*, 1993) or the *M. fortuitum* group (Brown-Elliott & Wallace, 2002). Some phenotypic characteristics support our analysis, notably the highly mucoid aspect of the colonies and the susceptibility of *M. mucogenicum* to polymyxin B (Springer *et al.*, 1995). Species of the *M. mucogenicum* group utilize mannitol as a carbon source, in contrast to species of the *M. chelonae–abscessus* group, and also citrate, in contrast to species of the *M. fortuitum* group (Springer *et al.*, 1995). Likewise, comparative thin-layer and gas chromatography of fatty esters and alcohols showed clear differentiation of *M. mucogenicum* isolates from isolates belonging to the *M. fortuitum* and *M. chelonae–abscessus* groups (Muñoz *et al.*, 1997; Wallace *et al.*, 1993). In addition, *M. mucogenicum* strain ATCC 49649 differs from *M. mucogenicum* ATCC 49650$^T$ by exhibiting high level of semi-quantitative catalase activity, cephalothin resistance and negative acetamidase (Wallace *et al.*, 1993). The 16S rRNA gene sequence of *M. mucogenicum* ATCC 49649 (GenBank accession no. AY457073) obtained in our laboratory showed a double C and T (Y) peak at position 448, suggesting the presence of two 16S rRNA gene alleles. This was in accordance with the fact that RGM have two 16S rRNA gene copies with the exception of *M. abscessus* and *M. chelonae*, which have only one copy (Domenech *et al.*, 1994). Double peaks were not observed in other RGM 16S rRNA genes sequenced in our laboratory. Furthermore, *M. mucogenicum* ATCC 49650$^T$ was the only strain to contain a lipooligosaccharide composed of glucose and galactose (Muñoz *et al.*, 1998). The percentage *rpoB* gene sequence similarity between these two strains suggested that, like *M. peregrinum* and *M. septicum*, they may be representatives of two closely related species, as they occupy similar positions in the *M. mucogenicum* group of the phylogenetic tree (Fig. 6). Further evidence is provided by a 3-bp deletion at positions 34–36 of the *rpoB* gene and a 1-bp deletion at position 818 of the 16S rRNA gene.

The *M. fortuitum* group (cluster III) group included three RGM of veterinary interest, *M. porcinum*, *M. senegalense* and *M. farcinogenes*, as determined by 16S rRNA, *hsp65*, *sodA*, *recA* and *rpoB* gene sequencing. This result is in good agreement with DNA–DNA relatedness studies, which indicated a moderate level of relatedness between *M. fortuitum*, *M. peregrinum*, *M. farcinogenes* and *M. senegalense* (Lévy-Frébault *et al.*, 1986; Schinsky *et al.*, 2000). Numerical analyses (Ridell & Goodfellow, 1983; Tsukamura & Ichiyama, 1986), comparative immunology (Lopez-Marin *et al.*, 1993; Lanéelle *et al.*, 1996) and chemotaxonomic properties (Kirschner *et al.*, 1992; Wallace *et al.*, 2002) revealed a high degree of similarity within the species of the *M. fortuitum* group. The 99 % *rpoB* gene sequence similarity between *M. houstonense* and *M. fortuitum* suggested that these strains may be closely related subspecies, although *M. houstonense* showed resistance to pipemidic acid, biochemical differences such as mannitol, inositol, sorbitol and trehalose utilization (Kirschner *et al.*, 1992) and three base differences in the 16S rRNA gene sequence. The combined datasets delineated two subclusters, IIIa (*M. peregrinum*, *M. septicum*, *M. neworleansense* and *M. porcinum*) and IIIb (*M. fortuitum*, *M. houstonense*, *M. senegalense* and *M. farcinogenes*), within this group. However, a minor discrepancy was observed within these subclusters. In contrast to the 16S rRNA, *hsp65*, *sodA* and *recA* gene sequence analyses and the combined datasets, where *M. porcinum* belongs to subcluster IIIa, *rpoB* gene sequence analysis placed it in subcluster IIIb with a 65 % bootstrap value, although it possessed a 3-bp deletion at positions 22–24 characteristic of the species of subcluster IIIa (Fig. 6). These data suggested an intraspecific horizontal *rpoB* gene fragment exchange, as recently demonstrated by Lorenz & Sikorski (2000). In this group, the two subclusters IIIa and IIIb can be differentiated in that species of subcluster IIIb grow at 42 °C (Ridell & Goodfellow, 1983; Tsukamura & Ichiyama, 1986; Schinsky *et al.*, 2000) and utilize sorbitol as a carbon source. Species of this group are responsible for animal infections (Youssef *et al.*, 2002). Furthermore, *M. fortuitum* was the most commonly isolated mycobacterium obtained from chronic non-healing skin lesions in dogs and cats (Jang & Hirsh, 2002). Youssef *et al.* (2002) proposed the name '*M. fortuitum–peregrinum* group' on the basis of tissue culture, chromatography and histopathological examination of pyogranulomatous panniculitis in a 6-year-old cat. This name seems to be appropriate, since *M. fortuitum* and *M. peregrinum* were the first species described in subclusters IIIa and IIIb, respectively. A recent study of the ITS sequence showed that the *M. fortuitum* group was a separate group (Hamid *et al.*, 2002). In this group, *M. senegalense* was indistinguishable from *M. farcinogenes* on the basis of all genes analysed except *rpoB* (0·5 % divergence). This divergence was not in agreement with the suggestion that isolates may have species status when they exhibit $\geqslant$3 % *rpoB* gene sequence divergence (Adékambi *et al.*, 2003). However, these two mycobacteria can be differentiated on the basis of their DNA–DNA relatedness, growth rates, biochemical activity and histopathological behaviour (Chamoiseau, 1979) as well as sequencing of the ITS (Hamid *et al.*, 2002).

