Journal of Medical Microbiology (2007), **56**, 1033–1041 DOI 10.1099/jmm.0.47105-0

Identification of *Mycobacterium* using the EF-Tu encoding (*tuf*) gene and the tmRNA encoding (*ssrA*) gene

Sophie Mignard[1,2,3] and Jean-Pierre Flandrois[1,2,3]

Correspondence
Sophie Mignard
sophie.mignard@chu-lyon.fr

[1]Universite de Lyon, Universite de Lyon 1, CNRS, UMR5558, Biometrie et Biologie Evolutive, Villeurbanne F-69622, France

[2]Hospices Civils de Lyon, Lyon F-69001, France

[3]Laboratoire de Bacteriologie, Centre Hospitalier Lyon Sud, Hospices Civils de Lyon, Chemin du Grand Revoyet, Pierre-Benite 69495, France

The partial nucleotide sequences encoding the elongation factor Tu (*tuf* gene) (652 bp) and transfer-mRNA (tmRNA or *ssrA* gene) (340 bp) were determined to assess the suitability of these two genes as phylogenetic markers for the classification of mycobacteria, and thus as alternative target molecules for identifying mycobacteria. A total of 125 reference strains of the genus *Mycobacterium* and 74 clinical isolates were amplified by PCR and sequenced. Phylogenies of the two genes constructed by the neighbour-joining method were created and compared to a concatenated tree of 16S rDNA, *hsp65*, *sodA* and *rpoB* genes. The phylogenetic trees revealed the overall natural relationships among *Mycobacterium* species. The tmRNA phylogeny was similar to that of 16S rDNA, with low resolving power. The *tuf* gene provided better resolution of each mycobacterial species, with a phylogeny close to that of *hsp65*. However, none of these methods differentiated between the members of the *Mycobacterium tuberculosis* complex or the subspecies of the *Mycobacterium avium* complex. The correct identification of clinical isolates confirms the interest of these genes, especially *tuf*. It is suggested from these findings that tmRNA might be useful as another housekeeping gene in a polyphyletic approach to *Mycobacterium* species, but not as a first-line marker of species. *tuf* gene analysis suggests that this gene could be used effectively for phylogenetic analysis and to identify mycobacteria.

Received 30 November 2006
Accepted 14 March 2007

INTRODUCTION

The rising incidence of human infection by mycobacteria, especially among immunocompromised patients, is a serious public-health concern. Rapid identification is a high priority in order to control the infection and alternatives to traditional culture-based methods of identifying novel species or those that cannot be cultured have been developed. 16S rDNA-based phylogenetic analysis has contributed to establishing the systematics of *Mycobacterium* (Dobner *et al.*, 1996; Kirschner *et al.*, 1993; Kox *et al.*, 1995; Roth *et al.*, 1998; Telenti *et al.*, 1993). Other useful genotypic studies using protein-encoding genes *rpoB* (Kim *et al.*, 1999), *gyrB* (Kasai *et al.*, 2000), *hsp65* (Kim *et al.*, 2005; McNabb *et al.*, 2004; Ringuet *et al.*, 1999), *dnaJ* (Takewaki *et al.*, 1993), *recA* (Blackwood *et al.*, 2000; Van Soolingen *et al.*, 1997) and *sodA* (Zolg & Philippi-Schulz, 1994) have been published recently. However, all these methods have their limitations: for a few species no amplification is obtained, and some closely related species may not be distinguished, particularly in the case of the *Mycobacterium tuberculosis* complex (MTBC). It is still impossible to identify the full range of the *Mycobacterium* species using any single sequenced gene. It would be possible to optimize the identification strategy using only two targets, but this does not seem to provide a solution (Devulder *et al.*, 2005).

We have studied two other genes *ssrA* (encoding transfer-mRNA) and *tuf* [encoding elongation factor Tu (EF-Tu)] that are implicated in the *trans*-translation process (Haebel *et al.*, 2004; Withey & Friedman, 2003; Keiler *et al.*, 1996), highly conserved phenomenon among bacteria. Both these genes have already been used to reconstruct bacterial phylogeny (Picard *et al.*, 2004; Schonhuber *et al.*, 2001; Andersen *et al.*, 2006), but not so far for mycobacteria. We tested these newly identified genes encoding tmRNA and EF-Tu for phylogeny, and to demonstrate the feasibility of the phylogeny-driven identification process by applying the

Abbreviations: MTBC, *Mycobacterium tuberculosis* complex; RG, rapidly growing; SG, slowly growing.

A table of the reference strains used is available as supplementary material with the online version of this paper.

procedure to clinical isolates. The aim of the study was to assess the suitability of these two genes as phylogenetic markers for the classification of mycobacteria, and thus as alternative target molecules for identifying mycobacteria.

METHODS

Mycobacterial strains and clinical isolates. A total of 125 reference strains of the genus *Mycobacterium* and 74 clinical isolates used in this study were provided by Collection de l'Institut Pasteur, National Collection of Type Cultures, Deutsche Sammlung von Mikroorganismen und Zellkulturen and by Hospices Civils de Lyon (reference strains used are shown in Supplementary Table S1 available with the online journal). The DNA of *Mycobacterium leprae* was extracted from a clinical sample; it had been obtained and amplified from skin-punch specimens of a patient diagnosed with leprosy on the basis of clinical and histological findings, and the amplification of 16S rRNA. Clinical isolates were identified by our routine genomic method (*hsp65* sequencing) (McNabb *et al.*, 2004), and provided for the blinded analysis of the tmRNA and *tuf* genes.

