ELSEVIER

Molecular and Cellular Probes 18 (2004) 399–408

MOLECULAR AND CELLULAR PROBES

www.elsevier.com/locate/ymcpr

Taxonomic and phylogenetic status of non-tuberculous mycobacteria in a Caribbean setting

Séverine Ferdinand[a], Eric Legrand[a], Khye Seng Goh[a], Mylène Berchel[a], Gianna Mazzarelli[b], Christophe Sola[a], Enrico Tortoli[b], Nalin Rastogi[a,*]

[a]*Unité de la Tuberculose et des Mycobactéries, Institut Pasteur de Guadeloupe, Morne Jolivière, BP 484, F-97165 Pointe-à-Pitre Cedex, Guadeloupe France*
[b]*Microbiology Laboratory, Regional Mycobacteria Reference Center, Careggi Hospital, Florence, Italy*

Received 30 April 2004; accepted for publication 21 June 2004

Abstract

This report describes detailed taxonomic and phylogenetic analysis of 15 non-tuberculous mycobacteria (NTMs) isolated from human pathological specimens in a Caribbean setting (12 slow-growers and three rapid-growers) that were not identified by cultural and biochemical tests and drug-susceptibility results. These isolates were further studied using PCR restriction fragment length polymorphism analysis (PRA) of a 441 bp *hsp65* fragment, as well as the sequencing of 16S rDNA and *hsp65* DNA, and HPLC of the mycolic acids. Our results showed that taxonomic position of well-defined NTMs was resolved by PRA and sequencing of *hsp65*, nonetheless, it was not suitable to investigate rarely observed or new strains that required 16S rDNA sequencing and HPLC for a definite response. Unrooted neighbor-joining phylogenetic trees were drawn based upon the 16S rDNA and *hsp65* sequences of the 15 NTMs compared with those from described species (73 for 16S rDNA and 45 for *hsp65*). For most of the NTMs not showing an exactly matching sequence with either *hsp65* or 16S rDNA in the GenBank, the phylogenetic tree was able to provide with useful indications about their relatedness to known species. In such a case, a concording HPLC pattern with the sequence data and the place of the strain within the tree could lead to a potential identification. We also identified three identical isolates that define a new mycobacterial species within the group of *M. simiae*-related mycobacteria. The isolation and characterization of mycobacteria from new settings may lead to identify potential pathogens that may propogate in future because of increased human migration, travels, and climatic and ecological changes of the modern world.

Keywords: Taxonomy; Phylogeny; Non-tuberculous mycobacteria; PCR-restriction analysis; DNA sequencing; HPLC

1. Introduction

The genus *Mycobacterium* encompasses more than 90 species that include pathogenic or potentially pathogenic species both for humans, and animals [1,2]. Almost all of the non-tuberculous mycobacteria (NTMs) are frequently isolated from diverse environmental sources such as water, soil, aerosols [3], and many studies have shown epidemiological links between strains isolated in tap water, and those isolated from patients [4]. Unlike for tubercle bacilli, human-to-human transmission of NTMs has never been established; though potential animal sources of infection have been suggested [5] both in HIV immunocompromised, as well as in apparently healthy individuals [6]. Previous studies have shown that identification of mycobacteria by phenotypical tests alone may result in erroneous identification; not only as the tests used may not be fully reproducible but also as phenotypic features of rare and new species are poorly known and often overlap the ones of best known mycobacteria. The HPLC of mycolic acids in conjunction with new molecular methods such as PCR restriction fragment length polymorphism analysis (PRA) of *hsp65*, 16S rRNA and *hsp65* genes sequencing, may help to define taxonomical status of difficult to identify NTMs. In this study, we report our results on 15 NTMs isolated, between 1993 and 1998 from Caribbean patients, that could not be precisely identified by standard biochemical tests.

* Corresponding author. Tel.: +590-590-893881; fax: +590-590-893880.
E-mail address: nrastogi@pasteur-guadeloupe.fr (N. Rastogi).

0890-8508/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcp.2004.06.006

2. Materials and methods

2.1. Mycobacterial isolates and phenotypic identification and DNA probe assay

Clinical isolates investigated in this study were either growth from pathological specimens at Institut Pasteur de Guadeloupe, or sent, for identification purposes, from regional microbiology laboratories in French Guyana and Martinique. The origin of specimens, demographic and clinical characteristics of the patients, biochemical tests performed and drug-susceptibility testing results (using the 1% proportion method) are summarized in Table 1. The strains were also tested with the AccuProbe rRNA hybridization assay (Gen-Probe Inc., San Diego, USA), using probes for identification of *M. kansasii*, *M. gordonae*, *M. avium* complex (MAC), *M. avium*, and *M. intracellulare*, respectively. If necessary, the INNO LiPA MYCOBACTERIA (Innogenetics, Ghent, Belgium) identification kit was also used [7]. Analysis of mycolic acids was initially performed by thin-layer chromatography (TLC) [8] followed by HPLC [9] if necessary.

2.2. DNA extraction

DNA was prepared by a glass bead method as previously reported [10]. Briefly, one loopful of bacteria was suspended in 300 µl of Tris–EDTA (10 mM Tris, 1 mM EDTA, pH 7.4), and 100 µl of acid-washed glass beads were added to the tubes (diameter 106 µm; Sigma, St Louis, USA), heated at 94 °C for 15 min, and sonicated at 35 kHz for 15 min in a water bath sonicator (Gen-Probe Inc., San Diego, USA). An aliquot (5 µl) of the supernatant containing the crude extract was used for PCR.

2.3. PCR restriction fragment length polymorphism analysis (PRA)

PRA based on the amplification of a 441-bp fragment of the *hsp65* gene (positions 396–836 of the published sequence from *M. tuberculosis*) was performed with forward primer Tb11 (5′-ACCAACGATGGTGTGTCCAT-3′), and reverse primer Tb12 (5′-CTTGTCGAACCGCATACCCT-3′) described by the method of Telenti et al. [11]. Twenty-two microliters of the PCR product was digested by *Bst*EII (Promega, Madison, WI, USA), and *Hae*III (Biolabs Inc., Beverly, MA, USA), and all the restriction digestion was loaded on a 3% (wt/vol.) Metaphor agarose gel (FMC Bioproducts, Rockland, USA). A 100-bp ladder (Pharmacia-Biotech, Uppsala, Sweden) served as an external molecular size marker, and was added after every sixth lane of migration to reduce migration-related errors. The fragments were visualized by ethidium bromide staining, images were videocaptured, and analyzed using the Gel-Analyst software (Bioprobe Systems, Montreuil, France). The Taxotron package software (P.A.D. Grimont, Institut Pasteur, Paris, France) was used to convert the migration file into a molecular weight data file as reported by Devallois et al. [10].

