Mycobacterium celeriflavum sp. nov., a rapidly growing scotochromogenic bacterium isolated from clinical specimens

Abdolrazagh Hashemi Shahraki,[1] Cengiz Çavuşoğlu,[2] Emanuele Borroni,[3] Parvin Heidarieh,[4] Orhan Kaya Koksalan,[5] Andrea Maurizio Cabibbe,[3] Mohamad Hashemzadeh,[6] Alessandro Mariottini,[7] Ehsan Mostafavi,[1] Davide Cittaro,[8] Mohamad Mehdi Feizabadi,[9] Dejan Lazarevic,[8] Farhad Yaghmaei,[1] Gian Lorenzo Molinari,[10] Anna Camaggi[10] and Enrico Tortoli[3]

Correspondence
Enrico Tortoli
e.tortoli@libero.it

[1]Department of Epidemiology, Pasteur Institute of Iran, Tehran, Iran

[2]Department of Medical Microbiology, Ege University, Izmir, Turkey

[3]Emerging Bacterial Pathogens Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy

[4]Department of Bacteriology and Virology, Alborz University of Medical Science, Karaj, Iran

[5]Institute for Medical Experimental Research, Istanbul University, Istanbul, Turkey

[6]Infectious and Tropical Diseases Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

[7]Diagnostic Genetics Unit, Careggi University Hospital, Florence, Italy

[8]Center for Translational Genomics and Bioinformatics, IRCCS San Raffaele Scientific Institute, Milan, Italy

[9]Department of Microbiology, Tehran University of Medical Science, Tehran, Iran

[10]Microbiology Laboratory, Ospedale Maggiore Carità, Novara, Italy

Six strains of a rapidly growing scotochromogenic mycobacterium were isolated from pulmonary specimens of independent patients. Biochemical and cultural tests were not suitable for their identification. The mycolic acid pattern analysed by HPLC was different from that of any other mycobacterium. Genotypic characterization, targeting seven housekeeping genes, revealed the presence of microheterogeneity in all of them. Different species were more closely related to the test strains in various regions: the type strain of *Mycobacterium moriokaense* showed 99.0 % 16S rRNA gene sequence similarity, and 91.5–96.5 % similarity for the remaining six regions. The whole genome sequences of the proposed type strain and that of *M. moriokaense* presented an average nucleotide identity (ANI) of 82.9 %. Phylogenetic analysis produced poorly robust trees in most genes with the exception of *rpoB* and *sodA* where *Mycobacterium flavescens* and *Mycobacterium novocastrense* were the closest species. This phylogenetic relatedness was confirmed by the tree inferred from five concatenated genes, which was very robust. The polyphasic characterization of the test strains, supported by the ANI value, demonstrates that they belong to a previously unreported species, for which the name *Mycobacterium celeriflavum* sp. nov. is proposed. The type strain is AFPC-000207^T (=DSM 46765^T=JCM 18439^T).

Abbreviation: ANI, average nucleotide identity.

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequences of strains AFPC-000207^T, FI-09258 and FI-10161 are KJ607136, HM770867 and KJ586590, respectively; those for the *hsp65* sequences are KJ586615, KJ586613 and KJ586614, respectively; and those for the *rpoB* sequences are KJ607137, HM807427 and KJ586585, respectively. The GenBank/EMBL/DDBJ accession numbers for the *sodA* sequences of strains AFPC-000207^T, FI-09258, FI-10161 and E498 are KJ586625, KJ586627, KJ586628 and KM396308, respectively; and those for the *rpoBC* sequences are KJ586618, KJ586621, KJ586622 and KM396307, respectively. The GenBank/EMBL/DDBJ accession numbers for the *gyrB* sequences of strains AFPC-000207^T and FI-09258 are KJ586606 and KJ586609, respectively; and those for the *DnaK* sequences are KJ586601 and KJ586604, respectively.

Two supplementary tables and three supplementary figures are available with the online Supplementary Material.

Table 1. Microbiological features of the novel strains and patient information

All strains were from sputum samples. COPD, Chronic obstructive pulmonary disease; ACER, association of amikacin, clarithromycin, ethambutol and rifampicin.

Strain	Patient age (years)	Sex	Microscopy	Culture*	Disease	Year	Locality	Treatment	Follow-up
AFPC-000207^T	44	M	+	4/4	COPD	2010	Ahwaz (Iran)	Anti-TB standard replaced by ACER	Resolved
AFPC-00088	59	F	+	3/3	Bronchiectasis	2011	Ahwaz (Iran)	ACER	Resolved
E498	5	M	+	1/2	Recurrent parotitis	2002	Izmir (Turkey)	None	Resolved
13DK204	26	F	+	1/6	Tuberculosis (TB)	2013	Edirne (Turkey)	Anti-TB standard	Resolved
FI-09258	73	M	−	1/1	Lung carcinoma	2009	Novara (Italy)	None	Death
FI-10161	63	M	−	1/1	Unknown	2010	Florence (Italy)	None	Unknown

*No. of positive cultures/total no. of cultures.

