



Careful use of 16S rRNA gene sequence similarity values for the identification of *Mycobacterium* species.

Mamadou Beye, Nidal Fahsi, Didier Raoult, Pierre-Edouard Fournier

► To cite this version:

Mamadou Beye, Nidal Fahsi, Didier Raoult, Pierre-Edouard Fournier. Careful use of 16S rRNA gene sequence similarity values for the identification of *Mycobacterium* species.. *New Microbes and New Infections*, 2018, 22, pp.24-29. 10.1016/j.nmni.2017.12.009 . hal-01795912

HAL Id: hal-01795912

<https://amu.hal.science/hal-01795912>

Submitted on 1 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Careful use of 16S rRNA gene sequence similarity values for the identification of *Mycobacterium* species

M. Beye¹, N. Fahsi¹, D. Raoult^{1,2} and P.-E. Fournier¹

1) URMITE, UM63, CNRS7278, IRD198, Inserm1095, Institut Hospitalo-Universitaire Méditerranée-infection, Aix-Marseille Université, Faculté de Médecine, Marseille, France and 2) Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia

Abstract

In order to evaluate the suitability of 16S rRNA nucleotide sequence similarity for the classification of new *Mycobacterium* isolates at the species level, we systematically studied the pairwise identity values of this gene for 131 *Mycobacterium* species with standing in nomenclature. Only one of the studied species, *M. poriferae* (0.76%), strictly respected the 95% and 98.65% threshold values currently recommended to determine the affiliation of bacterial isolates to an existing or new genus or species, respectively. All other species exhibited at least an identity value >98.65% and/or <95% with another *Mycobacterium* species. Therefore, we suggest that interpretation of interspecies 16S rRNA identity values should be made cautiously when classifying a new mycobacterial isolate at the species level.

© 2017 The Authors. Published by Elsevier Ltd.

Keywords: 16S rRNA, *Mycobacterium*, sequence similarity, taxonomy, threshold value

Original Submission: 20 October 2017; **Accepted:** 20 December 2017

Article published online: 29 December 2017

Corresponding author: P.-E. Fournier, URMITE, UM63, CNRS 7278, IRD198, Inserm1095, Institut Hospitalo-Universitaire Méditerranée-infection, 19-21 Bd Jean Moulin, 13005 Marseille, France
E-mail: pierre-edouard.fournier@univ-amu.fr

Introduction

Taxonomy provides scientists with essential information, enabling them to understand the relationships between living organisms and their different ecosystems [1]. For prokaryotes, taxonomy allows the reliable identification of microbial strains from clinical or environmental samples [2]. Bacterial taxonomy was initiated in the late 19th century, when phenotypic characteristics were incorporated into bacterial description, including motility, growth requirements, morphology, staining properties, colony size and colour, and chemical reactions [3]. Between the mid-1950s and the 1980s, new parameters were progressively added, notably chemotaxonomy [4], numerical taxonomy, genomic DNA-DNA hybridization and G+C content [5]. In the 1980s, the advent of DNA amplification and sequencing techniques, in particular of the 16S rRNA gene,

constituted a major step forwards by facilitating bacterial classification [6,7]. The 16S rRNA gene is a highly conserved gene that is made of nine hypervariable domains separated by more preserved fragments in which universal primers can be designed. More than three million 16S rRNA gene sequences are currently available in public databases [8]. In 1996, Vandamme et al. [9] suggested that polyphasic taxonomy, which takes into account all available phenotypic and genotypic data and integrates them into a consensus classification, should include 16S rRNA gene sequence identity. In 2010, Tindall et al. [10], in a reevaluation of the various available methods, proposed a combination of phenotypic and genotypic criteria within which 16S rRNA gene sequence similarity, and phylogeny was included as a first-line tool.

In 1994, scientists considered two strains as belonging to different species if they shared 16S rRNA gene sequence similarity values <97% and to a distinct genus if this value was <95% [11]. The cutoff value at the species level was later reevaluated at 98.7% [12] and then 98.65% [13]. However, several authors have shown that these thresholds, originally designed to standardize the use of sequences of 16S rRNA genes in taxonomy, are not applicable to multiple genera. In 2015, we demonstrated that many of the current bacterial species with validly published names do not respect the 95% and 98.7% thresholds [14].