2.4. Automated DNA sequencing of 16S rDNA

PCR was performed in a 50 µl reaction mixture containing 50 mM KCl, 10 mM Tris–HCl (pH 9), 1.5 mM $MgCl_2$, 500 µM (each) deoxynucleoside triphosphates, 0.05 U/µl of rTaq polymerase (Amersham Pharmacia-Biotech Inc., Piscataway, USA), 0.6 pmol/µl of reverse primer 264 (5′-TGCACACAGGCCACAAGGGA-3′, corresponding to *Escherichia coli* 16S rDNA from position 1046 to 1027), and 0.6 pmol/µl of biotinylated forward primer 285 (5′-GAGAGTTTGATCCTGGCTCAG-3′, corresponding to *E. coli* 16S rDNA from position 9 to 29). A 7 min denaturation step at 94 °C was followed by 40 cycles of 94 °C for 30 s, 68 °C for 2 min, and 72 °C for 10 min in a 9600 Perkin–Elmer thermocycler (Perkin–Elmer Corporation, Norwalk, USA). PCR products were controlled on a 1% agarose gel, and purified with QIAquick columns (Qiagen, Courtaboeuf, France) in order to remove unincorporated nucleotides, and primers from the amplified DNA. Sequencing reactions were performed using the Thermo Sequenase fluorescent-labeled primer cycle sequencing kit (Amersham-Pharmacia Biotech, Buckinghamshire, England) following the manufacturer's specifications. A 1000 bp segment was amplified in a DNA thermal cycle using the following fluorescent primers: forward Cy5.5-284 (5′-XCCTGGCTCAGGACGAACGCT-3′), and reverse Cy5.0-244 (5′-XCCCACTGCTGCCTCCCGTAG-3′) (Oligo Express, Montreuil, France). The labeled extension products were denatured at 94 °C for 5 min, and 1.5 µl of each sample was loaded onto a 6% acrylamide gel for analysis. Electrophoretic separation was performed with an automated Opengene® Long Read Tower DNA sequencer (Visible Genetics Inc., Toronto, Canada) operated at 1500 V for 240 min. All the sequencing data were from two independent experiments, and any discrepancy was systematically checked by performing a third experiment.

2.5. Automated DNA sequencing of hsp65

The first PCR was performed in the same reaction mixture conditions as the first PCR for 16S rDNA sequencing. The forward M13-Tb11 (5′-TGTAAAACGACGGCCAGTACCA-ACGATGGTGTGTCCAT-3′), and reverse M13-Tb12 (5′-CAGG-AAACAGCTATGACCCTTGTCGAACCGCATACCCT-3′) primers were used at 0.8 pmol/µl according to Ringuet et al. [12]. A 7 min denaturation step at 94 °C was followed by 40 cycles of 94 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min, and 72 °C for 10 min. In a second PCR, a 441-bp segment (nucleotides 396–836) of *hsp65* was amplified in a DNA thermal cycle (Perkin–Elmer 9600) using the following fluorescent

Table 1
A summary of clinical and epidemiological data, cultural and biochemical characteristics, and drug-susceptibility results of 15 isolates of non-tuberculous mycobacteria

CLINICAL AND EPIDEMIOLOGICAL DATA								CULTURAL AND BIOCHEMICAL TESTS										DRUG SUSCEPTIBILITY						
Strain	Isolated from	Month and year of isolation	Age	Sex	Origin	HIV status	Smear	Growth (°C)	Growth (days) 7	Growth (days) 21	Pigment formation	Niacin	Aryl-sulfatase	Catalase 22°C	Catalase 68°C	Nitrate Reduction	Urease	Tween 80 Hydrolysis (days)	Mycolic acid type	INH	SM	RIF	EMB	TB1	NAP	Accuprobe
94008	Sputum	Jan 1994	NA	M	GPE	+	ND	30, 37	-	+	P	-	-	+	+	+	+	+ (5d)	1, 3, 4	R	R	S	S	S	S	*M. kansasii* +
95001	Gastric washing	Dec 1993	NA	M	GPE	+	ND	30, 37	-	+	Sc	-	-	+	+	-	+	+ (5d)	1, 3, 4	R	R	R	S	R	R	*M. gordonae* +
95077	Sputum	Oct 1995	64	M	GPE	NA	+	30, 37	-	+	Sc	-	+	+	+	-	-	+ (5d)	1, 3, 4	R	R	R	S	R	R	*M. gordonae* +
96048	Gastric washing	May 1996	90	M	MTQ	NA	-	30, 37	-	+	N	-	+	+	+	+	-	+ (5d)	NON	R	S	S	S	R	R	NEG
96071	Gastric washing	July 1996	41	M	GUF	NA	+	30, 37	-	+	Sc	-	-	+	+	-	+	-	1, 4, 6	R	R	R	S	R	R	NEG
97047	Urine	June 1997	43	F	GUF	NA	+	30, 37, 42	-	+	N	-	-	+	+	-	+	-	1, 3, 4	R	R	S	R	R	R	NEG
97050	Gastric washing	May 1997	57	M	GUF	+	+	30, 37, 42	-	+	N[a]	-	-	+	+	-	-	Variable	1, 6	R	S	S	R	R	R	NEG
97056	Bronchial aspirate	July 1997	60	M	GUF	+	+	30, 37	-	+	N	-	-	+	+	-	-	-	1, 6	R	R	R	R	R	R	NEG
98002	Gastric washing	Jan 1998	74	M	GUF	NA	-	30, 37	-	+	N	-	+	+	+	+	-	+ (1d)	1, 6	R	S	R	S	S	R	NEG
98075	Gastric washing	Sept 1998	11	M	GUF	NA	+	30, 37	-	+	N	-	-	+	+	-	+	-	1, int 4/5[b]	R	R	S	S	R	R	NEG
98082	Gastric washing	Oct 1998	72	F	GUF	NA	+	30, 37	-	+	N	-	-	+	+	-	+	-	1, int 4/5[b]	R	R	S	S	R	R	NEG
98095	Bronchial aspirate	Oct 1998	54	M	GUF	NA	+	30, 37	-	+	N	-	-	+	+	-	+	-	1, int 4/5[b]	R	R	S	S	R	R	NEG
96138	Urine	Dec 1996	55	M	GUF	NA	+	30, 37	+	+	N	-	+	+	+	+	+	-	1, 5	R	R	R	S	R	R	NEG
97078[c]	Milk (Cow)	Sept 1997	NA	NA	MTQ	NAP	-	30, 37, 42	+	+	Sc	-	-	+	+	+	+	-	NON	R	S	R	S	R	R	NEG
98007	Gastric washing	Jan 1998	64	M	GUF	-	+	30, 37, 42	+	+	N	-	+	+	+	+	+	-	1, 5	R	R	R	R	R	R	NEG

Abbreviations: +, positive result; -, negative result; ND, non done; NA, not available; NON, non-interpretable; NEG, negative results for all Accuprobes available; NAP, non applicable; M, male; F, female; N, non pigmented; P, photochromogenic; Sc, scotochromogenic; S, drug susceptible; R, drug resistant; INH, isoniazid (1 μg/ml); SM, streptomycin (2 μg/ml); RIF, rifampin (1 μg/ml); EMB, ethambutol (2 μg/ml); TB1, thiacetazone (10 μg/ml); NAP *p*-nitro-α-acetylamino-β-hydroxypropiophenone (5 μg/ml). Drug susceptibility and growth experiments were performed using 7H11 agar medium except for tests for growth in the presence of TB1 and susceptibility to ethambutol that were performed on Löwenstein-Jensen medium.

[a] The pigmentation for this strain was variable as successive cultures varied from nonpigmented to pale yellow.

[b] intermediate spot position between type 4 and 5.