Non-tuberculous mycobacteria are typically environmental organisms that can cause opportunistic infections in humans and animals (Tortoli, 2003). Among them, the rapidly growing species have been long considered non-pathogenic but, in recent years, they have been frequently reported to be responsible for disease, in particular of cutis and soft tissues and of bone and joints, and also of the lung (Tortoli, 2009). Six strains of mycobacteria characterized by yellow, scotochromogenic pigmentation were isolated from respiratory specimens of the same number of independent patients. Their characterization, conducted using a polyphasic approach including biochemical and cultural tests, mycolic acid analysis, multi-locus sequence investigation and whole genome sequencing, led us to infer that the strains belong to a hitherto unrecognized species of the genus *Mycobacterium*.

Two strains, AFPC-000207^T and AFPC-00088, were isolated from independent patients, 1 year apart, in the same city (Ahvaz) of Iran. Strains E498 (Cavuşoğlu & Tortoli, 2006) and 13DK204 were isolated, over 12 years, from patients living in geographically distant cities in Turkey. Two further strains (FI-09258 and FI-10161) were isolated from unrelated Italian patients. Microbiological features of the strains and clinical/epidemiological information of the patients are reported in Table 1.

Major biochemical tests recommended for the identification of mycobacteria were performed as described by Kent & Kubica (1985); they included niacin accumulation, nitrate reduction, Tween 80 hydrolysis (10 days), urease, β-glucosidase, tellurite reduction and catalase. The six strains showed negative results for the majority of the tests performed. Only thermostable catalase, nitrate reduction and tellurite reduction were uniformly positive (Table 2). The strains grew rough, yellow, scotochromogenic colonies on Löwenstein–Jensen medium after 5–7 days at 37 °C. Growth was slower at 30 °C and was absent at 42 °C. No colonies developed on MacConkey agar without crystal violet or on media supplemented with *p*-nitrobenzoate (500 μg ml^{-1}), hydroxylamine (500 μg ml^{-1}) or isoniazid (1 μg ml^{-1}). All the strains grew in the presence of thiacetazone (10 μg ml^{-1}).

For HPLC of cell-wall mycolic acids, colonies were grown for 7 days at 37 °C on Middlebrook 7H11 agar and the mycolic acids were saponified with KOH (25 % in H_2O), extracted with chloroform, derivatized according to the manufacturer's instructions (MIDI) and loaded onto an Agilent ChemStation HPLC (Agilent Technologies). Mycolic acids were separated with a gradient of methanol and 2-propanol (starting ratio 75 : 25 %, end ratio 95 : 5 %) and analysed using the software Sherlock, version Myco 1.0, and the database MICAG1 1.02. The system identified the strains as representing *Mycobacterium flavescens* (chromatotype 2) with a low similarity index (<40 %). The profile was characterized by a continuous series of peaks eluting between 2.7 and 8.7 min, with the most prominent clustering before 4 min and around 8 min (Fig. 1). No similar

Table 2. Biochemical test results

Test	Result
Niacin accumulation	Negative
Nitrate reduction	Positive
Tween 80 hydrolysis (10 days)	Negative
Urease	Variable
β-Glucosidase	Negative
Tellurite reduction	Positive
Thermostable catalase (68 °C)	Positive
Catalase (semi-quantitative)	Variable

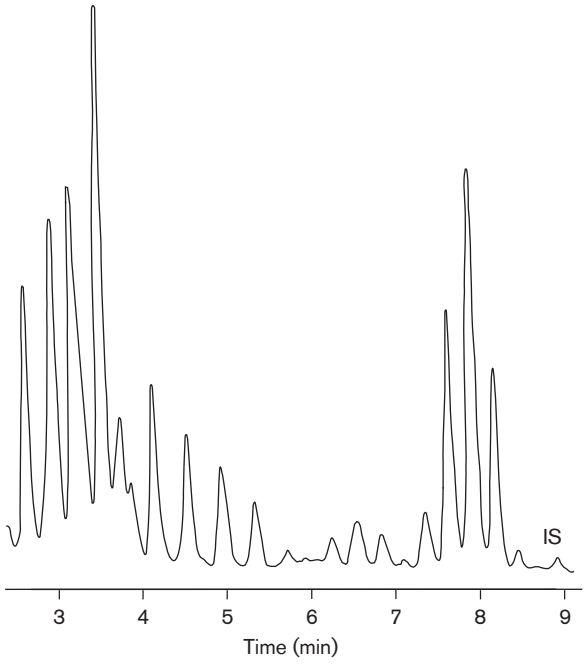


Fig. 1. Representative mycolic acid pattern of strain AFPC-000207^T. IS, internal standard.

pattern is present in the HPLC database at http://www.mycobactoscana.it/page4.htm.

Antibiotic MICs were determined, according to CLSI recommendations (CLSI, 2011), using a commercially available microdilution method (RAPMYCOI; Sensititer) including drugs with potential activity against rapidly growing non-tuberculous mycobacteria. The strains presented a quite uniform pattern characterized by susceptibility to amikacin, clarithromycin, linezolid, quinolones and doxycycline and resistance to cefoxitin. The only exception was represented by trimethoprim/sulfamethoxazole, to which the two Iranian strains were highly resistant and the others extremely susceptible (Table 3).