TABLE I. List of species with standing in nomenclature used in our study

Species	Accession no.	Size (bp)
<i>Mycobacterium abscessus</i> subsp. <i>bolletii</i>	AY859681	1481
<i>Mycobacterium africanum</i>	AF480605	1433
<i>Mycobacterium agri</i>	AJ429045	1456
<i>Mycobacterium aichiense</i>	X55598	1456
<i>Mycobacterium alvei</i>	AF023664	1465
<i>Mycobacterium anyangense</i>	KJ855063	1420
<i>Mycobacterium arosiense</i>	EF054881	1493
<i>Mycobacterium arupense</i>	DQ157760	1487
<i>Mycobacterium asiaticum</i>	AF480595	1466
<i>Mycobacterium aubagnense</i>	AY859683	1482
<i>Mycobacterium aurum</i>	X55595	1458
<i>Mycobacterium austroafricanum</i>	X93182	1462
<i>Mycobacterium avium</i> subsp. <i>avium</i>	AJ536037	1472
<i>Mycobacterium avium</i> subsp. <i>silvaticum</i>	EF521891	1442
<i>Mycobacterium bouchodurhonense</i>	EF591053	1498
<i>Mycobacterium branderi</i>	AF480574	1469
<i>Mycobacterium brisbanense</i>	AY012577	1499
<i>Mycobacterium brumae</i>	AF480576	1449
<i>Mycobacterium canariense</i>	AY255478	1433
<i>Mycobacterium caprae</i>	AJ131120	1524
<i>Mycobacterium celatum</i>	L08169	1460
<i>Mycobacterium celeriflavum</i>	KJ607136	1442
<i>Mycobacterium chelonae</i>	AY457072	1481
<i>Mycobacterium chitae</i>	X55603	1457
<i>Mycobacterium chlorophenolicum</i>	X79292	1466
<i>Mycobacterium chubuense</i>	AF480597	1472
<i>Mycobacterium conceptionense</i>	AY859684	1483
<i>Mycobacterium confluentis</i>	AJ634379	1504
<i>Mycobacterium conspicuum</i>	X88922	1433
<i>Mycobacterium cosmeticum</i>	AY449728	1507
<i>Mycobacterium crocinum</i>	DQ534008	1398
<i>Mycobacterium diernhoferi</i>	AF480599	1458
<i>Mycobacterium doricum</i>	AF264700	1450
<i>Mycobacterium duvalii</i>	U94745	1502
<i>Mycobacterium elephantis</i>	AJ010747	1517
<i>Mycobacterium fallax</i>	AF480600	1470
<i>Mycobacterium farcinogenes</i>	AY457084	1483
<i>Mycobacterium flavescens</i>	X52932	1454
<i>Mycobacterium fluoranthenorans</i>	AJ617741	1494
<i>Mycobacterium fortuitum</i> subsp. <i>acetamidolyticum</i>	FR733720	1505
<i>Mycobacterium fortuitum</i> subsp. <i>fortuitum</i>	AY457066	1483
<i>Mycobacterium fragae</i>	JQ898451	1452
<i>Mycobacterium frederiksbergense</i>	AJ276274	1474
<i>Mycobacterium gadium</i>	X55594	1456
<i>Mycobacterium gastri</i>	AF480602	1469
<i>Mycobacterium genavense</i>	X60070	1449
<i>Mycobacterium goodii</i>	Y12872	1417
<i>Mycobacterium gordonae</i>	X52923	1461
<i>Mycobacterium hassiacum</i>	U49401	1491
<i>Mycobacterium heidelbergense</i>	AJ000684	1445
<i>Mycobacterium hodleri</i>	X93184	1459
<i>Mycobacterium holsaticum</i>	AJ310467	1526
<i>Mycobacterium houstonense</i>	AY457067	1483
<i>Mycobacterium interjectum</i>	HM037998	1431
<i>Mycobacterium intermedium</i>	X67847	1441
<i>Mycobacterium intracellulare</i>	AJ536036	1440
<i>Mycobacterium iranica</i>	HQ009482	1450
<i>Mycobacterium kansasii</i>	AJ536035	1470
<i>Mycobacterium komossense</i>	X55591	1462
<i>Mycobacterium koreense</i>	JF271826	1474
<i>Mycobacterium kubicae</i>	AF133902	1321
<i>Mycobacterium kyorinense</i>	AB370111	1470
<i>Mycobacterium lacus</i>	AF406783	1470
<i>Mycobacterium lentiflavum</i>	AF480583	1452
<i>Mycobacterium litorale</i>	GU997640	1380
<i>Mycobacterium llatzerense</i>	AJ746070	1397
<i>Mycobacterium madagascariense</i>	AB537170	1470
<i>Mycobacterium mageritense</i>	AJ699399	1497
<i>Mycobacterium malmoense</i>	X52930	1457
<i>Mycobacterium manteni</i>	FJ042897	1471
<i>Mycobacterium marinum</i>	AF456240	1522
<i>Mycobacterium marseillense</i>	EU266632	1440
<i>Mycobacterium microti</i>	AF480584	1484
<i>Mycobacterium monacense</i>	AF107039	1473
<i>Mycobacterium morioakaense</i>	AJ429044	1493
<i>Mycobacterium mucogenicum</i>	AY457074	1482
<i>Mycobacterium murale</i>	AB537171	1459
<i>Mycobacterium nebraskense</i>	AY368456	1506
<i>Mycobacterium neoaurum</i>	AF480593	1470
<i>Mycobacterium neworleansense</i>	AY457068	1483
<i>Mycobacterium noviomagense</i>	EU239955	1478
<i>Mycobacterium novocastrense</i>	U96747	1513
<i>Mycobacterium obuense</i>	X55597	1458
<i>Mycobacterium pallens</i>	DQ370008	1435
<i>Mycobacterium paraense</i>	KJ948996	1480