[c] animal isolate (cow).

primers as described previously: forward Cy5.5-Tb11 (5′-ACCAACGATGGTGTGTCCAT-3′), and reverse Cy5.0-Tb12 (5′-CTTGTCGAACCGCATACCCT-3′). PCR products (10 μl) were controlled on 1% agarose gel, and purified with QIAquick columns (Qiagen, Courtaboeuf, France). The electrophoresis operated at 1300 V for 35 min. The sequences of both strands were determined.

2.6. Analysis of sequence data

The sequences of 16S rDNA, and *hsp65* for various mycobacterial species extracted from data banks were edited, and aligned by the CLUSTAL method using Megalign software (DNAStar Inc., Madison, USA). Phylogenetic analysis was performed using the PHYLIP version 3.57c package as recommended [13]. Distance matrices based on Kimura's two parameter model were produced by using the DNADIST program, and a neighbor-joining tree was constructed by using the NEIGHBOR program. The resulting tree was plotted by using the DRAWTREE program. The stability of the grouping was assessed by bootstrapping with SEQBOOT, and CONSENSE programs. A total of 100 bootstrapped trees were generated, and averaged to give the phylogenetic trees. For analysis of 16S rDNA sequencing data, the results obtained in this investigation were compared to known sequences in the GenBank database, and interpreted using the BlastN algorithm (available on http://ww.ncbi.nlm.nih.gov/BLAST).

3. Results

3.1. Routine identification

The biochemical characteristics of the 15 clinical isolates are summarized in Table 1. Twelve strains were slow-growing mycobacteria, and three strains were rapid-growing mycobacteria. All the strains showed a thermoresistant catalase activity, and were not be able to produce niacin. Among the 12 slow-growers, eight strains were non-pigmented, three were scotochromogenic, and one was photochromogenic. Three non-pigmented strains (98075, 98082, and 98095) shared the same biochemical characteristics. For a detailed analysis, the readers are referred to Table 1. TLC analysis of mycolic acids showed five distinct patterns (Table 1), with a unique profile for three identical non-pigmented strains (98075, 98082, and 98095). It consist of a type 1 spot with an additional one intermediate between spot types 4 and 5 (Table 1). Cultural and biochemical tests and drug-susceptibility results did not help to identify these 15 isolates upon initial isolation from pathological samples. However, upon reculturing, three isolates provided with modified test results with identification compatible with *M. kansasii*, and *M. gordonae*. AccuProbe analysis was able to confirm the identification of these three isolates correctly: 94008 as *M. kansasii*; 95001 and 95077 as *M. gordonae*. The remaining 12 isolates gave negative results to all commercialized AccuProbes (Table 1). Two other strains (96138 and 98007) were tentatively identified as members of the *M. fortuitum* complex based on their biochemical characteristics.

3.2. HPLC analysis

The HPLC analysis was performed for 10 unidentifiable isolates (see above). The three isolates previously showing a unique TLC pattern (98075, 98082, and 98095), shared an unique HPLC pattern represented in Fig. 1 and Table 2. This profile differed from all the 99 distinct HPLC profiles from known mycobacterial species (available online from http://www.ao-careggi.toscana.it/microbiologia/CRRM/hplc_engl.htm). Two other strains (97050 and 97056) had a pattern compatible with *M. interjectum* [14]. A third strain (97047) was also compatible with *M. interjectum*, although the HPLC profile was different from the one observed for strains 97050 and 97056. It was not surprising as two different HPLC patterns have been reported for *M. interjectum* [15]. The chromatograms of remaining four strains were typical of known species (patterns available online from http://www.ao-careggi.toscana.it/microbiologia/CRRM/hplc_engl.htm): 96048 was typical of *M. triviale*, 96071 either of *M. avium* complex (MAC) or *M. scrofulaceum*, 98002 of *M. nonchromogenicum*, and 97078 of *M. duvalii*.

3.3. PRA analysis

The *Bst*EII and *Hae*III digestion patterns of *hsp65* PRA are summarized in Table 3. These were compared to

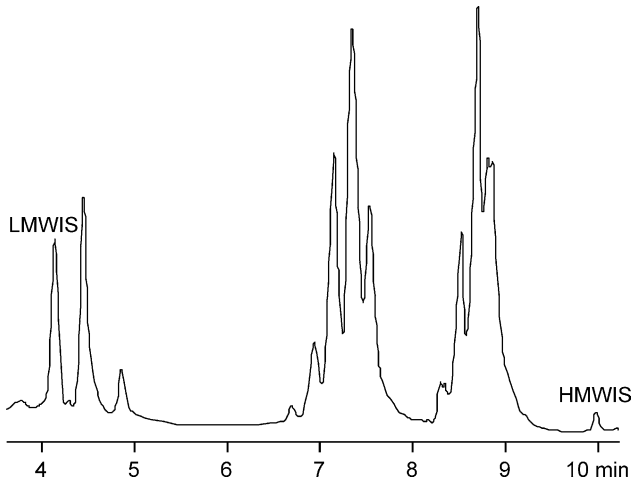


Fig. 1. Representative HPLC pattern of the strains 98075, 98082, and 98095 (only a single pattern is shown here as all the three resulted in an identical pattern). This profile differed from all the 99 distinct HPLC profiles reported so far from known mycobacterial species (available online from http://www.ao-careggi.toscana.it/microbiologia/CRRM/hplc_engl.htm). LMWIS, low molecular weight internal standard; HMWIS, high molecular weight internal standard.

Table 2
Features of the HPLC patterns of mycobacterial strains that were difficult to identify by routine tests and 16S rDNA sequencing

Strains	Profile description	Best resemblance to	References
96048	Single, late, poorly defined cluster	*M. triviale*	[9,14]
96071	Major, early cluster + two minor, close, late clusters[a]	*M. scrofulaceum*	[9,14]
97047	Two late, close clusters	*M. interjectum* type II	[15]
97050, 97056	Major, early cluster + two minor, close, late clusters[a]	*M. interjectum* type I	[15]
97078	Major, early cluster closely followed by two minor peaks	*M. duvalii*	[27]
98002	Single,early cluster	*M. nonchromogenicum*[b]	[9,14]
98075, 98082, 98095	Two very early peaks followed by two late clusters	None	Unreported (this study)

The results obtained were compared with published literature as well as with a collection of 99 HPLC profiles reported so far from 99 mycobacterial species, available online from http://www.ao-careggi.toscana.it/microbiologia/CRRM/hplc_engl.htm.

[a] The profile of *M. scrofulaceum* differs from the one of *M. interjectum* type I mainly for presenting a wider gap between the first and the second cluster of peaks.

[b] The HPLC pattern of *M. engbaekii* has not been reported so far, but may be consulted online from http://www.ao-careggi.toscana.it/microbiologia/CRRM/hplc_engl.htm).

