Emended description of *Mycobacterium abscessus*, *Mycobacterium abscessus* subsp. *abscessus* and *Mycobacterium abscessus* subsp. *bolletii* and designation of *Mycobacterium abscessus* subsp. *massiliense* comb. nov.

Enrico Tortoli,¹ Thomas A. Kohl,² Barbara A. Brown-Elliott,³ Alberto Trovato,¹ Sylvia Cardoso Leão,⁴ Maria Jesus Garcia,⁵ Sruthi Vasireddy,³ Christine Y. Turenne,⁶ David E. Griffith,⁷ Julie V. Philley,⁷ Rossella Baldan,¹ Silvia Campana,⁸ Lisa Cariani,⁹ Carla Colombo,¹⁰ Giovanni Taccetti,⁸ Antonio Teri,⁹ Stefan Niemann,² Richard J. Wallace Jr.³ and Daniela M. Cirillo¹

¹Emerging Bacterial Pathogens Unit, San Raffaele Scientific Institute, Milan, Italy
²Molecular and Experimental Mycobacteriology, Forschungszentrum Borstel, Leibniz-Zentrum für Medizin und Biowissenschaften, Borstel, Germany
³Mycobacteria/Nocardia Research Laboratory, Department of Microbiology, The University of Texas Health Science Center at Tyler, Tyler, TX, USA
⁴Departamento de Microbiologia, Imunologia e Parasitologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil
⁵Department of Preventive Medicine, Public Health and Microbiology, Autonoma University of Madrid, Madrid, Spain
⁶Diagnostic Services Manitoba, Winnipeg, Manitoba, Canada
⁷Department of Pulmonary Medicine, The University of Texas Health Science Center at Tyler, Tyler, TX, USA
⁸Regional Reference Center for Cystic Fibrosis, Meyer University Hospital, Florence, Italy
⁹Cystic Fibrosis Microbiology Laboratory, IRCCS Ca’ Granda, Milan, Italy
¹⁰Cystic Fibrosis Center, IRCSS Ca’ Granda, Milan, Italy

The taxonomic position of members of the *Mycobacterium abscessus* complex has been the subject of intensive investigation and, in some aspects confusion, in recent years as a result of varying approaches to genetic data interpretation. Currently, the former species *Mycobacterium massiliense* and *Mycobacterium bolletii* are grouped together as *Mycobacterium abscessus* subsp. *bolletii*. They differ greatly, however, as the former *M. bolletii* has a functional *erm(41)* gene that confers inducible resistance to macrolides, the primary therapeutic antimicrobials for *M. abscessus*, while in the former *M. massiliense* the *erm(41)* gene is non-functional. Furthermore, previous whole genome studies of the *M. abscessus* group support the separation of *M. bolletii* and *M. massiliense*. To shed further light on the population structure of *Mycobacterium abscessus*, 43 strains and three genomes retrieved from GenBank were subjected to pairwise comparisons using three computational approaches: verage ucleotide

**Abbreviations:** ANI, average nucleotide identity; DDH, DNA-DNA hybridization; GGD, genome to genome distance; SNP, single nucleotide polymorphism; WGS, whole genome sequencing.

The GenBank/EMBL/DDBJ accession numbers for the whole genome sequences of strains investigated in the study are ERS1157708–ERS1157749. One supplementary table is available with the online Supplementary Material.
Single gene sequencing, such as the 16S rRNA or rpoB genes, has contributed to the discovery of novel taxa and to re-defining bacterial phylogenetic relationships as a whole. The increasing accessibility of whole genome sequencing (WGS) will influence and help define bacterial speciation, sub-speciation and epidemiology in novel ways.

The first report of *Mycobacterium abscessus* dates back to 1953 (Moore & Frerichs, 1953). In 1972, it was re-classified as a subspecies of *Mycobacterium chelonae* (Kubica et al., 1972). Twenty years later it again was elevated to species level (Kusunoki & Ezaki, 1992).