TABLE I. Continued

Species	Accession no.	Size (bp)
<i>Mycobacterium paraffinicum</i>	GQ153270	1492
<i>Mycobacterium parafortuitum</i>	X93183	1460
<i>Mycobacterium paragordoniae</i>	KC525204	1393
<i>Mycobacterium parakoreense</i>	JF271823	1465
<i>Mycobacterium parascrofulaceum</i>	AY337273	1468
<i>Mycobacterium paraseoulense</i>	DQ536404	1522
<i>Mycobacterium paratuberculosis</i>	X52934	1458
<i>Mycobacterium parmense</i>	AF466821	1529
<i>Mycobacterium peregrinum</i>	AY457069	1483
<i>Mycobacterium phlei</i>	AF480603	1461
<i>Mycobacterium phocaicum</i>	AY859682	1482
<i>Mycobacterium porcinum</i>	AY457077	1483
<i>Mycobacterium poriferae</i>	AF480589	1449
<i>Mycobacterium pseudoshottii</i>	AY570988	1453
<i>Mycobacterium psychrotolerans</i>	AJ534886	1516
<i>Mycobacterium pulveris</i>	AJ429046	1492
<i>Mycobacterium pyrenivorans</i>	AJ431371	1481
<i>Mycobacterium rhodesiae</i>	AJ429047	1485
<i>Mycobacterium riyadhense</i>	EU274642	1475
<i>Mycobacterium rufum</i>	AY943385	1322
<i>Mycobacterium rutilum</i>	DQ370011	1417
<i>Mycobacterium scrofulaceum</i>	AF480604	1466
<i>Mycobacterium sediminis</i>	KC010490	1515
<i>Mycobacterium senegalense</i>	AY457081	1483
<i>Mycobacterium sensuense</i>	DQ536408	1526
<i>Mycobacterium seoulense</i>	DQ536403	1522
<i>Mycobacterium septicum</i>	AY457070	1483
<i>Mycobacterium setense</i>	EF138818	1336
<i>Mycobacterium sherrisii</i>	AY353699	1510
<i>Mycobacterium shinjukuense</i>	AB268503	1505
<i>Mycobacterium shottii</i>	AY005147	1491
<i>Mycobacterium simiae</i>	X52931	1479
<i>Mycobacterium smegmatis</i>	AJ131761	1482
<i>Mycobacterium sphagni</i>	FR733719	1505
<i>Mycobacterium stomatopiae</i>	AM884331	1471
<i>Mycobacterium szulgai</i>	X52926	1454
<i>Mycobacterium thermoresistibile</i>	X55602	1464
<i>Mycobacterium tokaiense</i>	AF480590	1451
<i>Mycobacterium triplex</i>	U57632	1474
<i>Mycobacterium triviale</i>	DQ058405	1362
<i>Mycobacterium tusciae</i>	AF058299	1409
<i>Mycobacterium ulcerans</i>	AB548725	1475
<i>Mycobacterium vaccae</i>	AF480591	1439
<i>Mycobacterium vulneris</i>	EU834055	1471
<i>Mycobacterium wolinskyi</i>	AY457083	1485
<i>Mycobacterium yongonense</i>	JF738056	1395

In 2000, Woo et al. [15] proposed that 16S rRNA gene sequencing was the reference standard for the identification of *Mycobacterium* species. Genotypic investigations based on the sequencing of the 16S rRNA gene have played a significant role in the taxonomic classification of members of the genus *Mycobacterium* [16]. However, to date, no systematic study of the degree of 16S rRNA divergence among *Mycobacterium* species has been conducted.