Table 3
Identification and percentage of similarity of 15 unidentifiable NTM isolates according to the GenBank or RIDOM databases for 16S rDNA and *hsp65* sequences

Isolate	Growth	PRA fragments (bp)		PRA identification[a]	*hsp65*[b] sequencing (441 bp)	16S rDNA sequencing (960 bp)
		*Bst*EII	*Hae*III			
94008	slow-grower	240/210	126/103/77	*M. kansasii* I	*M. gordonae* (98.7%)[c]	*M. kansasii* (99.3%)
95001	slow-grower	240/120/100	130/110/(40/30)	*M. gordonae* III	*M. gordonae* (100%)	*M. gordonae* (98.8%)
95077	slow-grower	240/120/100	130/110/95/(40/30)	Unknown pattern[d]	*M. gordonae* (98.6%)	IWGMT 90161 (98.3%)
96048	slow-grower	440	170/140/80/(40)	*M. triviale*[e]	*M. terrae* (90.3%)	*M. triviale* (98.6%)
96071	slow-grower	240/210	145/130	*M. avium* III[f]	*M. avium* (96.8%)	IWGMT 90236 (99.8%)
97047	slow-grower	[240/210][g]	[185/130/(30)][g]	*M. simiae* I[h]	*M. simiae* (99.4%)	*M. interjectum* (99.8%)
97050	slow-grower	240/210	130/115	*M. gordonae* V	*M. avium* (97.8%)	IWGMT 90203 (100%)
97056	slow-grower	440	130/105	Unknown pattern	*M. genavense* (96.7%)	IWGMT 90093 (100%)
98002	slow-grower	325/120	140/85/60/(40)	Unknown pattern[i]	*M. nonchromogenicum* (96.5%)	*M. hiberniae* (99.9%), *M. engbaekii* (100%)[j]
98075	slow-grower	240/210	160/110/55	Unknown pattern	*M. avium* (95.4%)	IWGMT 90100 (98.9%)
98082	slow-grower	240/210	160/110/55	Unknown pattern	*M. avium* (95.8%)	IWGMT 90100 (98.9%)
98095	slow-grower	236/212	160/110/55	Unknown pattern	*M. avium* (95.9%)	IWGMT 90100 (98.9%)
96138	rapid-grower	240/210	145/140/100/(40)	*M. peregrinum* I	*M. peregrinum* (100%)	*M. fortuitum* (99.9%)
97078	rapid-grower	440	134 or [145/140/(40)][g]	*M. flavescens* or new pattern[k]	*M. vaccae* (95.1%)	*M. duvalii* (99.9%)
98007	rapid-grower	240/135/85	145/125/60/(40)	Unknown pattern[l]	*M. fortuitum* (98.9%)	*M. peregrinum* (98.4%)

[a] Based on published data and our own results.

[b] Numbers indicate percentages of similarity between test strains and related species according to the GenBank database.

[c] Values obtained by sequencing on a 1st isolate gave a hypothetical pattern of 240/210 upon *Bst*EII digestion and 130/115/100/40 upon *Hae*III, which is close to *M. gordonae* V *hsp65*. This 1st isolate appeared fortuitously close to *M. gordonae* in the *hsp65* tree. However, another strain from the same patient (not sequenced) gave the PRA values of 240/211 and 126/103/77 compatible with *M. kansasii* I.

[d] An unknown pattern close to *M. gordonae* III (240/120/100 upon *Bst*EII digestion and 130/110 upon *Hae*III digestion); this strain is close to *M. gordonae* in the *hsp65* tree. It probably represents a new *M. gordonae* PRA pattern (pattern X), not previously sequenced for *hsp65*.

[e] Close to *M. triviale* I (440 upon *Bst*EII digestion and 170/140 upon *Hae*III digestion; previously unsequenced, values below 50 bp were not provided), it may represent a new PRA pattern *M. triviale* II.

[f] This strain which gives a *M. avium* III profile, is close to strains IWGMT 90160 and IWGMT 90236 that are within the same group as *M. avium–M. intracellulare–M. scrofulaceum*. Negative to AccuProbe tests using MAC, *M. avium*, and *M. intracellulare* probes, this isolate represents an atypical *M. avium* complex organism.

[g] When in brackets, the *hsp65* PCR product was not well-digested for visual PRA interpretation, hence the values shown are based on restriction sites following *hsp65* sequencing.

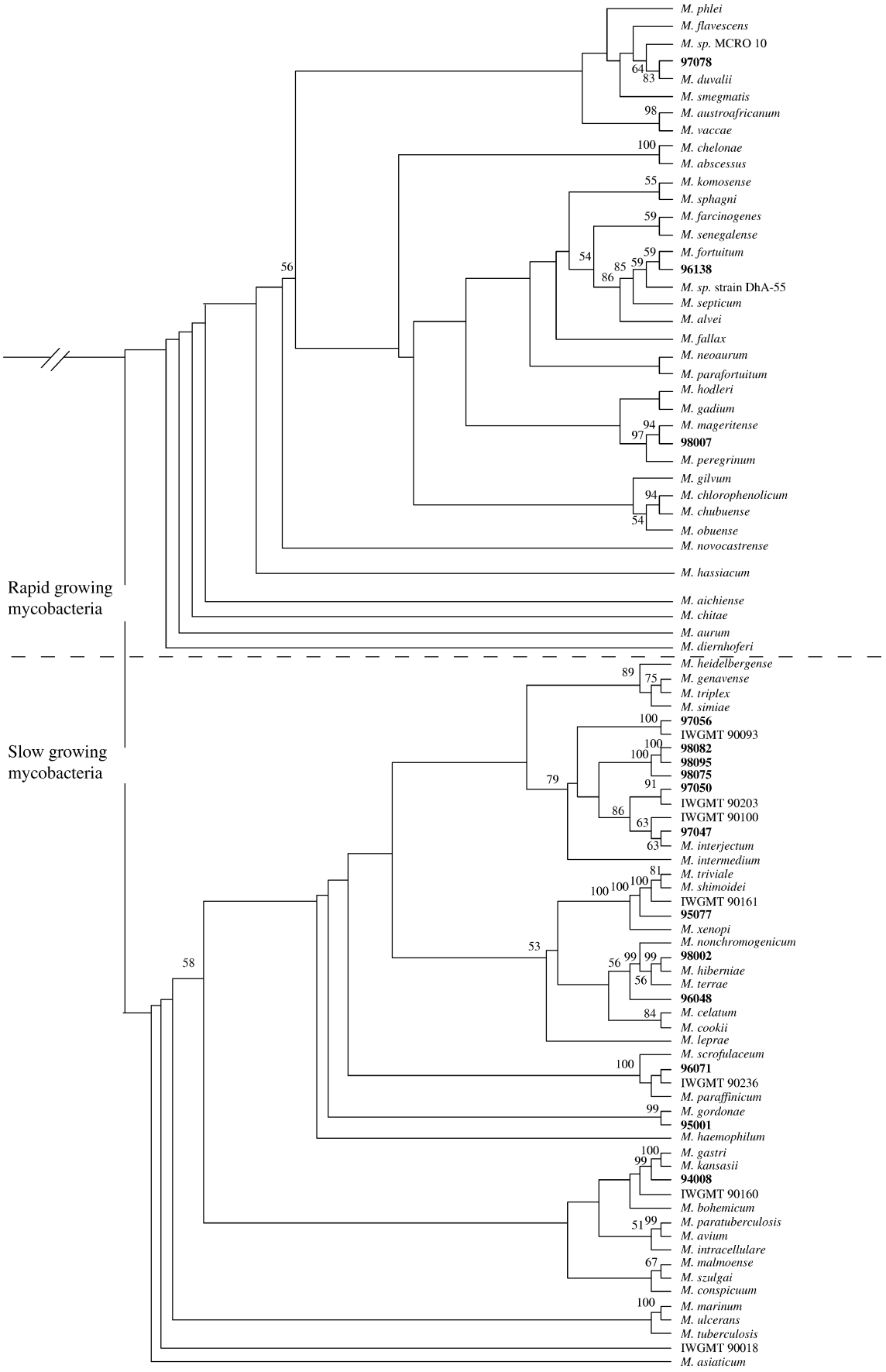
[h] This strain with a *M. simae* I profile, is also close to *M. simiae* by its *hsp65* sequence. However, its 16S rDNA sequence classifies it as the most closely related species *M. interjectum*, as well as other biochemical properties including its susceptibility to rifampin. It may therefore represent a previously unreported *M. interjectum* II PRA profile that is also common to *M. simiae* I.