Preparation of DNA and PCR. Mycobacterial DNA was prepared by heating to 100 °C, as described by Afghani & Stutman (1996). The amplification was done by a PCR general protocol in a 50 µl final volume as described below.

We designed the following primers for tmRNA amplification: r1 5′-TGG AGC TGC CGG GAA TCG AAC-3′ and r2 5′-GGG GCT GAA ACG GTT TCG-3′. For *tuf* amplification, the primers designed were T1 5′-CAC GCC GAC TAC ATC AAG AA-3′ and T2 5′-GAA CTG CGG ACG GTA GTT GT-3′. Primers were designed after comparison of aligned *tuf* and *ssrA* genes of published mycobacteria complete genomes. DNA (4 µl) was added to reaction mixtures containing (at the final concentrations shown): PCR master mix (Fermentas Life Sciences) (1.25 U *Taq* polymerase, 200 µM each dNTP and 2 mM $MgCl_2$), 1 mM $MgCl_2$ (Roche Diagnostics) and 500 nM each forward and reverse primer (MWG Biotech) and 4 % DMSO. The PCR program used consisted of an initial denaturing step at 95 °C for 3 min, 30 cycles of denaturing at 95 °C for 30 s, annealing at 57 °C for tmRNA and 55 °C for *tuf* for 30 s, and extension at 72 °C for 30 s, and a final extension step at 72 °C for 7 min. A sample of complete mix and DNA-free water was used as a negative control. PCR products were electrophoresed in a 1.5 % agarose gel, and visualized with ethidium bromide (Eurogentec) under UV light to check the correct size of the product amplified.

The sequencing step was performed by a service provider (Biofidal) where the amplicons were purified.

Sequence analysis. Sequences were aligned using MUSCLE (Edgar, 2004), and then any non-matching ends were deleted to obtain a homogeneous set and thus increase the reliability of the tree obtained. Similarity and distance matrices were computed using Clustal (data not shown) (Saitou & Nei, 1987; Thompson *et al.*, 1997). The tmRNA and *tuf* sequences of *Nocardia farcinica* were extracted from GenBank (accession no. NC_006361) to constitute the external outgroup. Phylogenetic trees were inferred using the neighbour-joining method (Saitou & Nei, 1987) with the Kimura two-parameter distance (Kimura, 1980) global gap removal option performed by Phylo_win (Galtier *et al.*, 1996), and using *Nocardia* as an outgroup. The resulting trees were evaluated by bootstrap analysis (Felsenstein, 1993) based on 1000 resamplings. The GenBank accession nos of *ssrA* and *tuf* are shown in Supplementary Table S1 available with the online journal.

Clinical isolates. Seventy-four isolates were provided without information from the Clinical Microbiology Laboratory of Lyon-Sud Hospital, Hospices Civils de Lyon. They were all identified by *hsp65*, *tuf* and *ssrA* sequencing. A profile alignment of the unidentified sequence and the previously aligned set had been done using MUSCLE, and a phylogenetic tree was then constructed using Clustal and plotted using NJplot (Saitou & Nei, 1987). The identification was deduced from the phylogenetic position and from the computed identities between sequences. A comparison of the identities deduced from *hsp65* versus tmRNA, and from *hsp65* versus *tuf* was carried out in order to validate identities given by the two genes studied (Table 1). Intraspecific variations in the sequence similarities and phylogenies of tmRNA and *tuf* genes among the clinical isolates of the same species were investigated by constructing similarity matrices and species-only phylogenetic trees (data not shown).

Phylogeny analysis. We compared the phylogenies of the new genes *ssrA* and *tuf*. We noted the robustness of the tree evaluated with bootstrap values of more than 75 %, the good resolving power between species, and the clear distinction between slowly growing (SG) and rapidly growing (RG) species. The congruence of the new trees was evaluated by comparison with the concatenated tree of the 16S rDNA, *hsp65*, *sodA* and *rpoB* genes (Devulder *et al.*, 2005), as this had been shown to be the best way to proceed with the higher bootstrap values. Only 103 strains were included because datasets for these 4 genes are incomplete (*sodA* displays a particularly high level of variance and this has not been amplified successfully). We had to complete Devulder's database with the new strains described since 2004, by inserting 14 more species (*Mycobacterium pinnipedii* had not been added) and create a new concatenated tree (Fig. 1).