Here we evaluate the value of current 16S rRNA cutoff values at the species and genus levels by systematically calculating the pairwise degree of 16S rRNA similarity between all *Mycobacterium* species with standing in nomenclature.

Methods

Collection of 16S rRNA gene sequences from members of the genus *Mycobacterium*

Within the List of Prokaryotic Names with Standing in Nomenclature website (<http://www.bacterio.net/mycobacterium.html>),

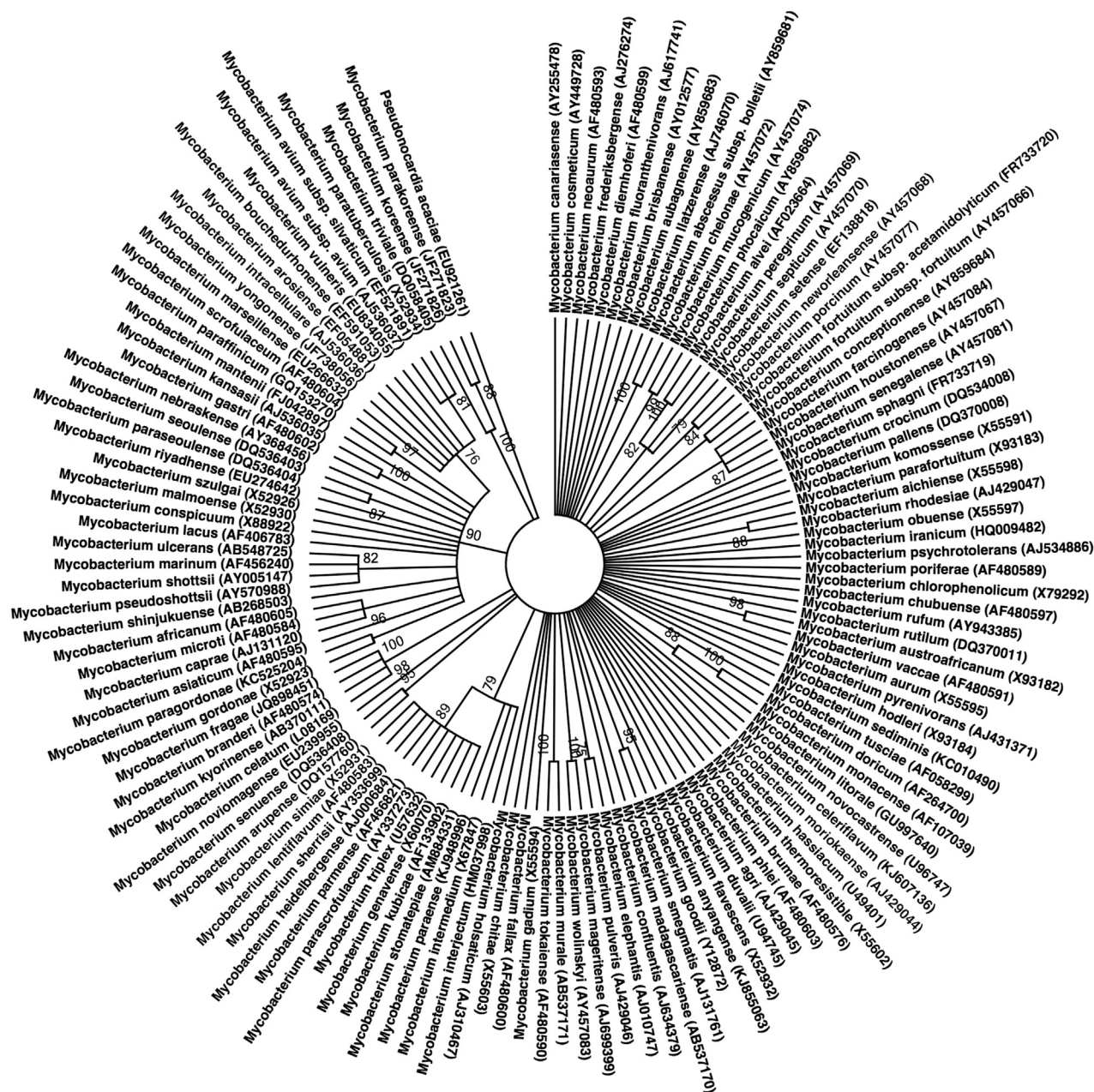


FIG. 1. Phylogenetic distribution of *Mycobacterium* species used in present study based on comparison of 16S rRNA sequences. Sequences were aligned by MUSCLE [14], and phylogenetic inferences were obtained by maximum likelihood method and Kimura two-parameter model in MEGA software. Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 1000 times to generate majority consensus tree. *Pseudonocardia acaciae* (EU921261) was used as outgroup.

we selected all *Mycobacterium* species with a validly published name as of 25 March 2016, and we collected the 16S rRNA gene accession numbers from type strains.

