ORIGINAL ARTICLE BACTERIOLOGY

Re-analysis of 178 previously unidentifiable *Mycobacterium* isolates in the Netherlands in 1999–2007

J. van Ingen[1,2], R. de Zwaan[2], M. Enaimi[2], P. N. Richard Dekhuijzen[1], M. J. Boeree[1] and D. van Soolingen[2]

1) *Department of Pulmonary Diseases, Radboud University Nijmegen Medical Centre, Nijmegen and* 2) *National Mycobacteria Reference Laboratory, National Institute for Public Health and the Environment, Bilthoven, the Netherlands*

Abstract

Nontuberculous mycobacteria (NTM) that cannot be identified to the species level by reverse line blot hybridization assays and sequencing of the 16S rRNA gene comprise a challenge for reference laboratories. However, the number of 16S rRNA gene sequences added to online public databases is growing rapidly, as is the number of *Mycobacterium* species. Therefore, we re-analysed 178 *Mycobacterium* isolates with 53 previously unmatched 16S rRNA gene sequences, submitted to our national reference laboratory in 1999–2007. All sequences were again compared with the GenBank database sequences and the isolates were re-identified using two commercially available identification kits, targeting separate genetic loci. Ninety-three out of 178 isolates (52%) with 20 different 16S rRNA gene sequences could be assigned to validly published species. The two reverse line blot assays provided false identifications for three recently described species and 22 discrepancies were recorded in the identification results between the two reverse line blot assays. Identification by reverse line blot assays underestimates the genetic heterogeneity among NTM. This heterogeneity can be clinically relevant because particular sub-groupings of species can cause specific disease types. Therefore, sequence-based identification is preferable, at least at the reference laboratory level, although the exact targets needed for clinically useful results remain to be established. The number of NTM species in the environment is probably so high that unidentifiable clinical isolates should be given a separate species status only if this is clinically meaningful.

Keywords: Identification, mycobacteria, nontuberculcus mycobacteria, 16SrRNA
Original Submission: 2 September 2009; **Revised Submission:** 30 October 2009; **Accepted:** 5 November 2009
Editor: M. Drancourt
Article published online: 20 November 2009
Clin Microbiol Infect 2010; **16**: 1470–1474
10.1111/j.1469-0691.2010.03123.x

Corresponding author: J. van Ingen, RIVM; LIS (pb 22), National Mycobacteria Reference Laboratory, Antonie van Leeuwenhoeklaan 9, PO Box 1, 3720 BA Bilthoven, the Netherlands
E-mail: jakko.van.ingen@rivm.nl

Introduction

The isolation frequency of nontuberculous mycobacteria (NTM) increases in many countries where the incidence of tuberculosis is in decline [1,2]. The NTM are ubiquitous in the environment and their presence in clinical samples does not necessarily indicate NTM disease. Assessment of the clinical relevance of these isolates is not straightforward; the American Thoracic Society has published diagnostic criteria to aid in this assessment [3].

Clinical relevance differs significantly depending upon the NTM species, which makes species identification crucial [3]. Currently, NTM are mostly identified using molecular tools, such as species-specific probes, which are often incorporated in commercial line probe assays, or the direct sequencing of semi-conserved genes with proven taxonomic value. Among these, the 16S rRNA, *rpoB* and *hsp65* genes and the16S-23S internal transcribed spacer region are most commonly used [4–7]. However, not all clinical NTM isolates can be convincingly identified by molecular identification. The ongoing increase in the number of newly recognized species is testimony to this phenomenon [8].

To establish the magnitude of the problem of unidentifiable NTM, we re-investigated all NTM submitted to the national reference laboratory in the Netherlands in 1999–2007 that could previously not be convincingly identified by partial 16S rRNA gene sequencing. All sequences were re-analysed

and the isolates re-identified applying two commercially available identification kits that target two separate genetic loci.

Materials and Methods

At the national mycobacteria reference laboratory (National Institute for Public Health and the Environment; RIVM) in the Netherlands, NTM are identified using the Inno-Lipa Mycobacteria v2 (Innogenetics, Ghent, Belgium) reverse line blot assay, in accordance with the manufacturer's instructions. If no identification to the species level is obtained, additional sequencing of the hypervariable region A (151 bp) of the 16S rRNA gene is performed, but only upon request by the referring clinician. Prior to 2004, we used the AccuProbe (GenProbe, San Diego, CA, USA) assay as a first line of NTM identification.