The *M. wolinskyi* group (cluster IV) and *M. mageritense* group (cluster V) represent independent branches within the trees, although they appear together in the 16S rRNA and *hsp65* gene sequence trees with low bootstrap values of 66 and 55 %, respectively. These two genes were found to be less informative than *sodA*, *recA* and *rpoB*. The relative branching order of these sequences was not well supported by bootstrap values. Determining the relationships between these organisms and other species under study will require sequencing of additional RGM isolates such as *M. wolinskyi* ATCC 70009, whose 16S rRNA gene sequence (GenBank accession no. Y12871) differs from that of the type strain (Tortoli, 2003). *M. wolinskyi* and *M. mageritense* formed two separate groups, separated from the *M. smegmatis* and *M. fortuitum* groups, respectively. This result is supported by

DNA–DNA hybridization values of $<28\%$ between *M. mageritense* and the *M. fortuitum* group and of $<20\%$ between *M. wolinskyi* and *M. smegmatis* group species (Domenech *et al.*, 1997; Brown *et al.*, 1999). Furthermore, *M. wolinskyi* was distinguished from the *M. smegmatis* group due to the absence of pigmentation. Similarly, *M. mageritense* was distinguished from the *M. fortuitum* group by exhibiting an intermediate amikacin MIC, resistance to kanamycin, a cefoxitin MIC of $\leqslant 32$ μg ml$^{-1}$ and the use of L-rhamnose as a carbon source (Wallace *et al.*, 2002). Moreover, comparison of enzymic activities (*M. mageritense* but not *M. wolinskyi* had a positive 3-day arylsulfatase reaction) and DNA–DNA relatedness suggested that these two groups were intermediate between the members of the *M. fortuitum* and thermotolerant *M. smegmatis* groups (Domenech *et al.*, 1997; Brown *et al.*, 1999; Wallace *et al.*, 2002).

Species of the *M. smegmatis* group (cluster VI) were thermotolerant (*M. smegmatis* and *M. goodii*) with the exception of *M. wolinskyi*, as previously described (Brown-Elliott & Wallace, 2002). This proposition is in good agreement in respect of DNA–DNA relatedness, general absence of susceptibility to new macrolides including clarithromycin and positive 3-day arylsulfatase reaction. However, *M. smegmatis* is highly susceptible to tobramycin (MIC$\leqslant 1$ μg ml$^{-1}$), whereas *M. goodii* has intermediate susceptibility (MIC 2–8 μg ml$^{-1}$) and *M. wolinskyi* is resistant to tobramycin (MIC$>8$ μg ml$^{-1}$) (Brown *et al.*, 1999). Further evidence was provided by a 3-bp deletion at positions 28–30 in the *rpoB* gene sequence and 2-bp insertions at positions 82–83 and 185–186 and a 1-bp insertion at position 202 and a 1-bp deletion at position 178 of the 16S rRNA gene sequence.