As a result of the wide heterogeneity in length and quality of the 16S rRNA gene sequences of type strains, we did not use sequences shorter than 1320 nt. We created a FASTA format file containing all selected sequences.

16S rRNA gene sequence analysis: calculation of pairwise 16S rRNA gene sequence similarities

Sequences were aligned using Muscle software with default settings [17]. In this study, pairwise 16S rRNA gene sequence similarities between all species of the genus *Mycobacterium* were first estimated by MEGA 5 phylogeny software [18]. Then the highest and lowest values computed by this software were

TABLE 2. Species that do not respect pairwise similarity thresholds of <95% and >98.65%

Species (accession no.)	No. for <95% threshold	No. for >98.65% threshold
<i>Mycobacterium abscessus</i> subsp. <i>bolletii</i> (AY859681)	41	2
<i>Mycobacterium africanum</i> (AF480605)	23	7
<i>Mycobacterium agri</i> (AJ429045)	32	0
<i>Mycobacterium aichiense</i> (X55598)	18	1
<i>Mycobacterium alvei</i> (AF023664)	17	13
<i>Mycobacterium anyangense</i> (KJ855063)	3	5
<i>Mycobacterium arsiense</i> (EF054881)	16	11
<i>Mycobacterium arupense</i> (DQ157760)	12	0
<i>Mycobacterium asiaticum</i> (AF480595)	6	4
<i>Mycobacterium aubagnense</i> (AY859683)	6	3
<i>Mycobacterium aurum</i> (X55595)	23	0
<i>Mycobacterium austroafricanum</i> (X93182)	17	1
<i>Mycobacterium avium</i> subsp. <i>avium</i> (AJ536037)	30	9
<i>Mycobacterium avium</i> subsp. <i>silvaticum</i> (EF521891)	23	13
<i>Mycobacterium bouchodurhanense</i> (EF591053)	20	11
<i>Mycobacterium branderi</i> (AF480574)	66	0
<i>Mycobacterium brisbanense</i> (AY012577)	1	1
<i>Mycobacterium brumae</i> (AF480576)	32	0
<i>Mycobacterium canariense</i> (AY255478)	20	4
<i>Mycobacterium caprae</i> (AJ131120)	17	7
<i>Mycobacterium celatum</i> (L08169)	23	0
<i>Mycobacterium celeriflavum</i> (KJ607136)	0	1
<i>Mycobacterium chelonae</i> (AY457072)	47	2
<i>Mycobacterium chitae</i> (X55603)	17	0
<i>Mycobacterium chlorophenolicum</i> (X79292)	0	3
<i>Mycobacterium chubuense</i> (AF480597)	0	1
<i>Mycobacterium conceptionense</i> (AY859684)	2	17
<i>Mycobacterium confluentis</i> (AJ634379)	1	0
<i>Mycobacterium conspicuum</i> (X88922)	8	3
<i>Mycobacterium cosmeticum</i> (AY449728)	6	6
<i>Mycobacterium crocinum</i> (DQ534008)	5	6
<i>Mycobacterium diernhoferi</i> (AF480599)	23	5