If the 16S rRNA sequences yielded no full match with GenBank (National Center for Biotechnology Information, NCBI; http://www.ncbi.nlm.nih.gov) sequences or a match with that of a species not validly published at the time, the respective isolates were designated 'Unknown *Mycobacterium* species (UMS)' and numbered consecutively. We extracted these sequences from our database and subjected them to a new comparison with the GenBank sequences in June 2008.

To assess the impact of reverse line blot assays on the disclosure of genetic diversity among NTM, at least one isolate from each UMS was subjected to additional identification using the GenoType Mycobacterium CM/AS (Hain Lifescience, Nehren, Germany) and Inno-Lipa Mycobacteria v2 reverse line blot assays in accordance with the manufacturers' instructions.

In our re-analysis, we considered UMS with a > 1-bp sequence difference from a species type strain, or with a reverse line blot identification conflicting with the 16S rRNA gene sequence result, as being related to that species. UMS with 16S rRNA gene sequences 1 bp divergent from that of an established species were considered sequevars of that species, unless the reverse line blot identification results were discordant. UMS with 16S rRNA gene sequences identical to that of a type strain of a species available in GenBank and concordant reverse line blot results were considered to be representative of that species.

Results

We found 178 clinical isolates with 53 different 16S rRNA gene sequences not matching those of validly published species available in the GenBank database at the time of referral to the RIVM. These comprise 4% of the 4481 NTM isolates referred to the RIVM in the study period. During this period, another 913 NTM isolates (20%) were identified no further than *Mycobacterium avium* complex, except *M. avium* or *Mycobacterium intracellulare* (n = 302), *Mycobacterium fortuitum* complex (n = 104), NTM not reacting with the *M. avium* complex and the *Mycobacterium tuberculosis* complex AccuProbe kits (n = 347) or NTM reacting only with the *Mycobacterium* species probe of the Inno-Lipa assay (n = 70).

Identification by 16S rRNA gene sequencing for these 913 isolates was not requested by the referring clinicians and thus was not conducted.

The isolation frequency of the UMS and results of the new identification efforts are detailed in Table 1. Based on our re-analysis of the partial 16S rRNA gene sequence, 20 UMS (yielding 93 isolates; 52%) could be assigned to validly published species, including *Mycobacterium noviomagense* (UMS1; Table 1), which we described recently [9]. The remaining 85 isolates, comprising 33 UMS, were related to the *M. avium* complex (n = 10), *M. fortuitum* complex (n = 7), *Mycobacterium xenopi* (n = 3), *Mycobacterium terrae* complex (n = 4), *Mycobacterium gordonae* (n = 3), *Mycobacterium simiae* (n = 2), *Mycobacterium interjectum* (n = 2) or assigned to the slow or rapid growers, distantly related to established species (Table 1).

With the GenoType CM/AS assay, we identified 28 of the 53 UMS (53%) to species (n = 24) or complex level (*M. fortuitum* complex; n = 4). Twelve of these identifications were not in accordance with species or complex identifications based on the partial 16S rRNA gene sequence (Table 1). With the Inno-Lipa assay, we identified 19 UMS (36%) to species (n = 7) or complex levels (n = 12; six *M. avium* complex, six *M. fortuitum* complex). Eight identification results were discordant with species or complex identifications based on the partial 16S rRNA gene sequence (Table 1). Identifications as *M. fortuitum* complex or *M. gordonae* with the two assays are especially frequent.

Discordance between species or complex identifications with the two hybridization assays was noted in six different UMS (UMS 2, 10, 11, 13, 49 and 64; Table 1). This mainly involved isolates related to the *M. avium* complex, *M. interjectum* or *Mycobacterium scrofulaceum*.