RESULTS

The tmRNA DNAs (354 bp) from 124 isolates corresponding to a total of 125 reference strains of mycobacteria were amplified, sequenced and compared (the amplification of *Mycobacterium genavense* was unsuccessful). The *tuf* DNAs (652 bp) from 120 reference strains were successfully amplified. The five strains that were not amplified were *Mycobacterium botniense*, *Mycobacterium madagascariense*, *Mycobacterium hodleri*, '*Mycobacterium yunnanensis*' and *M. genavense*. The mean mol % G+C content was 61.7 % for tmRNA and 62.4 % for *tuf*. Some insertions were observed in tmRNA: a 21 bp insertion for three strains belonging to the *Mycobacterium terrae* complex (*M. terrae*, *Mycobacterium nonchromogenicum* and *Mycobacterium hiberniae*) at position 79 (*M. tuberculosis* H37Rv *ssrA* gene numbering, accession no. NC_000962.2), insertion of a codon in *Mycobacterium aubagnense*, *Mycobacterium mucogenicum*, *Mycobacterium diernhoferi*, *Mycobacterium fluoranthenivorans*, '*Mycobacterium hackensackense*', *Mycobacterium phocaicum* and *Mycobacterium fredericksbergense* at position 111 (*M. tuberculosis* H37Rv *ssrA* gene numbering). An additional band of amplification (1000 bp) was observed in *M. tuberculosis*, *Mycobacterium bovis*, *Mycobacterium africanum*, *Mycobacterium nebraskense* and *Mycobacterium gordonae*, due to the forward primer, which hybridized in the following gene (hypothetical protein Rv308c) from position 3 467 263 to 3 467 252 of *M. tuberculosis* H37rV genome (accession no. NC_000962.2). In the *tuf* gene, an insertion was observed for the *Mycobacterium abcessus*/*Mycobacterium chelonae*/*Mycobacterium*

Table 1. Comparison of identification of strains isolated from clinical specimens

The reference identification is given by *hsp65* sequencing. Identification given by tmRNA and *tuf* sequencing analysis is only based on the phylogenetic position.

Primary identification* (*n*)	tmRNA identification (*n*)	*tuf* identification (*n*)
MTBC (23)	MTBC (23)	MTBC (23)
Mycobacterium avium (8)	*M. avium* (8)	*M. avium* (8)
Mycobacterium chimaera (4)	*M. chimaera* or *M. intracellulare* (4)†	*M. chimaera* (4)
Mycobacterium gordonae (6)	*M. gordonae* (6)	*M. gordonae* (6)
Mycobacterium fortuitum (6)	*M. fortuitum* (6)	*M. fortuitum* (6)
Mycobacterium lentiflavum (2)	*M. lentiflavum* (2)	*M. lentiflavum* (2)
Mycobacterium abcessus (3)	*M. abcessus* (3)	*M. abcessus* (3)
Mycobacterium xenopi (7)	*M. xenopi* (7)	*M. xenopi* (7)
Mycobacterium kansasii (2)	*M. kansasii V* (1)	*M. kansasii V* (1)
	M. kansasii I (1)	*M. kansasii I* (1)
Mycobacterium szulgai (1)	*M. szulgai* (1)	*M. szulgai* (1)
Mycobacterium marinum (5)	*M. marinum* or *M. ulcerans* or *M. pseudoshottsii* (5)†	*M. ulcerans* or *M. pseudoshottsii* (5)†
Mycobacterium chelonae (1)	*M. chelonae* (1)	*M. chelonae* (1)
Mycobacterium novocastrense (1)	*M. novocastrense* (1)	*M. novocastrense* (1)
Mycobacterium intracellulare (2)	*M. intracellulare* or *M. chimaera* (2)	*M. intracellulare* (2)
Mycobacterium scrofulaceum (1)	*M. scrofulaceum* or*Mycobacterium parascrofulaceum* (1)†	*M. scrofulaceum* (1)
Mycobacterium spp. (2)	*Mycobacterium* spp. (2)	*Mycobacterium* spp. (2)

*Identification based on *hps65* sequencing performed in the Laboratoire de Bacteriologie, Hospices Civils de Lyon, France.
†The sequence could not discriminate the species.

bolletii/*Mycobacterium immunogenum* cluster (position 502 in the *M. tuberculosis* H37Rv *tuf* gene sequence).

For each gene, the sequences were compared pairwise for similarity. The mean percentage of similarity was 90.7 % for tmRNA and 90.2 % for *tuf*. The overall amino acid sequence identities of *tuf* ranged from 89 to 100 %. The phylogenetic trees had 24 and 34 % of nodes with a bootstrap value of more than 75 % for tmRNA and *tuf*, respectively (Figs 2 and 3).

The global phylogenies of each gene matched the conventional topology, with a division between RG and SG mycobacteria (Stahl & Urbance, 1990). The RG species seem to be older than the SG species, since the branching point of RG is deeper (assuming a constant rate of evolution). Nevertheless, the tmRNA phylogeny was of poor quality, with 26 strains mispositioned (SG among the RG or vice versa) (Fig. 2). The taxonomic position of the *M. terrae* complex is surprising in both phylogenies, but matches that described by Kim *et al.* (2005). However, *Mycobacterium triviale* was not correctly positioned: in the concatenated tree this species is clustered with the *M. terrae* complex where it belongs, whereas in tmRNA and *tuf* phylogenies it forms a single branch closer to the RG species.

With regard to the resolving power of the two genes, if we disregard the MTBC, *Mycobacterium avium* complex and *Mycobacterium fortuitum* complex, the tmRNA phylogenetic tree exhibits 11 clusters of identical sequences, including 24 species (Fig. 2). For *tuf*, five clusters of identical sequences were detected, including ten species. Most species showed good separation with *tuf*, which clearly has greater resolving power than tmRNA.