*Mycobacterium bolletii* (Adékambi et al., 2004, 2006b) were first described as novel species closely related to *M. abscessus* on the basis of 16S rRNA gene sequence. Upon observation that these two species, presenting complete identity of 16S rRNA gene sequence, could not be separated from *M. abscessus* either by phenotypic features or DNA–DNA hybridization (DDH; all three subspecies having >70% relatedness to each other), Leao et al. (2011) proposed combining *M. abscessus*, *M. bolletii* and *M. massiliense* in a single species: *M. abscessus*. Within this species two subspecies, namely *M. abscessus* subsp. *abscessus* and *M. abscessus* subsp. *bolletii* were distinguished, the latter including those isolates previously identified as *M. massiliense* or *M. bolletii*.

However, over recent years, mounting evidence has emerged that the strains included in the grouping of *M. abscessus* subsp. *bolletii* are not homogeneous (Cho et al., 2013; Sassi & Drancourt, 2014; Tan et al., 2015). While all of them cause infections in humans, differences in susceptibility to macrolides, based on sequence and functionality differences in an *erm* gene, greatly influence the clinical treatment outcome (Kim et al., 2010).

In 2006 it was for the first time reported that most clinically significant rapidly growing mycobacteria (RGM) in the *Mycobacterium fortuitum* group contain an inducible erythromycin ribosomal methylase (*erm*) gene that confers macrolide resistance (Nash et al., 2006). At that time, macrolides were the treatment of choice for most susceptible species of RGM, including *M. chelonae* and *M. abscessus* (Griffith et al., 2007; Wallace et al., 1993). Subsequently, in 2009, Nash et al. (2009) described a new *erm* gene in *M. abscessus*, designated *erm*(41), which is present in all three taxa of *M. abscessus*. However, the *erm*(41) gene of the former species *M. massiliense* is the shortest in sequence length of any of the previously evaluated *erm* genes as a consequence of two deletions of 2 bp and 274 bp, respectively. The deletions render the gene non-functional such that isolates of *M. massiliense* do not exhibit inducible clarithromycin resistance (Kim et al., 2010). *M. abscessus* subsp. *abscessus* and the former species *M. bolletii* have an intact *erm* gene that confers macrolide resistance.

Inducible macrolide resistance has long remained unrecognized using MIC determination with reading at 3 days as recommended by NCCLS (2003). The observation that MICs of most strains increase with additional incubation (Kim et al., 2010) led to move the reading to 14 days (CLSI, 2013).

Current knowledge distinguishes multiple mechanisms and phenotypes of macrolide resistance which are differently presented by the strains belonging to the *M. abscessus* complex. Inducible resistance is presented by *M. abscessus* subsp. *abscessus* and the previously recognized species *M. bolletii* having a wild-type *erm*(41) gene but is absent when the latter gene has a T28C mutation. In contrast, inducible resistance is not presented by the previously recognized species *M. massiliense*. Inducible resistance is presented by about 80% of *M. abscessus* subsp. *abscessus* and by almost the totality of the strains of the previously recognized species *M. bolletii* and is characterized by low or intermediate MICs.

Constitutive macrolide resistance, characterized by high MICs, may be present in the whole *M. abscessus* complex by consequence of a mutation at either the A2058 or the A2059 position of the 23S rRNA gene (Kim et al., 2010; Bastian et al., 2011; Brown-Elliott et al., 2015). The clinical impact of these MIC differences was corroborated in a recent study assessing the clinical significance in differentiating the former species *M. massiliense* from *M. abscessus* (Koh et al., 2011).

With the aim to clarify the taxonomic status of *M. abscessus* (especially the phylogenetic relationship of the former *M. massiliense* against the other two taxa within *M. abscessus*) we performed WGS of 41 strains cultured from sputum samples of 32 patients, mostly affected by cystic fibrosis. These strains had been identified as members of *M. abscessus* by rpoB gene...
sequencing. In addition, we sequenced the type strains of *M. abscessus* subsp. *bolletii* (CCUG 50184<sup>T</sup>) and of the former species *M. massiliense* (CCUG 48898<sup>T</sup>). The bioinformatic analysis also included the complete sequences, downloaded from GenBank, of *M. abscessus* subsp. *abscessus* ATCC 19977<sup>T</sup> (NC_010397) and of the strains M24 and M115 (NZ_AJLY00000000, NZ_AJLZ00000000) reported previously as possible novel subspecies of *M. abscessus* (Sassi & Dran-court, 2014).