Twenty-six (49%) of the UMS were only encountered once; the average number of isolates per UMS was 3.4 (range 1–24). Most UMS isolates were cultured from respiratory samples (n = 162; 129 sputa, 33 broncho-alveolar lavage fluid samples; 91%). Sixteen UMS isolates (9%) were cultured from normally sterile samples, including bone marrow (n = 1), lung (n = 2), lymph node (n = 4), pleura (n = 1) and joint biopsies (n = 1),

TABLE 1. Identification results of previously unidentifiable nontuberculous mycobacteria species

Unknown *Mycobacterium* species	*n*	16S hypervariable region A (GenBank)	Genotype CM/AS	Inno-Lipa	Interpretation
1	18	**100% *Mycobacterium noviomagense* NLA000500338T**	Myc	Myc	*M. noviomagense* [9]
2	2	99% *Mycobacterium intracellulare* W249st	*MSC*	MAIS	MAC
3	2	**100% *Mycobacterium nonchromogenicum* FI-06254**	Myc	Myc	*M. nonchromogenicum*
7	9	**100% *Mycobacterium arupense* DSM 44942**	Myc	Myc	*M. arupense*
10	6	99% *Mycobacterium colombiense* CIP 108962	*MINT*	MAIS	MAC
11	3	99% *Mycobacterium interjectum* DSM 44064	MIJ	*MAIS*	*M. interjectum* sqv.
12	4	**100% *Mycobacterium saskatchewanense* 00-250**	*MINT*	Myc	*M. saskatchewanense*
13	4	**100% *Mycobacterium seoulense* 03–19**	*MSC*	*MAIS*	*M. seoulense*
14	1	96% *Mycobacterium pyrenivorans* DSM 44605	Myc	Myc	RGM
15	3	**100% *Mycobacterium pulveris* CIP 106804**	Myc	Myc	*M. pulveris*
16	5	98% *M. saskatchewanense* 00–250	*MSC*	Myc	MAC
17	2	99% **Mycobacterium smegmatis** ATCC 700504	MSM	MSM	*M. smegmatis* sqv
18	3	**100% *Mycobacterium holsaticum* 1406**	Myc	*MGO*	*M. holsaticum*
19	4	**100% *Mycobacterium hiberniae* DSM 44241**	Myc	Myc	*M. hiberniae*
20	2	95% *M. nonchromogenicum* ATCC 19530	Myc	Myc	*M. terrae* related sp.
22	4	98% *Mycobacterium szulgai* CIP 104532	Myc	Myc	SGM
23	6	99% *Mycobacterium chlorophenolicum* CIP 104189	Myc	Myc	RGM
24	9	99% *Mycobacterium kumamotonense* CCUG 51961	Myc	Myc	*Mycobacterium terrae* related sp.
25	2	98% *Mycobacterium fallax* ATCC 35219	*MGO*	Myc	RGM
26	4	95% *Mycobacterium botniense* DSM 44537	*MFO2*	Myc	*M. xenopi* related sp.
27	24	**100% *Mycobacterium gordonae* FI-06271**	MGO	MGO	*M. gordonae*
28	5	**100% *Mycobacterium fortuitum* ATCC 49403**	MFO2	MFO	*M. fortuitum*
29	1	96% *Mycobacterium doricum* DSM 44339	Myc	*MFO*	RGM
30	3	98% *Mycobacterium mucogenicum* ATCC 49650	MFO2	MFO	*M. fortuitum* complex
31	1	**100% *Mycobacterium avium* 104**	MAV	MAV	*M. avium*
32	4	99% *M. terrae* ATCC 15755	Myc	Myc	*M. terrae* related sp.
33	4	99% *Mycobacterium mucogenicum* ATCC 49650	MMC	Myc	RGM
34	4	98% *Mycobacterium gordonae* CIP 104529	MGO	MGO	*M. gordonae* related sp.
35	1	**100% *Mycobacterium triviale* ATCC 23290**	Myc	Myc	*M. triviale*
36	1	98% *Mycobacterium simiae* CIP 104531	*MLE*	Myc	*M. simiae* related sp.
37	1	98% *M. simiae* CIP 104531	Myc	*MAIS*	*M. simiae* related sp.
38	4	**100% *Mycobacterium palustre* DSM 44572**	*M. palustre*	Myc	*M. palustre*
39	1	**100% *Mycobacterium lentiflavum* CIP 105465**	MLE	Myc	*M. lentiflavum*
40	1	99% *Mycobacterium gordonae* CIP 104529	MGO	MGO	*M. gordonae* sqv.
41	1	98% *Mycobacterium asiaticum* DSM 44297	*MGO*	Myc	*M. gordonae* related sp.
42	4	99% *Mycobacterium holsaticum* 1406	Myc	Myc	SGM
43	3	**100% *Mycobacterium branderi* CIP 104592**	Myc	Myc	*M. branderi*
44	2	**100% *Mycobacterium nebraskense* DSM 44803**	Myc	*MAIS*	*M. nebraskense*
45	2	**100% *Mycobacterium aurum* N196**	Myc	Myc	*M. aurum*
46	1	98% *Mycobacterium avium* ATCC 25291	MSC	Myc	MAC
48	2	94% *Mycobacterium branderi* CIP 104592	Myc	Myc	SGM
49	1	99% *Mycobacterium scrofulaceum* CIP 105416	MSC	*MINT1*	MAC
50	1	**100% *Mycobacterium cosmeticum* CIP 108169**	*MFO*	*MFO*	*M. cosmeticum*
52	1	98% *Mycobacterium sphagni* DSM 44076	MPE	MFO	*M. fortuitum* complex
53	1	97% *Mycobacterium doricum* DSM 44339	Myc	Myc	RGM
54	3	99% *Mycobacterium fortuitum* ATCC 49404	MPE	MFO	*M. fortuitum* complex
59	2	99% *Mycobacterium sphagni* DSM 44076	*MMC*	Myc	*M. fortuitum* complex
60	1	94% *Mycobacterium tusciae* CIP 106367	Myc	Myc	RGM
61	1	**100% *M. florentium* DSM 44852**	Myc	Myc	*M. florentinum*
62	1	98% *Mycobacterium porcinum* CIP 105392	*MMC*	Myc	*M. fortuitum* complex
63	1	**100% *Mycobacterium monacense* B9-21-178**	Myc	Myc	*M. monacense*
64	1	99% *Mycobacterium interjectum* DSM 44064	MIJ	Myc	*M. interjectum* sqv.
65	1	98% *Mycobacterium celatum* CIP 106109	Myc	Myc	*M. terrae* related sp.