Double-strand DNA sequences from the four strains included in the study were determined using BigDye Terminator chemistry on an AB3730 DNA sequencer (Applied Biosystems) following the standard protocol of the supplier. The regions investigated included the genes encoding 16S rRNA (Kirschner *et al.*, 1993), 65 kDa heat-shock protein (*hsp65*) (McNabb *et al.*, 2004), molecular chaperone DnaK (*dnaK*) (Dai *et al.*, 2011), DNA gyrase β-subunit (*gyrB*) (Gomila *et al.*, 2007), superoxide dismutase (*sodA*) (Zolg & Philippi-Schulz, 1994), RNA polymerase β-subunit (*rpoB*) (Adékambi *et al.*, 2003) and RNA polymerase β' subunit (*rpoBC*). The two Iranian strains had identical sequences in all these regions, the Turkish strains differed slightly from each other in two regions, while the Italian strains presented minor differences in six of the seven genetic targets (Table S1, available in the online Supplementary Material). All the strains clearly differed from any known species of the genus *Mycobacterium* and displayed, in various genes, closest similarity to different rapidly growing species (Table S2).

The whole genomes of the proposed type strain and of *Mycobacterium moriokaense* ATCC 53059^T, the species most closely related based on 16S rRNA gene sequence similarity, were sequenced. Paired-end libraries of read length 150 bp were prepared using a Nextera XT DNA Sample Preparation kit and Nextera XT Index kit (Illumina) according to the manufacturer's protocol. Libraries were then normalized to 2 nM, pooled for multiplexing in equal volumes and sequenced at 10 pM on an Illumina MiSeq platform with 300 nt paired end reads to achieve a coverage $>100\times$ per base. Read tags were processed with Trimmomatic (Bolger *et al.*, 2014) and then assembled with Velvet software (Zerbino & Birney, 2008). The resulting contigs were used to calculate average nucleotide identity (ANI), using the software available at http://enve-omics.ce.gatech.edu/ani/ (Konstantinidis & Tiedje, 2005). The resulting ANI was lower than 83 % (Fig. S1), clearly below the accepted cut-off of 95–96 % (Kim *et al.*, 2014).

PCR restriction analysis patterns (Telenti *et al.*, 1993) were deduced on the basis of restriction sites present in the

Table 3. MICs of antimycobacterial drugs potentially active on rapidly growing mycobacteria

S, susceptible; R, resistant.

Drug	Strain						Interpretation
	AFPC-000207^T	AFPC-00088	FI-09258	FI-10161	13DK204	E498	
Amikacin	$\leqslant$1	$\leqslant$1	$\leqslant$1	$\leqslant$1	$\leqslant$1	$\leqslant$1	S
Cefoxitin	>128	>128	128	>128	128	>128	R
Ciprofloxacin	1	1	0.5	0.5	1	0.5	S
Clarithromycin	1	1	1	0.5	0.25	0.12	S
Doxycycline	0.25	$\leqslant$0.12	$\leqslant$0.12	$\leqslant$0.12	0.25	0.5	S
Linezolid	$\leqslant$1	$\leqslant$1	$\leqslant$1	$\leqslant$1	$\leqslant$1	$\leqslant$1	S
Moxifloxacin	$\leqslant$0.25	$\leqslant$0.25	$\leqslant$0.25	$\leqslant$0.25	0.5	$\leqslant$0.25	S
Trimethoprim/sulfamethoxazole	>8/152	>8/152	$\leqslant$0.25/4.75	$\leqslant$0.25/4.75	$\leqslant$0.25/4.75	0.5/9.5	S–R

Table 4. PCR restriction analysis patterns of the test strains compared with the most closely related species

Species: 1, *M. celeriflavum* sp. nov.; 2, *M. flavescens*; 3, *M. moriokaense*; 4, *M. agri*; 5, *M. rutilum*; 6, *M. novocastrense*. Only fragments greater than 50 bp were taken into account. The fragment in parentheses is present in only one of the sequevars of the species.

Restriction enzyme	1	2	3	4	5	6
*Bst*EII	440	440	230–210	230–131–79	440	440
*Hae*III	174–87–(59)–58	138–58–51	196–87–58	160–145–59	196–78–51	138–58–51

hsp65 gene sequences. No restriction site for *Bst*EII was detected in any of the strains. The restriction enzyme *Hae*III produced fragments of 174, 87 and 58 bp in all of them, and in strains E498 and FI-10161 a 59 bp fragment was produced as well. A comparison of PCR restriction analysis patterns of the test strains and of closely related species is presented in Table 4.

For phylogenetic analysis, sequences of the type strains of closely related species were retrieved from the GenBank database, aligned using CLUSTAL W software (Thompson *et al.*, 1994) and trimmed to start and finish at the same position. Phylogenetic analysis was conducted for each investigated genetic region and *Mycobacterium tuberculosis* ATCC 27294^T was used as an outgroup. The neighbour-joining method (Saitou & Nei, 1987), supported by the MEGA 6 software (Tamura *et al.*, 2013), was used for the reconstruction of phylogenetic trees; 1000 bootstrap replications were implemented. The phylogenetic tree reconstructed using 16S rRNA gene sequences (Fig. 2) was characterized by very low robustness (almost 80 % of the nodes had bootstrap percentages below 75 %), which was particularly subtle for the branch including the test strains and *Mycobacterium brumae* ATCC 51384^T. The trees inferred from *hsp65*, *gyrB* and *dnaK* gene sequences showed similarly low robustness (Figs S2 and S3), in which the test strains clustered with different species of the genus *Mycobacterium*. Robust trees were obtained using the *rpoB*