<i>Mycobacterium doricum</i> (AF264700)	42	1
<i>Mycobacterium duvalii</i> (U94745)	3	2
<i>Mycobacterium elephantis</i> (AJ010747)	6	0
<i>Mycobacterium fallax</i> (AF480600)	2	0
<i>Mycobacterium farcinogenes</i> (AY457084)	2	21
<i>Mycobacterium flavescens</i> (X52932)	36	0
<i>Mycobacterium fluoranthenorans</i> (AJ617741)	10	5
<i>Mycobacterium fortuitum</i> subsp. <i>fortuitum</i> (AY457066)	6	16
<i>Mycobacterium fortuitum</i> subsp. <i>acetamidolyticum</i> (FR733720)	6	16
<i>Mycobacterium fragae</i> (JQ898451)	17	0
<i>Mycobacterium frederiksborgense</i> (AJ276274)	11	3
<i>Mycobacterium gadium</i> (X55594)	6	0
<i>Mycobacterium gastri</i> (AF480602)	26	7
<i>Mycobacterium genavense</i> (X60070)	11	4
<i>Mycobacterium goodii</i> (Y12872)	11	3
<i>Mycobacterium gordonae</i> (X52923)	27	1
<i>Mycobacterium hassiacum</i> (U49401)	52	0
<i>Mycobacterium heidelbergense</i> (AJ000684)	5	5
<i>Mycobacterium hodleri</i> (X93184)	5	0
<i>Mycobacterium holsaticum</i> (AJ310467)	1	0
<i>Mycobacterium houstonense</i> (AY457067)	2	21
<i>Mycobacterium interjectum</i> (HM037998)	1	1
<i>Mycobacterium intermedium</i> (X67847)	2	0
<i>Mycobacterium intracellulare</i> (AJ536036)	29	12
<i>Mycobacterium iranica</i> (HQ009482)	13	0
<i>Mycobacterium kansaii</i> (AJ536035)	26	7
<i>Mycobacterium komossense</i> (X55591)	10	0
<i>Mycobacterium koreense</i> (JF271826)	6	1
<i>Mycobacterium kubicae</i> (AF133902)	1	1
<i>Mycobacterium kyorinense</i> (AB370111)	70	0
<i>Mycobacterium lacus</i> (AF406783)	17	14
<i>Mycobacterium lentiflavum</i> (AF480583)	4	6
<i>Mycobacterium litorale</i> (GU997640)	20	0
<i>Mycobacterium llutzerense</i> (AJ746070)	14	1
<i>Mycobacterium madagascariense</i> (AB537170)	3	1
<i>Mycobacterium mageritense</i> (AJ699399)	0	1
<i>Mycobacterium malmoense</i> (X52930)	57	5
<i>Mycobacterium mantonii</i> (FJ042897)	5	7
<i>Mycobacterium marinum</i> (AF456240)	12	7
<i>Mycobacterium marseillense</i> (EU266632)	22	14
<i>Mycobacterium microti</i> (AF480584)	23	6
<i>Mycobacterium monacense</i> (AF107039)	35	1
<i>Mycobacterium moriokaense</i> (AJ429044)	0	3
<i>Mycobacterium mucogenicum</i> (AY457074)	4	11
<i>Mycobacterium murale</i> (AB537171)	1	0
<i>Mycobacterium nebraskense</i> (AY368456)	20	12
<i>Mycobacterium neoaurum</i> (AF480593)	5	4
<i>Mycobacterium neworleansense</i> (AY457068)	2	17
<i>Mycobacterium noviomagensis</i> (EU239955)	48	0