Myc, *Mycobacterium* species; MSC, *M. scrofulaceum*; MAIS, *M. avium-intracellulare-scrofulaceum* complex; MINT, *M. intracellulare*; MIJ, *M. interjectum*; MSM, *M. smegmatis*; MGO, *M. gordonae*; MFO, *M. fortuitum*; MAV, *M. avium*; MPE, *M. peregrinum*; MMC, *M. mucogenicum*; MLE, *M. lentiflavum*; MAC, *M. avium* complex; RGM, rapid-growing *Mycobacterium*; SGM, slow-growing *Mycobacterium*; sqv., sequevar; sp., species.
Bold indicates that identification to species level was obtained after re-analysis of the 16S RNA gene sequence. Italics indicates identification with hybridization assays not in accordance with partial sequence results.

as well as urine (n = 4), gastric aspirate (n = 2) and maxillary sinus lavage fluid (n = 1).

Discussion

Unidentifiable NTM are a significant phenomenon, comprising at least 4% of all NTM submitted to our national reference laboratory. Even though their number was halved, to 2%, by our re-analysis, this remains well above the 1% estimated by Tortoli *et al.* [10]. However, we most likely underestimate the number of UMS isolates. For isolates identified to genus or complex level with the hybridization assays, additional 16S rRNA gene sequencing is not free of charge and therefore not routinely performed. Yet, this may reveal novel sequences and thus UMS. In our situation, this appears to be most prominent in isolates identified with hybridization assays as *M. fortuitum* complex, *M. avium* complex or *M. gordonae*. In the 20% of all submitted isolates that were not identified to species level, many additional UMS may be identified.

By re-analysis of the 16S rRNA sequences, we identified 20 of our 53 UMS as validly published NTM species. Forty-eight percent of the unidentifiable isolates (n = 85) represent 33 novel species or variants of established species. Many UMS are genetically related to *M. avium* and *M. fortuitum* complex members. This implies that, within these specific complexes, more species exist than are currently described. Referral to our reference laboratory may be more likely for strains considered as possible pathogens and this creates a potential selection bias.