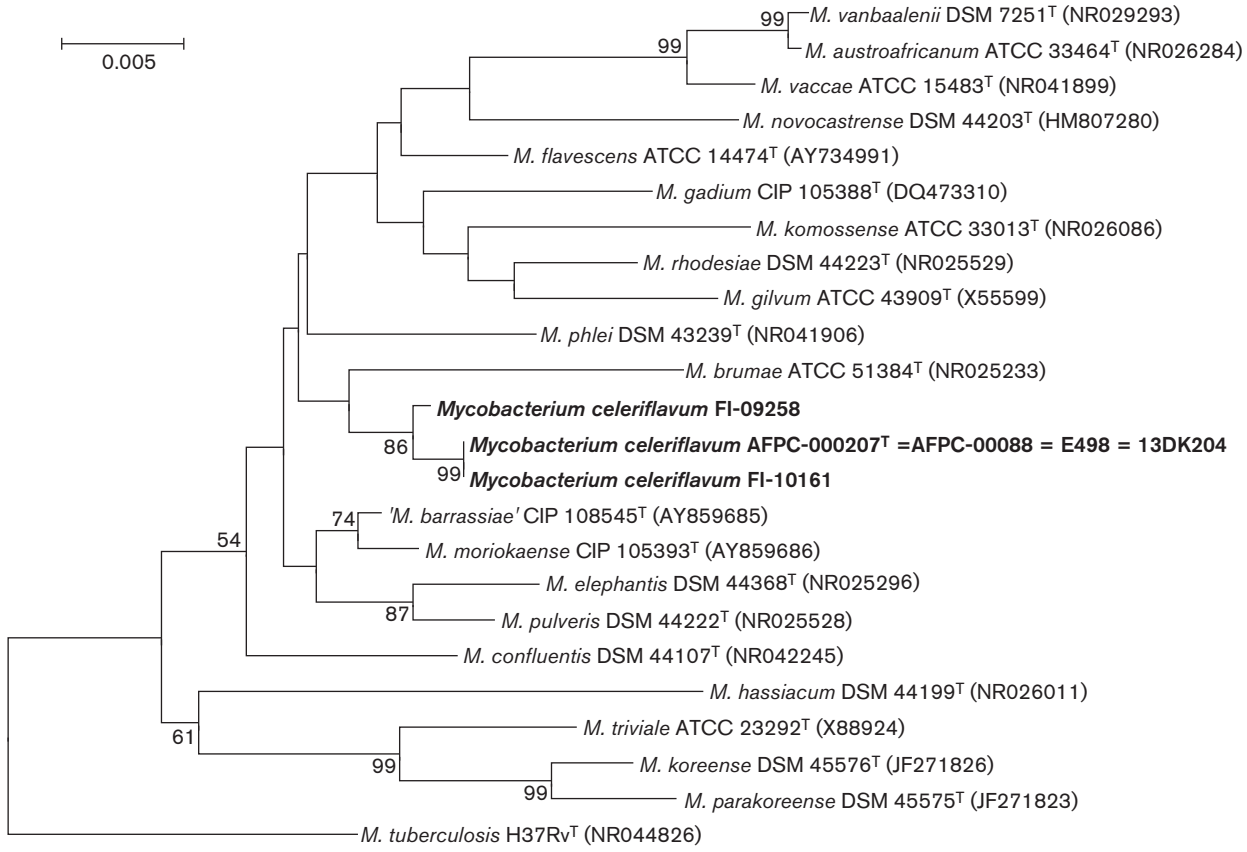


Fig. 2. Phylogenetic tree based on 16S rRNA gene sequences, reconstructed using the neighbour-joining method bootstrapped 1000 times. Bootstrap values >50 % are given at nodes. Bar, 0.005 substitutions per nucleotide position.