Since the initiation of our study, two species were described as emerging human RGM pathogens, *Mycobacterium boenickei* and *Mycobacterium brisbanense* (Schinsky *et al.*, 2004), which can be placed within existing clusters on the basis of their 16S rRNA gene sequences. It would be interesting to confirm these placements when using the molecular tools herein described. Among the genes we studied, the *rpoB* sequence-based relationships were in accordance with those published previously.

In conclusion, our study indicated that there was a robust and consistent phylogeny for the genomic backbone of housekeeping genes in bacteria. Because of the bootstrap values of the *rpoB*, *recA* and 16S rRNA gene sequences and their topology, all three genes could be used when describing a novel RGM species or when isolates are to be reported.

## ACKNOWLEDGEMENTS

We thank Christiane Bibard and Antoine Carlioz for technical assistance, Pierre Darlu (INSERM U535 Génétique épidémiologique et Structure des populations humaines) for access to PAUP4.0b and the ILD test and Babu V. A. K. Gundi for expert reviewing of the manuscript.

## REFERENCES

**Adékambi, T., Colson, P. & Drancourt, M. (2003).** *rpo*B-based identification of nonpigmented and late-pigmenting rapidly growing mycobacteria. *J Clin Microbiol* **41**, 5699–5708.

**Averof, M., Rokas, A., Wolfe, K. H. & Sharp, P. M. (2000).** Evidence for a high frequency of simultaneous double-nucleotide substitutions. *Science* **287**, 1283–1286.

**Blackwood, K. S., He, C., Gunton, J., Turenne, C. Y., Wolfe, J. & Kabani, A. M. (2000).** Evaluation of *rec*A sequences for identification of *Mycobacterium* species. *J Clin Microbiol* **38**, 2846–2852.

**Bromham, L., Penny, D., Rambaut, A. & Hendy, M. D. (2000).** The power of relative rates tests depends on the data. *J Mol Evol* **50**, 296–301.

**Brown, B. A., Springer, B., Steingrube, V. A. & 10 other authors (1999).** *Mycobacterium wolinskyi* sp. nov. and *Mycobacterium goodii* sp. nov., two new rapidly growing species related to *Mycobacterium smegmatis* and associated with human wound infections: a cooperative study from the International Working Group on Mycobacterial Taxonomy. *Int J Syst Bacteriol* **49**, 1493–1511.

**Brown-Elliott, B. A. & Wallace, R. J., Jr (2002).** Clinical and taxonomic status of pathogenic nonpigmented or late-pigmenting rapidly growing mycobacteria. *Clin Microbiol Rev* **15**, 716–746.

**Brunello, F., Ligozzi, M., Cristelli, E., Bonora, S., Tortoli, E. & Fontana, R. (2001).** Identification of 54 mycobacterial species by PCR-restriction fragment length polymorphism analysis of the *hsp65* gene. *J Clin Microbiol* **39**, 2799–2806.

**Bull, J. J., Huelsenbeck, J., Cunningham, C. W., Swofford, D. L. & Wadell, P. J. (1993).** Partitioning and combining data in phylogenetic analysis. *Syst Biol* **42**, 384–397.

**Chamoiseau, G. (1979).** Etiology of farcy in African bovines: nomenclatures of the causal organisms *Mycobacterium farcinogenes* Chamoiseau and *Mycobacterium senegalense* (Chamoiseau) comb. nov. *Int J Syst Bacteriol* **29**, 407–410.

**Cummings, M. P., Otto, S. P. & Wakeley, J. (1995).** Sampling properties of DNA sequence data in phylogenetic analysis. *Mol Biol Evol* **12**, 814–822.

**Cunningham, C. W. (1997).** Can three incongruence tests predict when data should be combined? *Mol Biol Evol* **14**, 733–740.

**Darlu, P. & Lecointre, G. (2002).** When does the incongruence length difference test fail? *Mol Biol Evol* **19**, 432–437.

**Domenech, P., Menendez, M. C. & Garcia, M. J. (1994).** Restriction fragment length polymorphisms of 16S rRNA genes in the differentiation of fast-growing mycobacterial species. *FEMS Microbiol Lett* **116**, 19–24.

**Domenech, P., Jimenez, M. S., Menendez, M. C., Bull, T. J., Samper, S., Manrique, A. & Garcia, M. J. (1997).** *Mycobacterium mageritense* sp. nov. *Int J Syst Bacteriol* **47**, 535–540.