Most of the UMS were single pulmonary isolates, which may reflect limited clinical relevance (i.e. patients failed to meet the American Thoracic Society diagnostic criteria for pulmonary NTM disease) [3] or a reluctance to submit further isolates for identification. In nine cases, UMS were isolated from normally sterile sites, and thus likely are causative agents of true NTM disease.

The current reverse line blot assays can only recognize a limited number of species. For manufacturers, it is a matter of choice to decide which species should be covered and whether recently described species should be added or replace species currently included in the assay. Many recently described species (e.g. *Mycobacterium monacense*, *Mycobacterium florentinum*; Table 1) are now identified as '*Mycobacterium* species other than those incorporated in the hybridization assay'. The results obtained in the present study demonstrate that hybridization assays also provide false identifications for some recently described species, including *Mycobacterium holsaticum*, *Mycobacterium cosmeticum* and *Mycobacterium saskatchewanense* (Table 1); users should be aware of these two pitfalls, perhaps through a note in the assay manuals.

Moreover, the hybridization assays are based on the detection of short DNA sequences. This monophasic approach precludes a high degree of genetic variation among identically identified strains and thus misses the distinction between related (sub)species. Importantly, we also recorded differences in identification results between the two hybridization assays tested. Because clinical relevance and drug susceptibility of NTM species differs, correct identification is important [3,8–11]. Sequence-based identification is more reliable than the limited approach of reverse line blot assays, although this requires sophisticated and expensive laboratory equipment and may be most suitable for reference laboratories.

For UMS, the 16S rRNA gene sequence similarity to that of an established species does not provide guidance for clinicians. In our previous identification of *M. noviomagense*, we noted that close genetic relationships with *M. xenopi* were associated with very different phenotypical features, drug susceptibility and clinical relevance [9]. Genetic relationships, based on a partial single target should, therefore, be interpreted with caution.

Identification based on DNA sequence analysis of a (partial) single gene disregards genetic variation in the rest of the genome. Our use of the hypervariable region A of the 16S rRNA gene only can be criticized because sequence variation outside this region may alter the species designation. Multigene identification may further improve our understanding of mycobacterial taxonomy and result in clinically relevant distinctions within species. *Mycobacterium kansasii* is a good example in this respect because seven subtypes have been described based on multiple genetic targets: one subtype causes pulmonary disease, one is a causative agent of HIV-related disseminated disease, whereas the five remaining types are environmental bacteria, not associated with human disease [12]. The maximum resolution of genetic identification will only be achieved after the introduction of routine sequencing of whole genomes of all available *Mycobacterium* isolates. This will lead to a robust phylogenetic tree that can be enriched with clinical data as a self-learning model to improve our understanding of mycobacterial virulence. It is conceivable that this will also lead to a complete reconsideration of the ever growing list of new species that are described on the basis of limited variation of semi-conserved genes and some degree of phenotypic variation.

It is questionable whether our UMS isolates with novel partial 16S rRNA sequences represent new species or variants of established species. Heterogeneity within the 16S rRNA gene has been described for multiple species, including *M. gordonae* [13]. Conversely, among rapid growers, new species have been described that share identical 16S rRNA genes but differ in other genetic and biochemical traits [8,14]. In this respect, *rpoB* sequences are increasingly used to define novel species [14].

What should constitute a new species? A unique 16S rRNA gene sequence remains the reference standard, although an exact cut-off point indicating distinct taxa has not been established for mycobacteria. A separate species status based entirely on unique 16S rDNA sequences would result in hundreds, if not thousands, of new species; it is doubtful whether this would serve clinicians or only add to the confusion. The results obtained in the present study demonstrate the presence of a large number of potentially new species. Moreover, human isolates represent only the tip of the 'NTM-iceberg' [8]. In this respect, we agree with Telenti [15], who has proposed that 'clinical meaningfulness should be the key to taxonomic precision', although this should include human and veterinary medicine.

Amidst an ever increasing number of species, a classification of NTM based on virulence factors, not unlike Runyon's

[16] classification based on growth rate and pigmentation, may be a future strategy. Such a classification could begin if more entire NTM genomes are sequenced [17].