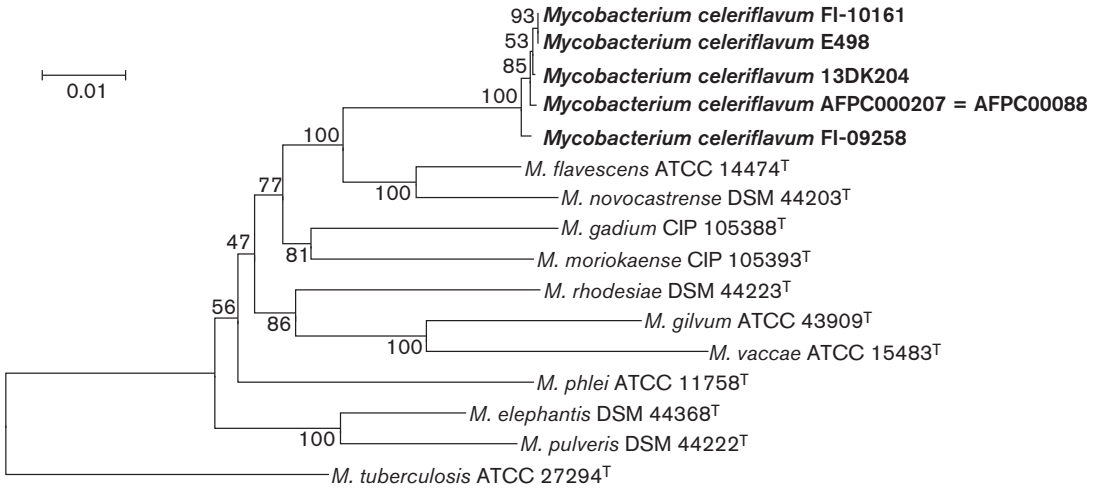


Fig. 3. Phylogenetic tree based on concatenated sequences of 16S rRNA, *hsp65*, *rpoB* and *sodA* genes, constructed using the neighbour-joining method bootstrapped 1000 times. Bootstrap values >50 % are given at nodes. Bar, 0.01 substitutions per nucleotide position.

and *sodA* gene sequences (Figs S2 and S3); in both cases, the closest species were *M. flavescens* and *Mycobacterium novocastrense*. A phylogenetic tree was not reconstructed for the *ropBC* gene because of the very limited number of species of the genus *Mycobacterium* for which such sequences are available. Phylogenetic trees based on concatenated sequences of various housekeeping genes are known to be characterized by high robustness (Devulder *et al.*, 2005; Mignard & Flandrois, 2008; Stackebrandt *et al.*, 2002). Sequences of the most closely related mycobacteria were available in the GenBank database for 16S rRNA, *hsp65*, *rpoB* and *sodA* genes; their concatenation produced strings 2947 bp long, from which a phylogenetic tree was reconstructed (Fig. 3). The bootstrap values were, in this case, very high and confirmed the relatedness with *M. flavescens* and *M. novocastrense*.

Based on the data presented, the six strains of mycobacteria described here are considered to represent a novel species of the genus *Mycobacterium*, for which the name *Mycobacterium celeriflavum* sp. nov. is proposed.

Description of *Mycobacterium celeriflavum* sp. nov.

Mycobacterium celeriflavum (ce.le.ri.fla′vum. L. adj. *celer* rapid; L. neut. adj. *flavum* yellow; N.L. neut. adj. *celeriflavum* referring to rapid growth and yellow pigmentation features of the species).

Cells are Gram-stain-positive, non-motile, non-spore-forming, acid–alcohol-fast bacilli. Yellow colonies develop at 25–37 °C both in the light and in the dark. Biochemical tests are not suitable to differentiate the novel species from other rapidly growing scotochromogenic species. Negative for niacin accumulation, Tween 80 hydrolysis and β-glucosidase, but positive for nitrate reduction, tellurite reduction and thermostable catalase. The HPLC profile of mycolic acids is characterized by an uninterrupted sequence of peaks with the major ones emerging early (elution minutes 2–4) and very late (around minute 8). Susceptible *in vitro* to amikacin, clarithromycin, doxycycline, linezolid and quinolons. Has unique 16S rRNA gene, *hsp65*, *rpoB*, *gyrB*, *dnaK* and *sodA* sequences that are clearly different from any other mycobacterial species, with *M. flavescens* being the most closely related to the majority of strains of the species.

The type strain is AFPC-000207^T (=DSM 46765^T=JCM 18439^T), which was isolated from a human pulmonary specimen in Iran. Strains AFPC-00088, E498, 13DK204, FI-09258 and FI-10161 are additional strains of the species.

References

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

Bolger, A. M., Lohse, M. & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Cavuşoğlu, C. & Tortoli, E. (2006). [Characterization of two new pigmented mycobacteria isolates]. *Mikrobiyol Bul* **40**, 185–194. [in Turkish] Medline

CLSI (2011). *Susceptibility Testing of Mycobacteria, Nocardiae and other Aerobic Actinomycetes*; Approved Standard, 2nd edn, M24-A2. Wayne, PA: CLSI.

Dai, J., Chen, Y., Dean, S., Morris, J. G., Salfinger, M. & Johnson, J. A. (2011). Multiple-genome comparison reveals new loci for *Mycobacterium* species identification. *J Clin Microbiol* **49**, 144–153.

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.

Gomila, M., Ramirez, A. & Lalucat, J. (2007). Diversity of environmental *Mycobacterium* isolates from hemodialysis water as shown by a multigene sequencing approach. *Appl Environ Microbiol* **73**, 3787–3797.

Kent, P. T. & Kubica, G. P. (1985). *Public Health Mycobacteriology. A Guide for the Level III Laboratory*. Atlanta: US Department of Health and Human Services.

Kim, M., Oh, H. S., Park, S. C. & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* **64**, 346–351.

Kirschner, P., Springer, B., Vogel, U., Meier, A., Wrede, A., Kiekenbeck, M., Bange, F. C. & Böttger, 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.

Konstantinidis, K. T. & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* **102**, 2567–2572.

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.

Mignard, S. & Flandrois, J. P. (2008). A seven-gene, multilocus, genus-wide approach to the phylogeny of mycobacteria using supertrees. *Int J Syst Evol Microbiol* **58**, 1432–1441.