[i] Pattern close to *M. nonchromogenicum* II (325/120 upon *Bst*EII digestion and 140/60/50/40 upon *Hae*III digestion).

[j] This strain also showed a similarity score of 100% with *M. engbaekii*, which is not yet an officially recognized species [17].

[k] Pattern close to *M. flavescens* I (440 upon *Bst*EII digestion and 145/50 upon *Hae*III digestion), with *hsp65* sequence close to *M. vaccae* on the *hsp65* tree. Based on16S rDNA sequencing, this isolate is identified as *M. duvalii*, a species within the group *M. duvalii–M. flavescens–M. vaccae*. It may represent a previously unreported PRA pattern for *M. duvalii*. A duplicate result gave values of 440 upon *Bst*EII, and 134 upon *Hae*III digestion, representative of *M. flavescens*.

[l] Pattern close to *M. fortuitum* II (235/120/80 upon *Bst*EII digestion and 145/125/60/50 upon *Hae*III digestion). Based on16S rDNA sequencing, this isolate is within a closely related group of *M. fortuitum–M. mageritense–M. peregrinum*.


M. phlei
M. flavescens
M. sp. MCRO 10
97078
M. duvalii
M. smegmatis
M. austroafricanum
M. vaccae
M. chelonae
M. abscessus
M. komosense
M. sphagni
M. farcinogenes
M. senegalense
M. fortuitum
96138
M. sp. strain DhA-55
M. septicum
M. alvei
M. fallax
M. neoaurum
M. parafortuitum
M. hodleri
M. gadium
M. mageritense
98007
M. peregrinum
M. gilvum
M. chlorophenolicum
M. chubuense
M. obuense
M. novocastrense
M. hassiacum
Rapid growing mycobacteria
M. aichiense
M. chitae
M. aurum
M. diernhoferi
M. heidelbergense
M. genavense
M. triplex
M. simiae
97056
IWGMT 90093
98082
98095
98075
97050
IWGMT 90203
IWGMT 90100
97047
M. interjectum
M. intermedium
M. triviale
M. shimoidei
IWGMT 90161
95077
M. xenopi
M. nonchromogenicum
98002
M. hiberniae
M. terrae
96048
M. celatum
M. cookii
M. leprae
M. scrofulaceum
96071
IWGMT 90236
M. paraffinicum
M. gordonae
95001
M. haemophilum
M. gastri
M. kansasii
94008
IWGMT 90160
M. bohemicum
M. paratuberculosis
M. avium
M. intracellulare
M. malmoense
M. szulgai
M. conspicuum
M. marinum
M. ulcerans
M. tuberculosis
IWGMT 90018
M. asiaticum
Slow growing mycobacteria

previously published results [10,16] for species attribution. The patterns for strains 94008, 95001, 96048, 96071, 97050, 96138, and 97078 were those reported for *M. kansasii* I, *M. gordonae* III, *M. triviale*, *M. avium* III, *M. gordonae V*, *M. peregrinum* I, and *M. flavescens*, respectively. Other strains showed previously unreported PRA patterns, including three strains (98075, 98082, and 98095) that showed an identical but unknown pattern with *Bst*EII or *Hae*III enzymes.

3.4. 16S rDNA sequencing and phylogenetic analysis

Approximatively 1000 bp of 16S rDNA were sequenced, and were compared with those available in GenBank (Table 3). Four groups of strains were distinguished. A 1st group of five strains showed a similarity score of $\geq$99% with known species: 94008 with *M. kansasii* (99.3%), 96138 with *M. fortuitum* (99.9%), 97047 with *M. interjectum* X70961 (99.8%, just a one nucleotide difference), 97078 with *M. duvalii* (99.9%), and 98002 with *M. hiberniae* (99.9%), and even 100% with the not officially recognized species *M. engbaekii* [17].

A 2nd group consisted of three strains that yielded a similarity score of $\geq$99% with strains described by the International Working Group on Mycobacterial Taxonomy (IWGMT) [18]; test strains 96071 with IWGMT 90236 (99.8%), and test strains 97050 and 97056 being 100% identical to IWGMT 90203, and IWGMT 90093, respectively. A parallel screening of the test strain 96071 by the INNO LiPA probe identified it as *M. scrofulaceum*.

The 3rd group of four strains (95001, 95077, 96048, and 98007), yielded a similarity score of $<$99% with, respectively, *M. gordonae*, IWGMT 90161, *M. triviale*, and *M. peregrinum*. The original misidentifications of strains 95001 and 95077 were attributed to misinterpretation of phenotypic tests (arylsulfatase, urease). Indeed, this investigation dealt with all strains that did not provide robust identification test results by phenotypic analysis and molecular probes.

Lastly, three strains (98075, 98082, and 98095) were certainly identical as they shared a common sequencing pattern with a similarity score of 98.9% with IWGMT 90100. These three strains probably represent a unique species as they shared common biochemical tests (Table 1), TLC and HPLC profiles (Fig. 1 and Table 2), and PRA patterns (Table 3).

Phylogenetic analysis of the tests strains versus GenBank data showed the presence of two major homogeneous clusters (the slow-growers, and the rapid-growers) (Fig. 2); numbers on the branches indicate bootstrap proportions (100 replicates). Among the 15 test strains, a single cluster of three strains with a bootstrap value of 100% was observed (98075, 98082, and 98095), suggesting that these could belong to a single new species. The phylogenetic tree provides also an indirect taxonomic status of the IWGMT strains previously described by Wayne et al. [18] by clustering them with established taxa. Similarly, five strains (97056, 98082, 98095, 98075, and 97050) clustered with closely related species *M. interjectum* and *M. intermedium*; strain 96048 is grouped with *M. terrae*, *M. hiberniae*, and *M. nonchromogenicum*, and lastly, strain 98007 was included in the same cluster as *M. mageritense*, and *M. peregrinum* (Fig. 2).