In conclusion, 4% of NTM isolates submitted to our reference laboratory were unidentifiable *Mycobacterium* species. A minority was isolated from normally sterile sites or samples and may comprise causative agents of human disease. Periodical re-analysis of UMS is warranted to re-classify them; 2% remained unidentifiable after re-analysis. Identification by reverse line blot assays underestimates the genetic heterogeneity among NTM. This heterogeneity can be clinically relevant because specific (sub)species can cause specific disease types. Sequence-based identification is preferable, at least at the reference laboratory level, although adequate targets and the number of targets needed for clinically useful results remain to be established. The number of NTM species in the environment is probably so high that clinical UMS isolates should be analysed and given a separate species status only if this is clinically meaningful.

Transparency Declaration

No funding was received for the current study. None of the authors has any relationship with any entity, commercial or non-commercial, that has an interest in the subject matter of this manuscript.

References

1. Martín-Casabona N, Bahrmand AR, Bennedsen J *et al.* Non-tuberculous mycobacteria: patterns of isolation. A multi-country retrospective survey. *Int J Tuberc Lung Dis* 2004; 8: 1186–1193.
2. Marras TK, Chedore P, Ying AM, Jamieson F. Isolation prevalence of pulmonary non-tuberculous mycobacteria in Ontario, 1997–2003. *Thorax* 2007; 62: 661–666.
3. Griffith DE, Aksamit T, Brown-Elliot BA *et al.* An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am J Respir Crit Care Med* 2007; 175: 367–416.
4. Kirschner P, Springer B, Vogel U *et al.* Genotypic identification of mycobacteria by nucleic acid sequence determination: report of a 2-year experience in a clinical laboratory. *J Clin Microbiol* 1993; 31: 2882–2889.
5. Adékambi T, Colson P, Drancourt M. *rpoB*-based identification of nonpigmented and late-pigmenting rapidly growing mycobacteria. *J Clin Microbiol* 2003; 41: 5699–5708.
6. 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: 175–178.
7. Roth A, Fischer M, Hamid ME, Michalke S, Ludwig W, Mauch H. Differentiation of phylogenetically related slowly growing mycobacteria based on 16S-23S rRNA gene internal transcribed spacer sequences. *J Clin Microbiol* 1998; 36: 139–147.
8. Tortoli E. The new mycobacteria: an update. *FEMS Immunol Med Microbiol* 2006; 48: 159–178.
9. Van Ingen J, Boeree MJ, De Lange WCM *et al. Mycobacterium noviomagense* sp. nov.; clinical relevance evaluated in 17 patients. *Int J Syst Evol Microbiol* 2009; 59: 845–849.
10. Tortoli E, Bartolini A, Böttger EC *et al.* Burden of unidentifiable mycobacteria in a reference laboratory. *J Clin Microbiol* 2001; 39: 4058–4065.
11. Van Ingen J, Bendien SA, De Lange WCM *et al.* Clinical relevance of nontuberculous mycobacteria isolated in the Nijmegen-Arnhem region, the Netherlands. *Thorax* 2009; 64: 502–506.
12. Taillard C, Greub G, Weber R *et al.* Clinical implications of *Mycobacterium kansasii* species heterogeneity: Swiss national survey. *J Clin Microbiol* 2003; 41: 1240–1244.
13. Kirschner P, Böttger EC. Microheterogeneity within rRNA of *Mycobacterium gordonae*. *J Clin Microbiol* 1992; 30: 1049–1050.
14. Adékambi T, Berger P, Raoult D, Drancourt M. *rpoB* gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of *Mycobacterium bolletii* sp. nov., *Mycobacterium phocaicum* sp. nov. and *Mycobacterium aubagnense* sp. nov. *Int J Syst Evol Microbiol* 2006; 56: 133–143.
15. Telenti A. More on 'What's in a name …' – pragmatism in mycobacterial taxonomy. *Int J Tuberc Lung Dis* 1998; 2: 182–183.
16. Runyon EH. Anonymous mycobacteria in pulmonary disease. *Med Clin North Am* 1959; 43: 273–290.
17. Behr MA. Mycobacterium du jour: what's on tomorrow's menu? *Microbes Infect* 2008; 10: 968–972.