**Drancourt, M., Bollet, C., Carlioz, A., Martelin, R., Gayral, J. P. & Raoult, D. (2000).** 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *J Clin Microbiol* **38**, 3623–3630.

**Fox, G. E., Wisotzkey, J. D. & Jurtshuk, P., Jr (1992).** How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* **42**, 166–170.

**Gatesy, J. & Arctander, P. (2000).** Hidden morphological support for the phylogenetic placement of *Pseudoryx nghetinhensis* with bovine bovids: a combined analysis of gross anatomical evidence and DNA sequences from five genes. *Syst Biol* **49**, 515–538.

**Gatesy, J., O'Grady, P. & Baker, R. (1999).** Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyls taxa. *Cladistics* **15**, 271–313.

**Gaunt, M. W., Turner, S. L., Rigottier-Gois, L., Lloyd-Macgilp, S. A. & Young, J. P. W. (2001).** Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *Int J Syst Evol Microbiol* **51**, 2037–2048.

**Hamid, M. E., Roth, A., Landt, O., Kroppenstedt, R. M., Goodfellow, M. & Mauch, H. (2002).** Differentiation between *Mycobacterium farcinogenes* and *Mycobacterium senegalense* strains based on 16S-23S ribosomal DNA internal transcribed spacer sequences. *J Clin Microbiol* **40**, 707–711.

**Jang, S. S. & Hirsh, D. C. (2002).** Rapidly growing members of the genus *Mycobacterium* affecting dogs and cats. *J Am Anim Hosp Assoc* **38**, 217–220.

**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.

**Kirschner, P., Kiekenbeck, M., Meissner, D., Wolters, J. & Böttger, E. C. (1992).** Genetic heterogeneity within *Mycobacterium fortuitum* complex species: genotypic criteria for identification. *J Clin Microbiol* **30**, 2772–2775.

**Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001).** MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

**Lanéelle, M. A., Silve, G., Lopez-Marin, L. M. & Daffe, M. (1996).** Structures of the glycolipid antigens of members of the third biovariant complex of *Mycobacterium fortuitum*. *Eur J Biochem* **15**, 270–279.

**Lévy-Frébault, V., Grimont, F., Grimont, P. A. D. & David, H. L. (1986).** Deoxyribonucleic acid relatedness study of the *Mycobacterium fortuitum-Mycobacterium chelonae* complex. *Int J Syst Bacteriol* **36**, 458–460.

**Lopez-Marin, L. M., Laneelle, M. A., Prome, D. & Daffe, M. (1993).** Structures of the glycopeptidolipid antigens of two animal pathogens: *Mycobacterium senegalense* and *Mycobacterium porcinum*. *Eur J Biochem* **215**, 859–866.

**Lorenz, M. G. & Sikorski, J. (2000).** The potential for intraspecific horizontal gene exchange by natural genetic transformation: sexual isolation among genomovars of *Pseudomonas stutzeri*. *Microbiology* **146**, 3081–3090.

**Muñoz, M., Julian, E., Garcia-Barcelo, M., Ausina, V. & Luquin, M. (1997).** Easy differentiation of *Mycobacterium mucogenicum* from other species of the *Mycobacterium fortuitum* complex by thin-layer and gas chromatography of fatty esters and alcohols. *J Chromatogr B Biomed Sci Appl* **689**, 341–347.

**Muñoz, M., Raynaud, C., Lanéelle, M.-A., Julián, E., López Marín, L. M., Silve, G., Ausina, V., Daffé, M. & Luquin, M. (1998).** Seroreactive species-specific lipooligosaccharides of *Mycobacterium mucogenicum* sp. nov. (formerly *Mycobacterium chelonae*-like organisms): identification and chemical characterization. *Microbiology* **144**, 137–148.

**Pfyffer, G. E., Brown-Elliot, B. A. & Wallace, R. J., Jr (2003).** *Mycobacterium*: general characteristics, isolation, and staining procedures. In *Manual of Clinical Microbiology*, 8th edn, pp. 532–559. Edited by P. R. Murray, E. J. Baron, J. H. Jorgensen, M. A. Pfaller & R. H. Yolken. Washington, DC: American Society for Microbiology.

**Ridell, M. & Goodfellow, M. (1983).** Numerical classification of *Mycobacterium farcinogenes*, *Mycobacterium senegalense* and related taxa. *J Gen Microbiol* **129**, 599–611.