3.5. hsp65 sequencing and phylogenetic analysis

A total of 441 bp of the *hsp65* fragment were sequenced, and the data for 401 bp were compared (20 bp per primer were not taken into account) with known *hsp65* sequences in GenBank (Table 3). In general, *hsp65* similarity scores were lower than those obtained by 16S rDNA sequencing. One of the reasons for lower score could be the paucity of *hsp65* data in GenBank, as only 35 sequences were available at the time of this comparison. Moreover, a number of different PRA patterns have been reported for some mycobacterial species, e.g. six different patterns have been reported for *M. kansasii* and nine for *M. gordonae* [10,16]. This heterogeneity of *hsp65* sequences does not permit to have an as accurate comparison for species designation as using 16S rDNA. For example, one of the isolates in our study (94008) turned out to be AccuProbe positive for *M. kansasii* and showed a 100% similarity to the 139 bp sequence of the hypervariable region A in *M. kansasii*, yet was not located close to *M. kansasii* on the *hsp65* tree (Fig. 3). A reconfirmation of PRA on another strain (Table 3)

Fig. 2. Phylogenetic tree based upon the 16S rDNA sequences for 15 non-tuberculous mycobacteria compared with 73, 16S rDNA sequences from described species. The tree was constructed by using the neighbor-joining method. Numbers on the dendrogram are the percentage of occurrence in 100 bootstrapped trees (only values above 50% are shown). GenBank accession numbers of various 16S rDNA sequence described from various mycobacterial species are as follows: *M. abscessus* X82235, *M. aichiense* X55598, *M. alvei* AF023664, *M. asiaticum* M29556, *M. aurum* M29558, *M. austroafricanum* X93182, *M. avium* X52918, *M. paratuberculosis* X52934, *M. bohemicum* U84502, *M. celatum* M08170, *M. chelonae* X82236, *M. chitae* X67874, *M. chlorophenolicum* X81926, *M. chubuense* X55596, *M. conspicuum* X88922, *M. cookii* M59278, *M. diernhoferi* X55593, *M. duvalii* U94745, *M. fallax* M29562, *M. farcinogenes* X55592, *M. flavescens* M29561, *M. fortuitum* X52921, *M. gadium* X55594, *M. gastri* X52919, *M. genavense* X60070, *M. gilvum* X55599, *M. gordonae* X52923, *M. haemophilum* L24800, *M. hassiacum* U49401, *M. heidelbergense* AJ000684, *M. hodleri* X93184, *M. interjectum* X70961, *M. intermedium* X67847, *M. intracellulare* X52927, *M. kansasii* X15916, *M. komosense* X55591, *M. leprae* X58891, *M. margeritense* AJ011335, *M. malmoense* X52930, *M. marinum* X52920, *M. neoaurum* M29564, *M. nonchromogenicum* X52928, *M. novocastrense* U96747, *M. obuense* X55597, *M. paraffinicum* X88925, *M. parafortuitum* X93183, *M. peregrinum* AF058712, *M. phlei* M29566, *M. scrofulaceum* X52924, *M. senegalense* M29567, *M. shimoidei* AJ005005, *M. simiae* X52931,

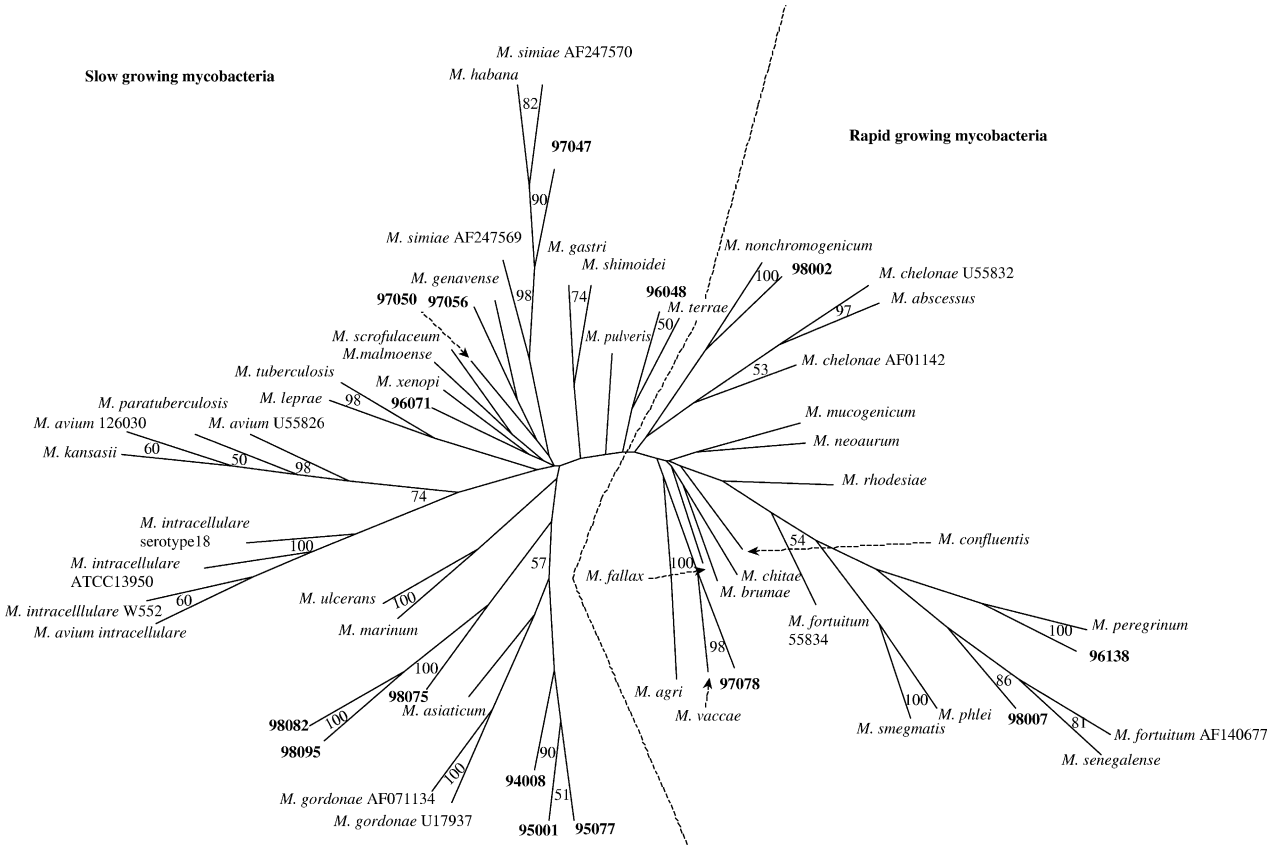


Fig. 3. Unrooted phylogenetic tree based upon the *hsp65* sequences for 15 non-tuberculous mycobacteria compared with 45 *hsp65* sequences from described species. The tree was constructed by using the neighbor-joining method. Numbers on the dendrogram are the percentage of occurrence in 100 bootstrapped trees (only values above 50% are shown). GenBank accession numbers of various *hsp65* sequences described from various mycobacterial species are as follows: *M. abscessus* AF071139, *M. agri* U17920, *M. asiaticum* U17921, *M. avium* ssp. *avium* U55826, *M. avium* ssp. *avium* AF126030, *M. avium* ssp. *paratuberculosis* U55827, *M. brumae* AF071129, *M. chelonae* AF072242, *M. chelonae* U55832, *M. chitae* AF071131, *M. confluentis* AF071132, *M. fallax* U17930, *M. fortuitum* U55834, *M. fortuitum* AF140677, *M. gastri* U17931, *M. genavense* U17932, *M. gordonae* AF071134, *M. gordonae* U17937, *M. habana* AF129011, *M. intracellulare* serotype 18 U55830, *M. intracellulare* U17924, *M. intracellulare* strain ATCC13950, *M. intracellulare* strain w-552 U85638, *M. kansasii* U17947, *M. leprae* M14341, *M. malmoense* U17948, *M. marnium* U55831, *M. mucogenicum* AF071135, *M. neoaurum* U17950, *M. nonchromogenicum* U17951, *M. peregrinum* AF071136, *M. phlei* U17952, *M. pulveris* U17953, *M. rhodesiae* U17954, *M. scrofulaceum* U17955, *M. senegalense* AF071137, *M. shimoidei* U17956, *M. simiae* AF247569, *M. simiae* AF247570, *M. smegmatis* AF071138, *M. tuberculosis* AL021932, *M. ulcerans* U34034, *M. vaccae* U17958, *M. xenopi* U17959, *M. terrae* AF257468.

recultured from the same patient resulted in a pattern typical of *M. kansasii* I (discussed below).