For WGS we used the Illumina MiSeq sequencer, Nextera XT library preparation kits and MiSeq reagent kits as instructed by the manufacturer (Illumina). The resulting reads were ‘quality trimmed’ using Trim Galore and assembled with the SPAdes software (Bankevich et al., 2012). The assembled genomes were assessed for quality with QUAST (Gurevich et al., 2013). The strains and genome information are presented in Table S1 (available in the online Supplementary Material). The annotation of genomes was accomplished with PROKKA (Seemann, 2014). The average nucleotide identity (ANI) and the genome to genome distance (GGD) were calculated among all the strains using the OrthoANI calculator (Lee et al., 2016). The calculation between each genome pair of the 46 strains produced 1035 values ($\binom{46}{2}$) for both GGD and ANI. Based on GGD data, a neighbour-joining tree, rooted using the genome of *Mycobacterium tuberculosis* H37Rv<sup>T</sup> as an outgroup, was reconstructed with MEGA version 6.06 software (Tamura et al., 2013).

The genomes above were mapped to *M. abscessus* ATCC 19977<sup>T</sup> using the exact alignment program SARUMAN (Blom et al., 2011). High-quality single nucleotide polymorphisms (SNPs) were extracted by customized Perl scripts, and the 97 162 SNP positions detected were concatenated into sequence alignments to be used as basis to infer a phylogenetic maximum parsimony tree with the Bionumerics software suite v7.5 (Applied Maths).

DDH is considered the reference for prokaryotic species circumscription. The ANI has been recently validated (Kim et al., 2014) as a computational alternative to DDH, where ANI values higher than 95–96 % are consistent with strains belonging to the same species. In contrast, no cut-off has been proposed yet to define the boundary of subspecies.

The distribution of the 1035 ANI values calculated between each genome pair analysed was typically bimodal and did not show any value between 97.7 and 98.4 (Fig. 1). Once the above-mentioned interval was adopted as cut-off, three clear-cut clusters emerged. Within each cluster, the ANI between single members was at least >98.4 % while the ANI between members of different clusters fell systematically under 96.7 %. The largest group encompassed 26 strains and included the type strain of *M. abscessus* subsp. *abscessus*. A second group consisted of 11 strains and included the type strain of *M. abscessus* subsp. *bolletii*, while nine strains, among which was the type strain of the former species *M. massiliense*, were present in the third group. The ranges of ANI values are reported in Table 1.

**Table 1.** Range of ANI values for different subspecies pairs

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>M. abscessus</em> subsp. <em>abscessus</em></td>
<td>98.8–100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. <em>M. abscessus</em> subsp. <em>bolletii</em></td>
<td>96.9–97.6</td>
<td>98.4–99.9</td>
<td></td>
</tr>
<tr>
<td>3. <em>M. abscessus</em> subsp. <em>massiliense</em></td>
<td>96.8–97.6</td>
<td>96.6–97.2</td>
<td>98.8–100</td>
</tr>
</tbody>
</table>
Fig. 2. Neighbour-joining tree inferred from the distance matrix of GGDs. *M. tuberculosis* H37RvT was used as an outgroup. Bar, 0.02 difference in GGD value.
GGD is the *in silico* equivalent of *in vitro* DDH; here, two strains characterized by values <0.0258 belong to the same species (Meier-Kolthoff et al., 2013). Again, the distribution of GGD values produced three groupings coinciding completely with the values that emerged from ANI data (data not shown). As expected, the GGDs between the strains within each single cluster fell firmly under the cut-off (<0.0170); the ranges of GGD values are reported in Table 2. The inclusion in the analysis (Fig. 2) of species (*M. chelonae, M. franklinii, M. immunogenum* and *M. saopaulense*) belonging to the *M. chelonae/M. abscessus* complex allows highlighting of the contrast between the closeness of the test strains and the clear-cut separation of independent, though closely related, species.

The SNP analysis (Fig. 3) confirmed three clusters embracing the same strains as the two previously described approaches, with an intra-cluster variability of <23 000 SNPs and an inter-cluster distance of >25 000 SNPs.

Out of five patients who yielded multiple isolates, three were shown to be infected by independent strains (4, 3 and 3 respectively). The remaining two subjects both had two isolates each (FRO13247, FRO14056 and MN6, MN352), that were concordant and recognized by the three methods as belonging to a single strain per patient.