The congruence of each tree was determined, and the grouping of strict pathogens was characteristic for the MTBC (all sequences were identical). Clustering was observed for *M. abcessus*/*M. chelonae*/*M. immunogenum*/*M. bolletii*, *Mycobacterium haemophilum* close to *M. leprae*, the *M. avium* complex, *Mycobacterium intracellulare* and *Mycobacterium chimaera*. *Mycobacterium kansasii* was distinguished from *Mycobacterium gastri* by the two genes, but *tuf* provided a better resolution of the *M. kansasii* complex. The *Mycobacterium marinum* group strains were clearly clustered, but *tuf* provided a better distinction between species, with *Mycobacterium pseudoshottsii* closer to *Mycobacterium ulcerans* than to *M. marinum*. Each gene displayed congruence similar to the concatenated tree, with a mispositioning of the *M. terrae* complex.

For identification of clinical isolates we applied a direct sequencing protocol targeting tmRNA and *tuf* of 74 clinical isolates. By referring to the phylogenetic trees constructed using 125 reference strains we were able to identify all 74 isolates to species level. The findings matched those obtained by our routine technique (*hsp65* sequencing). tmRNA analysis was unable to distinguish between some strains, notably *M. marinum* and *M. ulcerans* or *M. intracellulare* and *M. chimaera*, species that are often

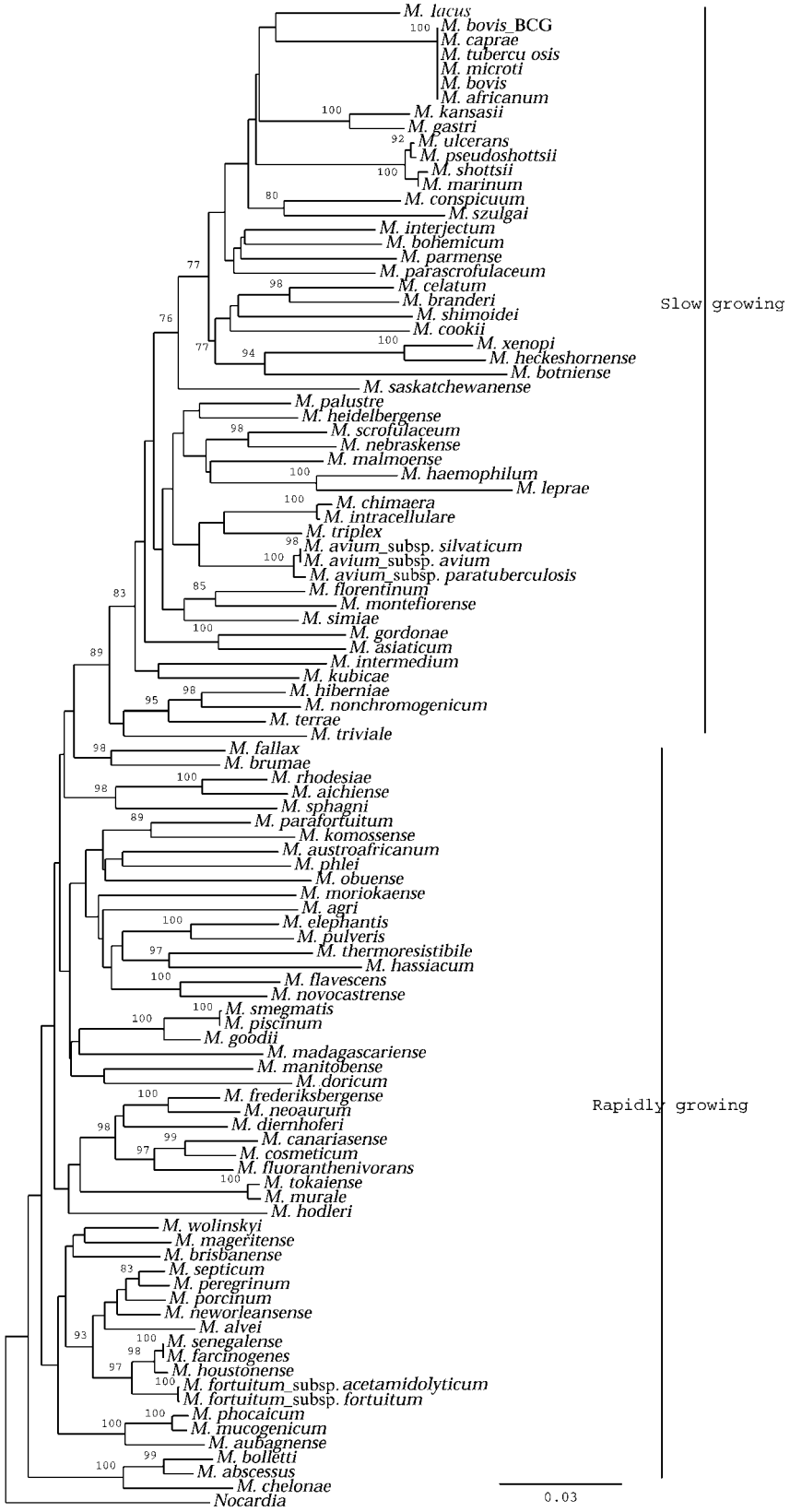

Fig. 1. Phylogenetic tree based on the concatenation of 16S rDNA, *hsp65*, *sodA* and *rpoB* genes by the neighbour-joining method and Kimura's two parameters model. Bootstrap values are indicated with numerical values in the tree. Values <75% are not shown. The tree is rooted using *Nocardia*.