**Rinquet, H., Akoua-Koffi, C., Honore, S., Varnerot, A., Vincent, V., Berche, P., Gaillard, J. L. & Pierre-Audigier, C. (1999).** *hsp65* sequencing for identification of rapidly growing mycobacteria. *J Clin Microbiol* **37**, 852–857.

**Roth, A., Reischl, U., Streubel, A., Naumann, L., Kroppenstedt, R. M., Habicht, M., Fischer, M. & Mauch, H. (2000).** Novel diagnostic algorithm for identification of mycobacteria using genus-specific amplification of the 16S-23S rRNA gene spacer and restriction endonucleases. *J Clin Microbiol* **38**, 1094–1104.

**Schinsky, M. F., McNeil, M. M., Whitney, A. M., Steigerwalt, A. G., Lasker, B. A., Floyd, M. M., Hogg, G. G., Brenner, D. J. & Brown, J. M. (2000).** *Mycobacterium septicum* sp. nov., a new rapidly growing species associated with catheter-related bacteraemia. *Int J Syst Evol Microbiol* **50**, 575–581.

**Schinsky, M. F., Morey, R. E., Steigerwalt, G. A. & 9 other authors (2004).** Taxonomic variation in the *Mycobacterium fortuitum* third biovariant complex: description of *Mycobacterium boenickei* sp. nov., *Mycobacterium houstonense* sp. nov., *Mycobacterium neworleansense* sp. nov. and *Mycobacterium brisbanense* sp. nov. and recognition of *Mycobacterium porcinum* from human clinical isolates. *Int J Syst Evol Microbiol* **54**, 1653–1667.

**Springer, B., Böttger, E. C., Kirschner, P. & Wallace, R. J., Jr (1995).** Phylogeny of the *Mycobacterium chelonae*-like organism based on partial sequencing of 16S rRNA gene and proposal of *Mycobacterium mucogenicum* sp. nov. *Int J Syst Bacteriol* **45**, 262–267.

**Steingrube, V. A., Gibson, J. L., Brown, B. A., Zhang, Y., Wilson, R. W., Rajagonpalan, M. & Wallace, R. J., Jr (1995).** PCR amplification and restriction endonuclease analysis of a 65-kilodalton heat shock protein gene sequence for taxonomic separation of rapidly growing mycobacteria. *J Clin Microbiol* **33**, 149–153.

**Swofford, D. L. (1998).** PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sunderland, MA: Sinauer Associates.

**Telenti, A., Marchesi, F., Balz, M., Bally, F., Böttger, E. & Bodmer, T. (1993).** Rapidly growing mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *J Clin Microbiol* **31**, 175–178.

**Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997).** The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.

**Tortoli, E. (2003).** Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. *Clin Microbiol Rev* **16**, 319–354.

**Tsukamura, M. & Ichiyama, S. (1986).** Numerical classification of rapidly growing nonphotochromogenic mycobacteria. *Microbiol Immunol* **30**, 863–882.

**Wallace, R. J., Jr, Silcox, V. A., Tsukamura, M., Brown, B. A., Kilburn, J. O., Butler, W. R. & Onyi, G. O. (1993).** Clinical significance, biochemical features, and susceptibility patterns of sporadic isolates of the *Mycobacterium chelonae*-like organism. *J Clin Microbiol* **31**, 3231–3239.

**Wallace, R. J., Jr, Brown-Elliot, B. A., Hall, L. & 9 other authors (2002).** Clinical and laboratory features of *Mycobacterium mageritense*. *J Clin Microbiol* **40**, 2930–2935.

**Wilson, R. W., Steingrube, V. A., Böttger, E. C. & 11 other authors (2001).** *Mycobacterium immunogenum* sp. nov., a novel species related to *Mycobacterium abscessus* and associated with clinical disease, pseudo-outbreaks and contaminated metalworking fluids: an international cooperative study on mycobacterial taxonomy. *Int J Syst Evol Microbiol* **51**, 1751–1764.

**Woese, C. R. (1987).** Bacterial evolution. *Microbiol Rev* **51**, 221–271.

**Youssef, S., Archambault, M., Parker, W. & Yager, J. (2002).** Pyogranulomatous panniculitis in a cat associated with infection by the *Mycobacterium fortuitum/peregrinum* group. *Can Vet J* **43**, 285–287.