Interestingly, two of the strains (BERN12216 and NOV11006) which fell in the *M. massiliense* group exhibited the *rpoB* gene sequence specific for *M. abscessus* subs. *abscessus* but a truncation in the *erm*(41) similar to that of *M. massiliense*.

Four of the strains clustering with *M. massiliense* and two belonging to the group that includes *M. abscessus* subs. *bolletii* presented borderline GGD values, suggesting the possibility of belonging to species different from *M. abscessus* (Table S1). This assignment was, however, not confirmed by ANI and was in clear contrast with the close relatedness characterizing the members of each of the three clusters that emerged from the phylogenetic trees based on GGD and SNP distance matrices (Figs 2 and 3).

Therefore, the analysis of 44 *M. abscessus* genomes (including the 3 type strains, but only two of the two couples of clonal isolates) clearly delineates three taxonomic entities whose most judicious delimitation is at the subspecies level. The two strains (M24 and M115) hypothesized to represent novel *M. abscessus* subspecies, clustered with *M. abscessus* subs. *bolletii* and *M. abscessus* subs. *massiliense*, respectively, such excluding hitherto further subspecies.

In accordance with our findings, Tettelin et al. (2014), based on the WGSs of strains related to the previously recognized species *M. massiliense*, showed the clear distance separating this group from the reference strains of both *M. abscessus* subs. *abscessus* and *M. abscessus* subs. *bolletii*. This study corroborates the previous independent study by Bryant et al. (2013) in the UK, who performed phylogenetic analysis based on whole genome sequence data of the *M. abscessus* group on isolates from clustered outbreaks. The investigators identified deep genetic divisions which corresponded to the presence of three subspecies with high overall genetic similarity (Bryant et al., 2013). They not only noted tight clustering from each patient indicating near-identical genomic sequences but also showed many examples of large genetic differences among isolates from different patients with *M. abscessus* and the former *M. massiliense*.

A recent study identified core genome SNPs in isolates of *M. abscessus* derived from the USA as well as other isolates from outside the USA, including the UK, France, Brazil, Malaysia and China. Their analysis again revealed three genetic subgroups of *M. abscessus* consistent with proposed subspecies classifications (Davidson et al., 2014).

Maintaining the proposed definition at the subspecies level rather than the species level is supported by the fact that *rpoB* and *hsp65* genes of the former *M. massiliense* and the former *M. bolletii* show higher similarities than the proposed 3 % species cut-offs (Adékambi et al., 2006a; McNabb et al., 2004). Specifically, the type strains for the former *M. bolletii* and former *M. massiliense* are only 1.46 % divergent by rpoB and only up to 2.13 % divergent when considering a larger collection of strains (Mačeras et al., 2011). In a previous investigation (Leao et al., 2009) the percentage of DDH among of the type strains of the three proposed subspecies was clearly >70 %, the universally recognized species threshold.

Several studies that have sought to define species in the context of WGS have shown that values of 95–96 % for both average amino acid identity (AAI) and ANI between the genomes of different species correlate well with the classic species definition of a 70 % DNA–DNA re-association threshold (Kim et al., 2014; Konstantinidis & Tiedje, 2005).

Inter-group AAI values for 14 published genomes consisting of three subsets of the *M. abscessus* complex range from 94.2 (between the only strains of former *M. bolletii* and *M. massiliense*) to 96.7 (between *M. bolletii* and *M. abscessus*). However, AAI values among strains within each subgroup are slightly higher; 96.2 to 97.6 (within *M. massiliense*), 97.3 to 98.6 (within *M. abscessus*) and 96.5 (M. bolletii) (Sassi &

---

**Table 2.** Range of GGD values for different subspecies pairs

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>M. abscessus</em> subs. <em>abscessus</em></td>
<td>0.0001–0.0107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. <em>M. abscessus</em> subs. <em>bolletii</em></td>
<td>0.0248–0.0319</td>
<td>0.0009–0.0169</td>
<td></td>
</tr>
<tr>
<td>3. <em>M. abscessus</em> subs. <em>massiliense</em></td>
<td>0.0237–0.0308</td>
<td>0.0290–0.0339</td>
<td>0.0001–0.0100</td>
</tr>
</tbody>
</table>
Fig. 3. Maximum-parsimony tree built from 97 162 SNP positions; logarithmic scale. White circles, *M. abscessus* subsp. *abscessus* strains; grey circles, *M. abscessus* subsp. *massilense* strains; black circles, *M. abscessus* subsp. *bolletii* strains.
Drancourt, 2014). These AAI values are consistent with the ANI values determined in this study, differing only marginally.