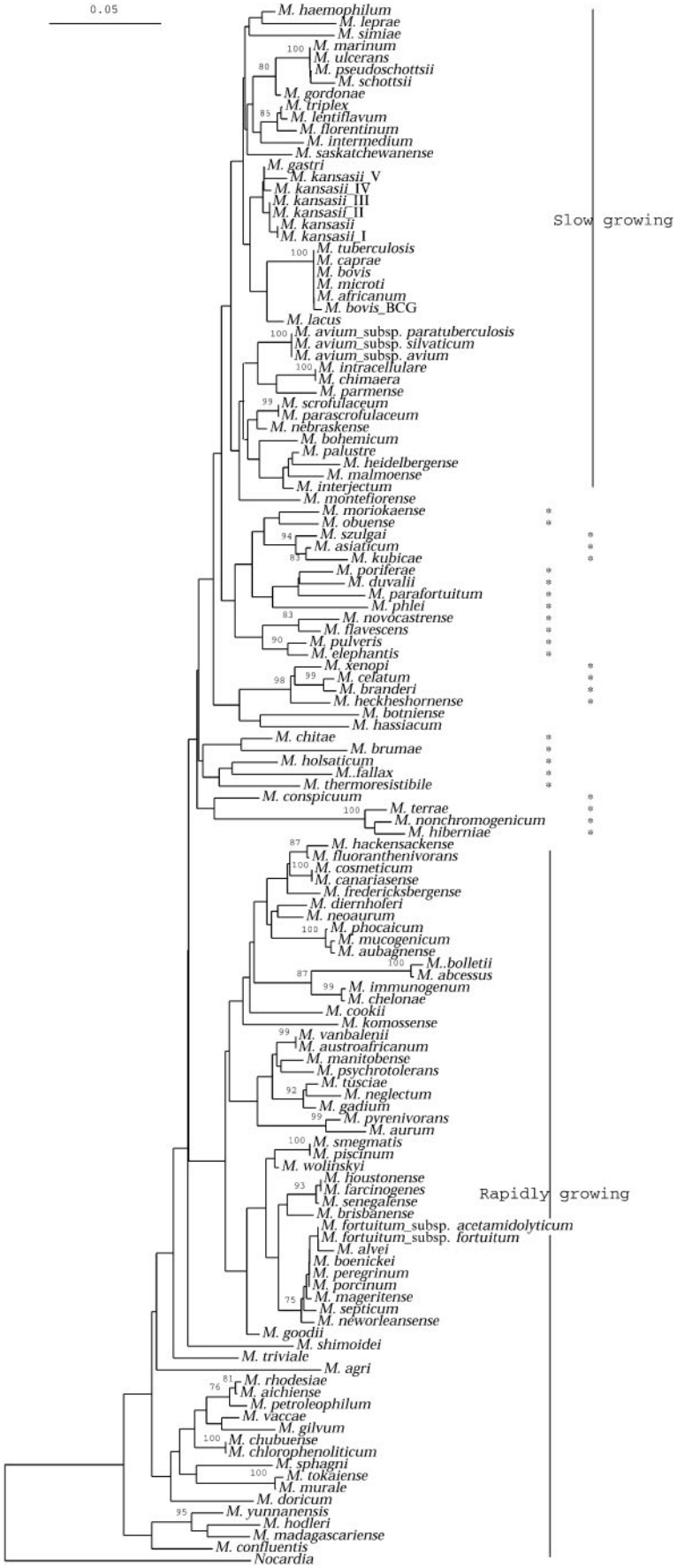

Fig. 2. Phylogenetic tree based on tmRNA sequences by the neighbour-joining method and Kimura's two parameters model. Bootstrap values are indicated with numerical values in the tree. Values <75 % are not shown. The tree is rooted using *Nocardia*. An asterisk indicates a mispositioning of the strain (an SG species among the RG species or the inverse).