Nonetheless, *hsp65* sequencing did provide results comparable to 16S rDNA sequencing for certain species, e.g. 95001 and 95077 as *M. gordonae*, 96138 and 98007 as *M. fortuitum*, and *M. peregrinum*, respectively (Table 3). Interestingly, the *hsp65* phylogenetic analysis (Fig. 3) also distinguished between slow, and rapid growing mycobacteria, with some exceptions, e.g. two slow-growers (*M. nonchromogenicum* and strain 98002) and one rapid-grower (*M. pulveris*) were not classified within the right groups. Strains 98075, 98082, and 98095 have the same *hsp65* sequence, and are on a separated branch from the other mycobacterial species. Sequence results for strain 96138 seems to be discordant; this strain yields a similarity score of 99.9% with *M. fortuitum* for 16S rDNA, and a similarity score of 100% with *M. peregrinum* for *hsp65* sequence. However, *M. peregrinum* was previously classified as a subspecies of *M. fortuitum* [19], and the two species are difficult to separate (Table 3).

4. Discussion

Our data show that taxonomic position of NTMs from clinical isolates, which remained unresolved by cultural and biochemical characteristics could be partially resolved by PRA and sequencing of *hsp65*, which are frequently used for mycobacterial identification [11,20]. Nonetheless, 16S rDNA sequencing [20], and HPLC of mycolic acids were needed for a definite response in many cases. Because of some unrelatedness between our patterns and corresponding reference strain patterns, the PRA turned out to be not suitable to investigate rare or new strains; as only the ones belonging to well-defined species (*M. gordonae*, *M. kansasii*, and *M. fortuitum*) were correctly identified.

This fact may be explained by a lack of current knowledge of all the variants of PRA patterns for a given species. Moreover, some of our PRA patterns were unreported patterns. Therefore, before describing a new PRA pattern, one should determine the identity of a species by complementary methods (e.g. 16S rDNA sequencing [20], HPLC of mycolic acids, etc.) so as not to report a faulty result. For example, isolate 94008 from a HIV-positive patient resulted in a *hsp65* sequence close to *M. gordonae* in the *hsp65* tree (Table 3 and Fig. 3). Nonetheless, 16S rDNA sequencing and AccuProbe confirmed it to be *M. kansasii*. A reconfirmation of PRA could not be performed on the initial isolate from this patient as it was lost upon subculturing. However, a strain recultured from another pathological sample from the same patient resulted in a PRA pattern typical of *M. kansasii* I. This suggested either an initial environmental contamination of the sample with *M. gordonae*, a mixed infection in this HIV-positive patient, or a laboratory error. For most of our isolates, which did not show an exactly matching sequence with either *hsp65* or 16S rDNA in the GenBank, a phylogenetic tree could provide with useful indications about their relatedness to known species. In such a case, a concording HPLC pattern with the sequence data and the place of the strain within the phylogenetic tree could lead to a potential identification.

The major discrepancies concern the three potential strains of *M. interjectum* (97047, 97050, and 97056) and the one belonging to the group *M. hiberniae* and *M. engbaekii* (98002). *M. interjectum* is scotochromogenic and our strains were either non-chromogenic or gave a variable pigmentation on successive cultures. However, a non-chromogenic variant of M. *interjectum* has already been reported [15]. Both *M. hiberniae* [21] and *M. engbaekii* [22] were described as pink, whereas in our report this strain was non-chromogenic. The TLC was not able to provide any aid as most of the strains within the genus *Mycobacterium* are known to share very common patterns. The only useful information concerned the three identical strains (98075, 98082, and 98095). They showed 98.9% similarity to IWGMT 90100 by 16S rDNA sequencing. Although, negative with the AccuProbe test for MAC, a GenBank comparison of *hsp65* showed that these three strains shared 96% similarity to *M. avium*. They all shared an identical, previously unreported HPLC pattern. Based on our results, these three isolates define a new mycobacterial species that will be described elsewhere.

Although we found HPLC to be very useful, it should be stressed that HPLC use alone may not have lead to a definite identification of mycobacteria because of the presence of overlapping patterns in some cases.

Regarding the 16S rDNA sequence comparison, Drancourt [23] suggested that $\geq$99% similarity between two isolates defines a similar species. Although acceptable for other bacteria, this definition does not apply to mycobacteria. Indeed, the latter are characterized by a unique genetic relatedness, e.g. two strains with a 99% similarity in their 16S rDNA may or may not belong to the same species [1,24]. For example, according to host-specificity, *M. tuberculosis* and *M. bovis* are considered as disctinct species, though at the molecular level these may be defined as subspecies of *M. tuberculosis* complex [2]. On the same lines, taxonomically and genetically identical mycobacterial species could present different clinical pathologies, e.g. *M. avium* and *M. paratuberculosis*, which have the same 16S rDNA sequence (both are taxonomically subspecies of *M. avium*) present distinct morphology and host-range [25]. *M. avium* leads to a disseminated disease among terminal AIDS patients while *M. paratuberculosis* causes John's disease among ruminants, and is implicated in Crohn's disease among humans [25]. In our opinion, both genotype and phenotype should be taken into account for describing new species [1]. Thus, if they share important phenotypic (HPLC pattern, growth rate, pigmentation) characteristics despite, minor sequence discrepancies, they should be considered sequevars of the same species, otherwise they should be considered to represent different species. In this view, the final identification of the strain we investigated should be: 94008 *M. kansasii*, 95001 *M. gordonae*, 95077 *M. gordonae*, 96138 *M. fortuitum*, 98007 *M. peregrinum*, 96048 *M. triviale*, new sequevar; 96071 *M. scrofulaceum*, new sequevar; 97047 *M. interjectum*, new sequevar, HPLC pattern II; 97050 *M. interjectum*, sequevar IWGMT 90203, HPLC pattern I; 97056 *M. interjectum*, sequevar IWGMT90093, HPLC pattern I; 98002 *M. hiberniae*, sequevar *M. engbaekii* (the non-official species *M. engbaekii* should be considered in our opinion as a sequevar of *M. hiberniae;* as HPLC pattern and major phenotypic characters are compatible); 97078 *M. duvalii*, new sequevar; while the cluster of three isolates (98075, 98082, and 98095) that represents a new species within the group of *M. simiae*-related mycobacteria.

The advent of sequencing technology has permitted the redefinition of existing species. However, one of the limits of the sequencing taxonomy is that numerous species remain undescribed, and that some intermediate strains may be misclassified. The *hsp65* sequencing at present has a very poor database permitting to correctly identify only well-defined species. The 16S rDNA sequencing is at present the touchstone. This report describes NTMs isolated from human pathological samples in a Caribbean setting, and the cluster of three strains (98075, 98082, and 98095) in our study may be referring to a new mycobacterial species with a tropism for such climate. Indeed, ecologic variations may influence emergence of new opportunistic pathogens, e.g. *M. avium* that was essentially a rare human pahogen prior to AIDS pandemy, turned out to be the most important opportunistic mycobacterial infection in terminal AIDS patients since late 1980s [26]. Thus, isolation and characterization of mycobacteria from new settings may lead to identify potential pathogens that may propogate in future because of increased human migration, travels, and climatic and ecological changes of the modern world.