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

Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A., Kämpfer, P., Maiden, M. C., Nesme, X., Rosselló-Mora, R., Swings, J. & other authors (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* **52**, 1043–1047.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.

Telenti, A., Marchesi, F., Balz, M., Bally, F., Böttger, 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., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

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

Tortoli, E. (2009). Clinical manifestations of nontuberculous mycobacteria infections. *Clin Microbiol Infect* **15**, 906–910.

Zerbino, D. R. & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821–829.

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

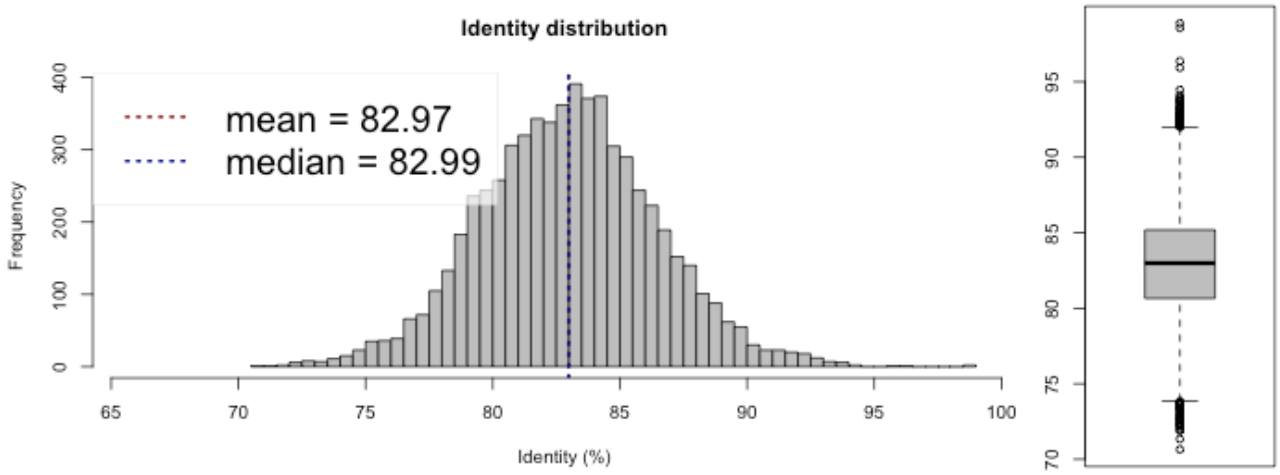
Supplementary table 1. Number of nucleotide substitutions (% similarity) of other test strains in comparison with the type strain (AFPC-000207^{T}) in different genomic regions.

Gene	AFPC-00088	FI-09258	FI-10161	E498	13DK204
16S rRNA	0/1441	5/1441* (99.6)	1/1441 (99.9)	0/1441	0/1441
dnaK	0/450	4/450 (99.1)	0/450	0/450	0/450
gyrB	0/281	3/281 (98.9)	3/281 (98.9)	3/281 (98.9)	3/281 (98.9)
hsp65	0/401	1/401 (99.7)	2/401 (99.5)	2/401 (99.5)	1/401 (99.7)
rpoB	0/710	3/710 (99.6)	1/710 (99.9)	1/710 (99.9)	1/710 (99.9)
rpoBC	0/478	6/478 (98.7)	4/478 (99.2)	3/478 (99.4)	4/478 (99.2)
sodA	0/389	2/389 (99.5)	1/389 (99.7)	2/389 (99.5)	0/389