TABLE 2. Continued

Species (accession no.)	No. for <95% threshold	No. for >98.65% threshold
<i>Mycobacterium novocastrense</i> (U96747)	44	0
<i>Mycobacterium obuense</i> (X55597)	21	0
<i>Mycobacterium pallens</i> (DQ370008)	5	6
<i>Mycobacterium paraense</i> (KJ948996)	0	1
<i>Mycobacterium paraffinicum</i> (GQ153270)	3	6
<i>Mycobacterium parafortuitum</i> (X93183)	3	6
<i>Mycobacterium paragordoniae</i> (KC525204)	17	2
<i>Mycobacterium parakoreense</i> (JF271823)	2	0
<i>Mycobacterium parascrofulaceum</i> (AY337273)	6	1
<i>Mycobacterium paraseoulense</i> (DQ536404)	10	16
<i>Mycobacterium paratuberculosis</i> (X52934)	45	8
<i>Mycobacterium parmense</i> (AF466821)	46	0
<i>Mycobacterium peregrinum</i> (AY457069)	9	15
<i>Mycobacterium phlei</i> (AF480603)	4	0
<i>Mycobacterium phocaicum</i> (AY859682)	4	10
<i>Mycobacterium porcinum</i> (AY457077)	4	16
<i>Mycobacterium poriferum</i> (AF480589)	0	0
<i>Mycobacterium pseudoshottisii</i> (AY570988)	22	6
<i>Mycobacterium psychrotolerans</i> (AJ534886)	6	1
<i>Mycobacterium pulvum</i> (AJ429046)	2	1
<i>Mycobacterium pyrenivorans</i> (AJ431371)	4	0
<i>Mycobacterium rhodesiae</i> (AJ429047)	6	12
<i>Mycobacterium riyadhense</i> (EU274642)	10	18
<i>Mycobacterium rufum</i> (AY943385)	37	0
<i>Mycobacterium rutilum</i> (DQ370011)	17	6
<i>Mycobacterium scrofulaceum</i> (AF480604)	7	8
<i>Mycobacterium sediminis</i> (KC010490)	37	0
<i>Mycobacterium senegalense</i> (AY457081)	2	21
<i>Mycobacterium senuense</i> (DQ536408)	30	0
<i>Mycobacterium seoulense</i> (DQ536403)	10	16
<i>Mycobacterium septicum</i> (AY457070)	8	13
<i>Mycobacterium setense</i> (EF138818)	14	16
<i>Mycobacterium sherrisii</i> (AY353699)	3	7
<i>Mycobacterium shinjukuense</i> (AB268503)	42	0
<i>Mycobacterium shottisii</i> (AY005147)	16	7
<i>Mycobacterium simiae</i> (X52931)	6	6
<i>Mycobacterium smegmatis</i> (AJ131761)	15	7
<i>Mycobacterium sphagni</i> (FR733719)	3	10
<i>Mycobacterium stomatopneumoniae</i> (AM884331)	2	4
<i>Mycobacterium szulgai</i> (X52926)	33	2
<i>Mycobacterium thermoresistibile</i> (X55602)	48	1
<i>Mycobacterium tokaiense</i> (AF480590)	0	1
<i>Mycobacterium triplex</i> (U57632)	3	6
<i>Mycobacterium triviale</i> (DQ058405)	2	0
<i>Mycobacterium tusciae</i> (AF058299)	25	0
<i>Mycobacterium ulcerans</i> (AB548725)	16	7
<i>Mycobacterium vaccae</i> (AF480591)	31	1
<i>Mycobacterium vulneris</i> (EU834055)	26	12
<i>Mycobacterium wolinskyi</i> (AY457083)	2	0
<i>Mycobacterium yangonense</i> (JF738056)	15	14

For each studied species, we indicate for each threshold numbers of pairwise comparisons for which abnormal values were observed.

more accurately determined by pairwise BLASTN. We defined as expected values of interspecies 16S rRNA gene sequence similarity percentages that were between 95% and 98.65% or intraspecies percentages that were greater than 98.65% [13], and as abnormal values interspecies percentages that were >98.65% or <95% [14] or intraspecies percentages of <98.65%.

Results

Of the 182 *Mycobacterium* species and subspecies with a validly published name at the time of our study and for which a 16S rRNA sequence was available, we included 131 species with 16S rRNA sequences longer than 1320 nt (Table 1). For two of those species, *M. avium* and *M. fortuitum*, we included two subspecies (Table 1). The phylogenetic distribution of the

studied *Mycobacterium* species is presented in Fig. 1. Among the 131 studied species, the pairwise 16S rRNA gene sequence similarity values ranged from 93.00% between *M. chelonae* and *M. kyorinense* to 100% between *M. fortuitum* subsp. *acetamidolyticum* and *M. fortuitum* subsp. *fortuitum*, *M. africanum* and *M. caprae*, *M. farcinogenes*, *M. houstonense* and *M. senegalense*, *M. gastri* and *M. kansasii*, *M. mucogenicum* and *M. phocaicum*, *M. murale* and *M. tokaiense*, and *M. paraseoulense* and *M. seoulense*, respectively (Supplementary Table S1).

Of the 131 studied *Mycobacterium* species, 90 (68.7%) exhibited at least one 16S rRNA gene sequence similarity value greater than 98.65% with another species in this genus (Table 2, Supplementary Table S1). Among 131 studied species, 123 (93.9%) exhibited at least one 16S rRNA gene sequence similarity value lower than 95% with another species in the genus (Table 2, Supplementary Table S1). Only one (0.76%) of the 131 studied species, i.e. *M. poriferae*, exhibited only expected values (Table 2, Supplementary Table S1). At the intraspecies level, only expected values were observed.