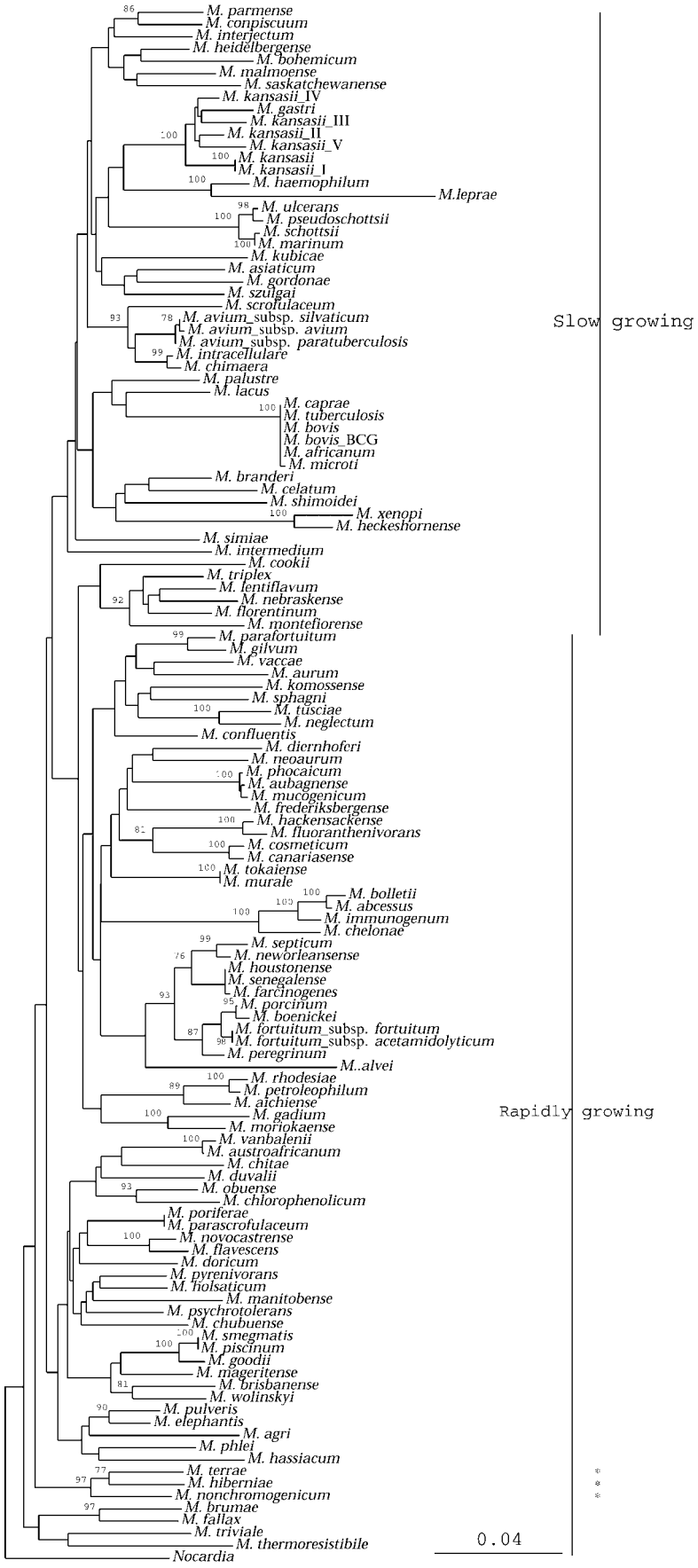

Fig. 3. Phylogenetic tree based on *tuf* sequences by the neighbour-joining method and Kimura's two parameters model. Bootstrap values are indicated with numerical values in the tree. Values $<75\%$ are not shown. The tree is rooted using *Nocardia*. An asterisk indicates a mispositioning of the strain (an SG species among the RG species or the inverse).

isolated from clinical specimens. *tuf* analysis of clinical specimens was better, the intraspecies percentage similarity ranging from 98.2 to 100 %. The main shortcoming was the poor distinction between clinical strains of *M. marinum*, which was attributable to the very small number of differences within the cluster containing *M. marinum*, *M. ulcerans*, *Mycobacterium shottsii* and *M. pseudoshottsii*.

DISCUSSION

This study assessed the feasibility of sequencing the EF-Tu and tmRNA encoding genes as suitable phylogenetic markers for identifying mycobacteria. The phylogenies of these genes are concordant with the distinctions based on growth rates, and the clustering of pathogenic species obtained by other methods (Adekambi & Drancourt, 2004; Kim *et al.*, 1999; McNabb *et al.*, 2004).

We assessed the characteristics that are usually required in phylogenetic studies. Firstly *tuf* and tmRNA are both single copy genes, which facilitates interpretation (duplication of the *tuf* gene is common among Gram-negative bacteria, but has not been reported for high mol % GC Gram-positive bacteria) (Kim *et al.*, 1999, 2005; Schonhuber *et al.*, 2001; Sela *et al.*, 1989; Zwieb *et al.*, 1999).

Good separation between closely related species is another constant requirement for phylogenetic studies, and this depends on the level of variability of the gene targeted. A high level of variability might be helpful for good discrimination between species, but it can also be a disadvantage, because of instability of species-specific signatures, and difficulties in developing reliable primers or probes (Kim *et al.*, 1999; Roth *et al.*, 1998). Genes do not all have the same resolving power. Despite high inter-species variability (9.3 %), tmRNA does not provide better resolution than 16S rDNA. *tuf* fulfils better the prerequisites for reliable phylogeny, with a resolving power close to those of the *hsp65* and *rpoB* genes. Moreover, it is useful for discriminating between closely related species, such as *M. kansasii* and *M. gastri* for instance. The *tuf* analysis supports the heterogeneity of *M. kansasii* revealed by molecular genotyping (Kim *et al.*, 2005; Picardeau *et al.*, 1997). However, as a consequence of this high level of variability, the higher the discriminating power, the higher the proportion of strains that are not amplified using a simple primer pair (Roth *et al.*, 1998). This might constitute a limitation for the use of the *tuf* gene, and indeed 5 out of the 125 reference strains were not amplified.