Acknowledgements

This work was supported by the Délégation Internationale au Réseau des Instituts Pasteur et Instituts Associés, and by the Fondation Française Raoul Follereau. S.F. benefitted from a PhD fellowship partially financed by Institut Pasteur, and Agence Nationale de Recherche contre le SIDA (ANRS), Paris, France.

References

[1] Tortoli E. Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. Clin Microbiol Rev 2003;16: 319–54.

[2] Rastogi N, Legrand E, Sola C. The mycobacteria: an introduction to nomenclature and pathogenesis. Rev Sci Tech Off Int Epiz 2001;20: 21–54.

[3] Falkinham 3rd JO. Epidemiology of infection by nontuberculous mycobacteria. Clin Microbiol Rev 1996;9:177–215.

[4] von Reyn CF, Maslow JN, Barber TW, Falkinham 3rd JO, Arbeit RD. Persistent colonisation of potable water as a source of *Mycobacterium avium* infection in AIDS. Lancet 1994;343:1137–41.

[5] Komijn RE, de Haas PE, Schneider MM, Eger T, Nieuwenhuijs JH, van den Hoek RJ, Bakker D, van Zijd Erveld FG, van Soolingen JDA. Prevalence of *Mycobacterium avium* in slaughter pigs in the Netherlands and comparison of IS*1245* restriction fragment length polymorphism patterns of porcine and human isolates. J Clin Microbiol 1999;37:1254–9.

[6] Tortoli E, Piersimoni C, Kirschner P, Bartoloni A, Burrini C, Lacchini C, Mantella A, Muzzi G, Tosi CP, Penati V, Scarparo C, Simonetti MT, Böttger EC. Characterization of mycobacterial isolates phylogenetically related to, but different from *Mycobacterium simiae*. J Clin Microbiol 1997;35:697–702.

[7] Tortoli E, Mariottini A, Mazzarelli G. Evaluation of INNO-LiPA MYCOBACTERIA v2: improved reverse hybridization multiple DNA probe assay for mycobacterial identification. J Clin Microbiol 2003;41:4418–20.

[8] Ausina V, Luquin M, Garcia-Barcelo M, Lanéelle MA, Lévy-Frébault V, Belda F, Prats G. *Mycobacterium alvei* sp. nov.. Int J Syst Bact 1992;42:529–35.

[9] Centers for Disease Control and Prevention. Standardized method for HPLC identification of mycobacteria. Atlanta: US Department of Health and Human Services, Public Health Service; 1996.

[10] Devallois A, Goh KS, Rastogi N. Rapid identification of mycobacteria to species level by PCR-restriction fragment length polymorphism analysis of the *hsp65* gene and proposition of an algorithm to differentiate 34 mycobacterial species. J Clin Microbio 1997;35: 2969–73.

[11] Telenti A, Marchesi F, Balz M, Bally F, Böttger EC, Bodmer T. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. J Clin Microbiol 1993; 31:175–8.

[12] Ringuet H, Akoua-Koffi C, Honore S, Varnerot A, Vincent V, Berche P, Gaillard JL, Pierre-Audigier C. *hsp65* sequencing for identification of rapidly growing mycobacteria. J Clin Microbiol 1999;37:852–7.

[13] Kirschner P, Meier A, Böttger EC. Genotypic identification and detection of mycobacteria-facing novel and uncultured pathogens. In: Persing DH, Smith TF, Tenover FC, White TJ, editors. Diagnostic molecular microbiology. Washington, DC: American Society for Microbiology; 1993. p. 173–90.

[14] Tortoli E, Bartoloni A. High-performance liquid chromatography and identification of mycobacteria. Rev Med Microbiol 1996;7: 207–19.

[15] Tortoli E, Kirschner P, Bartoloni A, Burrini C, Manfrin V, Mantella A, Scagnelli M, Scarparo C, Simonetti MT, Böttger EC. Isolation of an unusual mycobacterium from an AIDS patient. J Clin Microbiol 1996;34:2316–9.

[16] Rastogi N. An introduction to mycobacterial taxonomy, structure, drug resistance, and pathogenesis. In: Dionisio D, editor. Textbook-atlas of intestinal infections in AIDS. Milano: Springer; 2003. p. 89–115.

[17] Korsak T, Boisvert H. Mycobacteria with pink pigment. Ann Inst Pasteur 1972;122:31–41.

[18] Wayne LG, Good RC, Böttger EC, Butler R, Dorsch M, Ezaki T, Gross W, Jonas V, Kilburn J, Kirschner P, Krichevsky MI, Ridell M, Shinninck TM, Springer B, Stackebrandt E, Tarnok I, Tarnok Z, Tasaka H, Vincent V, Warren NG, Knott CA, Johnson R. Semantide- and chemotaxonomy-based analyses of some problematic phenotypic clusters of slowly growing mycobacteria, a cooperative study of the International Working Group on Mycobacterial Taxonomy. Int J Syst Bact 1996;46:280–97.

[19] Kusunoki S, Ezaki T. Proposal of *Mycobacterium peregrinum* sp. nov., nom. rev., and elevation of *Mycobacterium chelonae* subsp. *abscessus* (Kubica et al.) to species status: *Mycobacterium abscessus* comb. nov.. Int J Syst Bact 1992;42:240–5.

[20] Springer B, Stockman L, Teschner K, Roberts GD, Böttger EC. Two-laboratory collaborative study on identification of mycobacteria: molecular versus phenotypic methods. J Clin Microbiol 1996;34: 296–303.

[21] Kazda J, Cooney R, Monaghan M, Quinn PJ, Stackebrandt E, Dorsch M, Daffe M, Muller K, Cook BR, Tarnok ZS. *Mycobacterium hiberniae* sp. nov.. Int J Syst Bact 1993;43:352–7.

[22] Marranzano M. Atypical mycobacteria isolated from the water of the Simeto River at Catania. Boll Ist Sieroter Milan 1978;57:128–33.

[23] Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 2000;38:3623–30.

[24] Tortoli E, Bartoloni A, Böttger EC, Emler S, Garzelli C, Magliano E, Mantella A, Rastogi N, Rindi L, Scarparo C, Urbano P. Burden of unidentifiable mycobacteria in a reference laboratory. J Clin Microbiol 2001;39:4058–65.

[25] Legrand E, Sola C, Rastogi N. Le complexe *Mycobacterium avium–intracellulare*: marqueurs phénotypiques et génotypiques et les bases moléculaires de la transmision inter-espèces. Bull Soc Pathol Exot 2000;93:182–92.

[26] Rastogi N, McFadden JJ. 8th Forum in microbiology, mycobacteria and AIDS: epidemiological and genetic markers, virulence factors and interactions with immune system. Res Microbiol 1992;143: 361–436.

[27] Butler WR, Guthertz LS. Mycolic acid analysis by high-performance liquid chromatography for identification of Mycobacterium species. Clin. Microbiol. Rev. 2001;14:704–26.