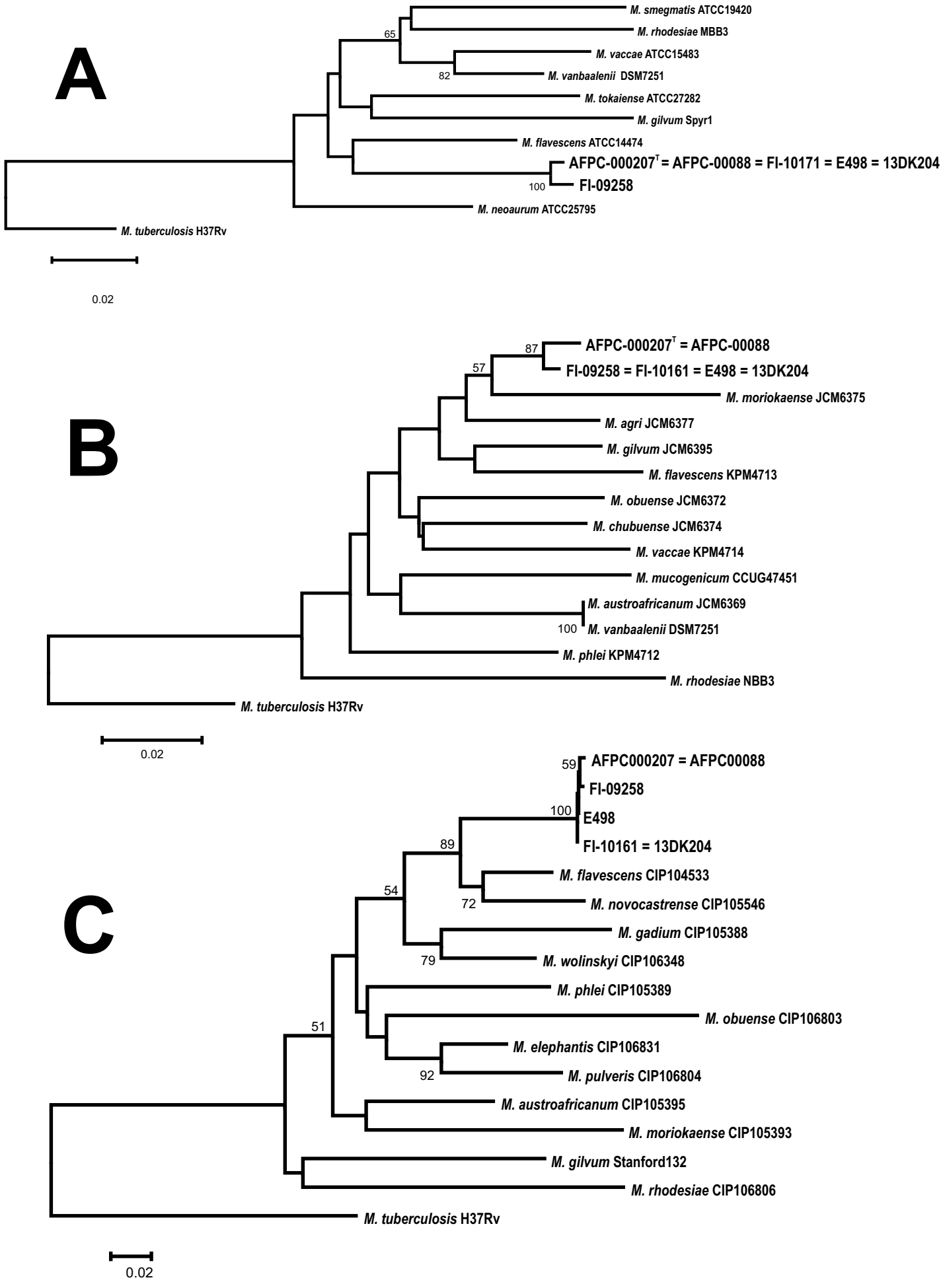
***** 4 out of 5 nucleotide substitutions concerned a 16S rRNA variable region corresponding to *Escherichia coli* positions 1133-1140

Supplementary table 2. Closest sequence similarities of test strains with other *Mycobacterium* species in different genomic regions.

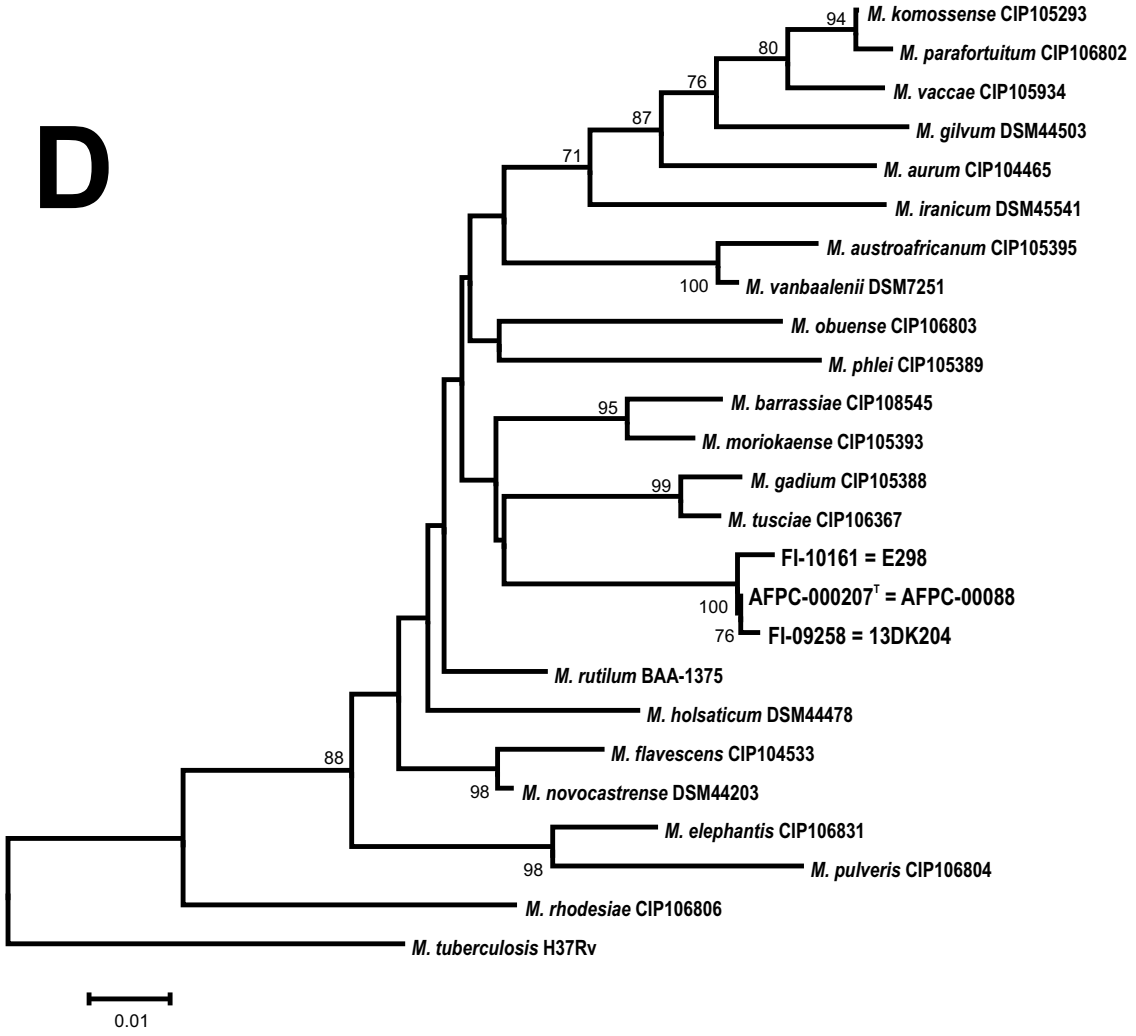
Gene	Closest species	Sequence similarity
16S rRNA	*M. moriokaense*	98.7-99.0%
dnaK	*M. flavescens*	91.5%
gyrB	*M. agri*	94.7-95.0%
hsp65	*M. rutilum*	95.2-96.5%
rpoB	*M. novocastrense*	93.7-93-8%
rpoBC	*M. flavescens*	92.3-93.1%
sodA	*M. flavescens*	92.0-92.3%