Congruence with phylogenies already described constitutes an important point in evaluating a new gene marker. For both the *tuf* and tmRNA genes, the congruence of pathogenic species was maintained. The 21 bp insertion in tmRNA in the *M. terrae* complex could be used as a molecular signature, and supports the clustering of the complex. *M. triviale* does not have this insertion, but has previously been shown to be slightly different, having less similarity with the other members of the *M. terrae* complex (Adekambi & Drancourt, 2004; Stahl & Urbance, 1990). We noted in this study that *M. hiberniae*, even though it has not been described as a member of the *M. terrae* complex, does in fact fall within this complex in all the phylogenies used in this work (Kazda *et al.*, 1993).

Another characteristic required for establishing the phylogeny of mycobacteria is the ability to distinguish between MTBC subspecies. Unfortunately, neither analysis of the *tuf* sequence nor of tmRNA was able to differentiate between the members of the MTBC, the subspecies of the *M. avium* complex or those of the *M. fortuitum* group. In fact few genes can achieve these separations (Blackwood *et al.*, 2000; Kim *et al.*, 2005; Takewaki *et al.*, 1993).

Clinical isolate studies are necessary to evaluate the potential of the method for bacterial identification. Correct identification depends on the degree of sequence homology between the isolate and the type strain. But the strain's position in the phylogenetic tree is another major factor. Attempts to identify strains in clinical isolates using tmRNA were not entirely satisfactory, mainly due to some ambiguous results due to sequence matches with two or more species. *tuf* results match those obtained by *hsp65* sequencing more closely. However, the same concern was observed for *tuf* with regard to the *M. marinum* strains. Surprisingly, *tuf* gene sequences of five routine strains misidentified them as being closer to *M. ulcerans* than to *M. marinum*. *M. marinum* and *M. ulcerans* are very closely related species, and neither ITS nor 16S rDNA sequences were able to distinguish between them (Portaels *et al.*, 1997; Roth *et al.*, 1998). *tuf* distinguishes well between the two references strains, and although we cannot rule out the possibility that the five routine strains constitute a variant closer to *M. ulcerans*, this is further evidence of the 'fuzziness' of the *M. marinum* group. The use of gene sequence as a routine means of identifying mycobacteria calls for a comprehensive database able to determine the percentage similarity between clinical isolates and validated species, and able to create a phylogenetic tree including the query sequence. We have therefore constructed *tuf* and *ssrA* databases, which are now included in BIBI database (http://umr5558-sud-str1.univ-lyon1.fr/lebibi/lebibi.cgi) (Devulder *et al.*, 2003).

The main limitation of our study is that only one strain from each species was tested (Blackwood *et al.*, 2000). The partial DNA sequence of a single gene cannot reflect the phylogenetic relationships between all mycobacteria, since only a small portion of the genome was used in this study. We included as many species as possible in order to reflect the phylogeny in as up-to-date a manner as possible [125 reference strains were tested out of the 127 species with standing in nomenclature (http://www.bacterio.cict.fr/m/mycobacterium.html) as of January 2006]. The inter- and intra-specific variability is therefore difficult to assess with limited data.

Conclusion

In conclusion, this study shows that *tuf* could be a good alternative molecular marker not only for phylogenetic analysis, but also for species identification of mycobacterial clinical isolates. It should be applied to phylogeny as a first-line genomic technique, as it displays interspecies divergence similar to that of *hsp65* or *rpoB*. tmRNA is not suitable for immediate phylogeny but it could be useful for some species that it distinguishes well. *tuf* and tmRNA genes should be included as housekeeping genes for a polyphasic approach to taxonomic analysis at and above the species level.

ACKNOWLEDGEMENTS

We are grateful to G. Fardel for technical assistance and to G. Carret for a critical review of the manuscript.

REFERENCES

Adekambi, T. & Drancourt, M. (2004). Dissection of phylogenetic relationships among 19 rapidly growing *Mycobacterium* species by 16S rRNA, *hsp65, sodA, recA, rpoB* gene sequencing. *Int J Syst Evol Microbiol* **54**, 2095–2105.

Afghani, B. & Stutman, H. R. (1996). Polymerase chain reaction for diagnosis of *M. tuberculosis*: comparison of simple boiling and a conventional method for DNA extraction. *Biochem Mol Med* **57**, 14–18.

Andersen, E. S., Rosenblad, M. A., Larsen, N., Westergaard, J. C., Burks, J., Wower, I. K., Wower, J., Gorodkin, J., Samuelsson, T. & Zwieb, C. (2006). The tmRDB and SRPDB resources. *Nucleic Acids Res* **34**, D163–D168.

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

Devulder, G., Perriere, G., Baty, F. & Flandrois, J. P. (2003). BIBI, a bioinformatics bacterial identification tool. *J Clin Microbiol* **41**, 1785–1787.

Devulder, G., Pérouse 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.

Dobner, P., Feldmann, K., Rifai, M., Loscher, T. & Rinder, H. (1996). Rapid identification of mycobacterial species by PCR amplification of hypervariable 16S rRNA gene promoter region. *J Clin Microbiol* **34**, 866–869.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), version 3.5c. Department of Genome Sciences, University of Washington, Seattle, WA, USA.

Galtier, N., Gouy, M. & Gautier, C. (1996). SeaView and Phylo_win: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* **12**, 543–548.

Haebel, P. W., Gutmann, S. & Ban, N. (2004). Dial tm for rescue: tmRNA engages ribosomes stalled on defective mRNAs. *Curr Opin Struct Biol* **14**, 58–65.