There is no question that there are three clades within the *M. abscessus* group. The proposal to maintain these within a single species rather than as three species is consistent with taxo-genomic findings to date, even though these are not well defined for the subspecies level. Probably, the best criterion that defines a subspecies was stated by Wayne et al. (1987): 'Subspecies designations can be used for genetically close organisms that diverge in phenotype' (Wayne et al., 1987). In this regard, the differences in macrolide susceptibility profiles expressed by the members of the *M. abscessus* complex is a phenotypic feature of notable clinical utility.

In summary, we propose that *M. abscessus* be divided into three separate subspecies. These are *M. abscessus* subsp. *abscessus* (for isolates with identity to ATCC 19977); *M. abscessus* subsp. *massiliense* (for isolates with identity to CIP 108297 and CCUG 48898); and *M. abscessus* subsp. *bolletii* (for isolates matching CIP 108541 or CCUG 50184).

**Emended description of Mycobacterium abscessus** (Moore & Frerichs 1953)

*Kusunoki and Ezaki 1992 Leao et al. 2011*


The following properties are displayed in addition to those listed in the species description in 1992 (Kusunoki & Ezaki, 1992) and emended in 2011 (Leao et al., 2011). *M. abscessus* has a genome size of approximately 5,000,000 bp and includes an average of 5,000 genes; the average DNA G+C content is 64 mol% (Table S1). In contrast to other closely related mycobacteria, *M. abscessus* is intrinsically resistant to most antimycobacterials; furthermore, it possesses the *erm*(41) gene. As previously discussed, this gene, when intact, encodes for a ribosomal methyltransferase which is responsible for inducible macrolide resistance (Nash et al., 2009). Emended from the previous taxon description, *Mycobacterium abscessus* encompasses three subspecies: *Mycobacterium abscessus* subsp. *abscessus* (Moore & Frerichs, 1953; Kusunoki & Ezaki, 1992; Leao et al., 2011), *Mycobacterium abscessus* subsp. *bolletii* (Adékambi et al., 2006a; Leao et al., 2011) and *Mycobacterium abscessus* subsp. *massiliense* subsp. nov.

The type strain is Hauduroy L948 (=TMC 1543 = ATCC 19977 = CCUG 20993 = CIP 104536 = DSM 44196 = JCM 13569 = NCTC 13031).

**Emended description of Mycobacterium abscessus subsp. bolletii** (Adékambi et al. 2006) Leao et al. 2011


The description is based on that given for *Mycobacterium bolletii* (Adékambi et al., 2006a). In contrast to earlier emendation (Leao et al., 2011), this subspecies does not include strains of the previously described species *M. massiliense* (Adékambi et al., 2004). The antimicrobial susceptibility is similar to that of *M. abscessus* subsp. *abscessus*, including inducible macrolide resistance, among all wild-type *erm*(41) genes sequenced to date.

The type strain is BD (= CCUG 50184 = CIP 108541 = JCM 15297).

**Description of Mycobacterium abscessus subsp. massiliense** comb. nov.


The description is based on that given for *Mycobacterium massiliense* (Adékambi et al., 2004). This subspecies, featuring a truncated *erm*(41) gene, is devoid of inducible macrolide resistance since the partial remaining gene is non-functional.

The type strain is CCUG 48898 (= CIP 108297 = KCTC 19086 = DSM 45103).

**Acknowledgements**

This study was supported by research grants FFC #27/2014 from Fondazione Ricerca Fibrosi Cistica, the European Union PathoNgenTrace project (FP7- 278864-2) and the German Center for Infection Research (DZIF).

**References**

'Mycobacterium massiliense' sp. nov. from the sputum of a patient with hemoptoic pneumonia. J Clin Microbiol 42, 5493–5501.