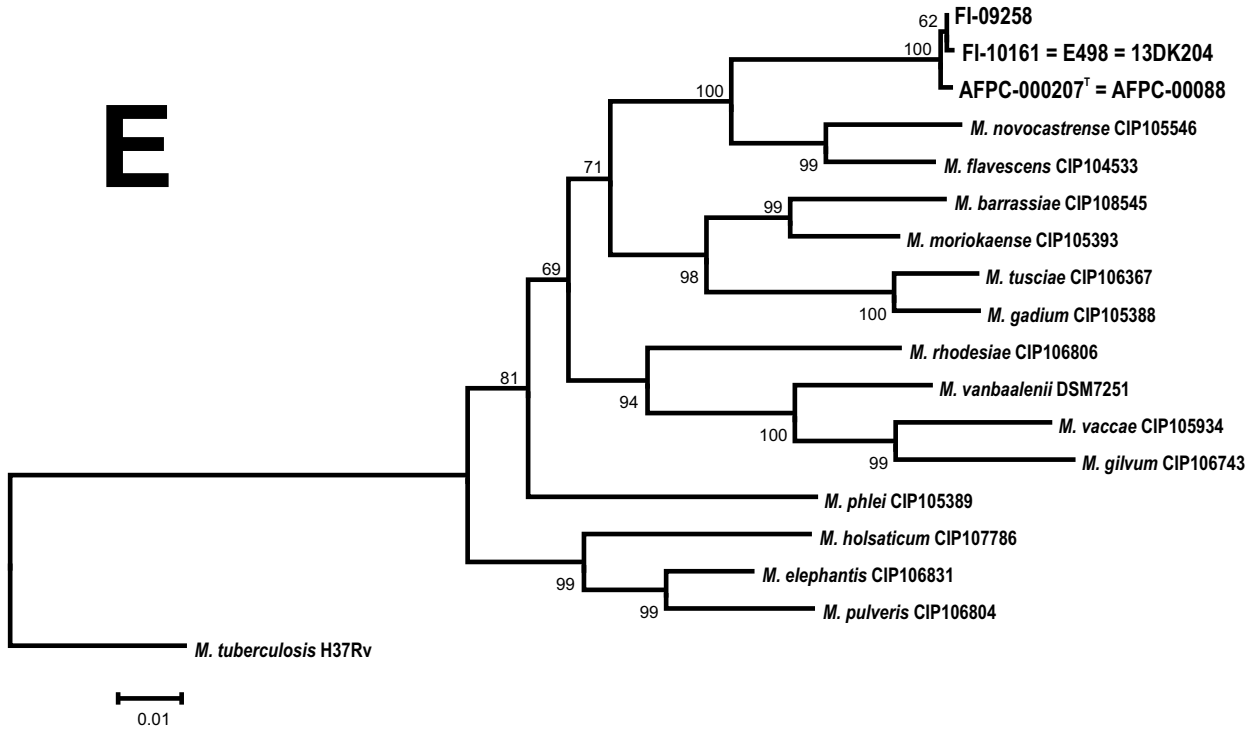


Supplementary figure 1. Average nucleotide identity between AFPC-000207^{T} and *Mycobacterium moriokaense*^{T}. The distribution of ANI is calculated using draft genomes obtained by assembly of short read tags.



Supplementary figure S2. Phylogenetic trees based on: A, DnaK; B, gyrB and C, sodA sequences, constructed using the neighbor-joining method bootstrapped 1000 times. Bootstrap values >50 are given at nodes.
Bars, 0.02 substitutions per nucleotide position.





Supplementary figure S3. Phylogenetic trees based on: D, *hsp65* and E, *rpoB* sequences, constructed using the neighbor-joining method bootstrapped 1000 times. Bootstrap values >50 are given at nodes.
Bars, 0.01 substitutions per nucleotide position.