Kasai, H., Ezaki, T. & Harayama, S. (2000). Differentiation of phylogenetically related slowly growing mycobacteria by their *gyrB* sequences. *J Clin Microbiol* **38**, 301–308.

Kazda, J., Cooney, R., Monaghan, M., Quinn, P. J., Stackebrandt, E., Dorsch, M., Daffe, M., Muller, K., Cook, B. R. & Tarnok, Z. S. (1993). *Mycobacterium hiberniae* sp.nov. *Int J Syst Bacteriol* **43**, 352–357.

Keiler, K. C., Waller, P. R. & Sauer, R. T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990–993.

Kim, B. J., Lee, S. H., Lyu, M. A., Kim, S. J., Bai, G. H., Kim, S. J., 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, H., Kim, S. H., Shim, T. S., Kim, M., Bai, G. H., Park, Y. G., Lee, S. H., Chae, G. T., Cha, C. Y. & other authors (2005). Differentiation of *Mycobacterium* species by analysis of the heat-shock protein 65 gene (*hsp65*). *Int J Syst Evol Microbiol* **55**, 1649–1656.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotides sequences. *J Mol Evol* **16**, 111–120.

Kirschner, P., Springer, B., Vogel, U., Meier, A., Wrede, A., Kiekenbeck, M., Bange, F. C. & Bottger, E. C. (1993). Genotypic identification of mycobacteria by nucleic acid sequence determination: report of a 2-year experience in a clinical laboratory. *J Clin Microbiol* **31**, 2882–2889.

Kox, L. F. F., Van Leeuwen, J., Jansen, H. M. & Kolk, A. H. J. (1995). PCR assay based on DNA coding for 16S rRNA for detection and identification of mycobacteria in clinical samples. *J Clin Microbiol* **33**, 3225–3233.

McNabb, A., Eisler, D., Adie, K., Amos, M., Rodrigues, M., Stephens, G., Black, W. A. & Isaac-Renton, J. (2004). Assessment of partial sequencing of the 65-kilodalton heat shock protein gene (*hsp65*) for routine identification of *Mycobacterium* species isolated from clinical sources. *J Clin Microbiol* **42**, 3000–3011.

Picard, F. J., Ke, D., Boudreau, D. K., Boissinot, M., Huletsky, A., Richard, D., Ouellette, M., Roy, P. H. & Bergeron, M. G. (2004). Use of *tuf* sequences for genus-specific PCR detection and phylogenetic analysis of 28 streptococcal species. *J Clin Microbiol* **42**, 3686–3695.

Picardeau, M., Prod'hom, G., Raskine, L., LePennec, M. P. & Vincent, V. (1997). Genotypic characterization of five subspecies of *Mycobacterium kansasii. J Clin Microbiol* **35**, 25–32.

Portaels, F., Aguiar, J., Fissette, K., Fonteyne, P. A., de Beenhouwer, H., de Rijk, P., Guedenon, A., Leman, R., Steunou, C. & other authors (1997). Direct detection and identification of *Mycobacterium ulcerans* in clinical specimens by PCR and oligonucleotide-specific capture plate hybridization. *J Clin Microbiol* **35**, 1097–1100.

Ringuet, 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., Fisher, 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 spaces sequences. *J Clin Microbiol* **36**, 139–147.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Schonhuber, W., le Bourhis, G., Tremblay, J., Amann, R. & Kulakauskas, S. (2001). Utilization of tmRNA sequences for bacterial identification. *BMC Microbiol* **1**, 20.

Sela, S., Yogev, D., Razin, S. & Bervovier, H. (1989). Duplication of the *tuf* gene: a new insight into the phylogeny of eubacteria. *J Bacteriol* **171**, 581–584.

Stahl, D. A. & Urbance, J. W. (1990). The division between fast- and slow-growing species corresponds to natural relationships among the mycobacteria. *J Bacteriol* **172**, 116–124.

Takewaki, S. I., Okuzumi, K., Ishiko, H., Nakahara, K. I., Ohkubo, A. & Nagai, R. (1993). Genus-specific polymerase chain reaction or the mycobacterial *dnaJ* gene and species-specific oligonucleotide probes. *J Clin Microbiol* **31**, 446–450.

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.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmongin, 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.

Van Soolingen, D., Hoogenboezem, T., de Haas, P. E. W., Hermans, P. W. M., Koedam, M. A., Teppema, K. S., Brennan, P. J., Besra, G. S., Portaels, F. & other authors (1997). A novel pathogenic taxon of the *Mycobacterium tuberculosis* complex, Canetti: characterization of an exceptional isolate from Africa. *Int J Syst Bacteriol* **47**, 1236–1245.

Withey, J. H. & Friedman, D. I. (2003). A salvage pathway for protein synthesis: tmRNA and *trans*-translation. *Annu Rev Microbiol* **57**, 101–123.

Zolg, J. W. & Philippi-Schulz, S. (1994). The superoxide dismutase gene as a target for detection and identification of mycobacteria by PCR. *J Clin Microbiol* **32**, 2801–2812.

Zwieb, C., Wower, I. & Wower, J. (1999). Comparative sequence analysis of tmRNA. *Nucleic Acids Res* **27**, 2063–2071.