Discussion

Over the past decade, several authors suggested that the inter- and intraspecies discriminatory power of 16S rRNA gene sequences was insufficient for some bacterial genera [19,20]. As examples, *Streptococcus pneumoniae* and *S. mitis* exhibit only a 3 nt difference (99.7% identity), which would classify them in the same species. In contrast, major interspecies differences may be observed, as is the case in the genus *Clostridium*, with *C. tetani* and *C. innocuum* exhibiting a 104 nt divergence (93.7% identity). The strict application of the 95% threshold would justify their classification in distinct genera [19]. In addition, in 2010, Pei et al. [21] identified an intragenomic sequence divergence greater than 1.3% among 16S rRNA genes copies in 11 bacterial species. Among these, *Borrelia afzelii*, an agent of Lyme disease in humans, exhibits a similarity of only 79.62% between its two 16S rRNA gene copies [21]. Thus, a strict application of the 98.65% threshold would classify these bacteria in different species depending on the 16S rRNA gene copy analysed [21,22]. According to Rossi-Tamisier et al. [14], among 158 studied bacterial genera, only members of 17 genera strictly respected the 95% and 98.65% thresholds. Among other studied genera, the percentage of species that respected strictly both thresholds varied from 0 (*Brucella*) to 93.9% (*Nocardia*) [14].

In the present report, we observed that the currently used 16S rRNA gene sequence similarity thresholds for delineating bacterial species are valid for only 0.76% of 131 studied *Mycobacterium* species with standing in nomenclature. Because our

study covers 71.97% of the currently validly published *Mycobacterium* species names, we believe that the 95% and 98.65% thresholds are not suitable for this genus and should at the maximum be used as indicators, not as a reference standard, for classifying new *Mycobacterium* species.

Acknowledgement

This study was funded by the Fondation Méditerranée Infection.

Conflict of interest

None declared.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.nmni.2017.12.009>.

References

- [1] Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, et al. Opinion: Re-evaluating prokaryotic species. *Nat Rev Microbiol* 2005;3: 733–9.
- [2] Moore ER, Mihaylova SA, Vandamme P, Krichevsky MI, Dijkshoorn L. Microbial systematics and taxonomy: relevance for a microbial commons. *Res Microbiol* 2010;161:430–8.
- [3] Lehmann KB, Neumann R. Atlas und Gundriss des Bakteriologie und Lehrbuch der speziellen bakteriologischen Diagnostik. 1896. Munich.
- [4] Minnikin DE, Alshamaony L, Goodfellow M. Differentiation of *Mycobacterium*, *Nocardia*, and related taxa by thin-layer chromatographic analysis of whole-organism methanolsates. *J Gen Microbiol* 1975;88: 200–4.
- [5] Brenner DJ, Fanning GR, Johnson KE, Citarella RV, Falkow S. Polynucleotide sequence relationships among members of *Enterobacteriaceae*. *J Bacteriol* 1969;98:637–50.
- [6] Coenye T, Vandamme P. Use of the genomic signature in bacterial classification and identification. *Syst Appl Microbiol* 2004;27:175–85.
- [7] Konstantinidis KT, Tiedje JM. Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Curr Opin Microbiol* 2007;10:504–9.
- [8] Quast C, Priesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;41:D590–6.
- [9] Vandamme P, Pot B, Gillis M, De Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* 1996;60:407–38.
- [10] Tindall BJ, Rossello-Mora R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 2010;60:249–66.
- [11] Stackebrandt E, Goebel BM. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bact* 1994;44:846–9.

- [12] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006;33:152–5.
- [13] Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–51.
- [14] Rossi-Tamisier M, Benamar S, Raoult D, Fournier PE. Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *Int J Syst Evol Microbiol* 2015;65: 1929–34.
- [15] Woo PC, Tsoi HW, Leung KW, Lum PN, Leung AS, Ma CH, et al. Identification of *Mycobacterium neoaurum* isolated from a neutropenic patient with catheter-related bacteremia by 16S rRNA sequencing. *J Clin Microbiol* 2000;38:3515–7.
- [16] Tortoli E. Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. *Clin Microbiol Rev* 2003;16:319–54.
- [17] Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform* 2004;5:113.
- [18] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;28:2731–9.
- [19] Bosshard PP, Zbinden R, Abels S, Boddingtonhaus B, Altwegg M, Bottger EC. 16S rRNA gene sequencing versus the API 20 NE system and the VITEK 2 ID-GNB card for identification of nonfermenting Gram-negative bacteria in the clinical laboratory. *J Clin Microbiol* 2006;44:1359–66.
- [20] Mignard S, Flandrois JP. 16S rRNA sequencing in routine bacterial identification: a 30-month experiment. *J Microbiol Methods* 2006;67: 574–81.
- [21] Pei AY, Oberdorf WE, Nossa CW, Agarwal A, Chokshi P, Gerz EA, et al. Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol* 2010;76:3886–97.
- [22] Acinas SG, Marcelino LA, Klepac-Ceraj V, Polz MF. Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J Bacteriol* 2004;186:2629–35.