

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. Phillely,⁷ Rossella Balzan,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

Correspondence

Enrico Tortoli
e.tortoli@libero.it
or
tortoli.enrico@hsr.it

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, IRCCS 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.

identity, enome to enome instance and single nucleotide polymorphism analysis. The three methods produced overlapping results, each demonstrating three clusters of strains corresponding to the same number of taxonomic entities. The distances were insufficient to warrant distinction at the species level, but met the criteria for differentiation at the subspecies level. Based on prior *erm*(41)-related phenotypic data and current genomic data, we conclude that the species *M. abscessus* encompasses, in adjunct to the presently recognized subspecies *M. abscessus* subsp. *abscessus* and *M. abscessus* subsp. *bolletii*, a third subspecies for which we suggest the name *M. abscessus* subsp. *massiliense* comb. nov. (type strain CCUG 48898^T=CIP 108297^T=DSM 45103^T=KCTC 19086^T).

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.*, 2006a) and *Mycobacterium massiliense* (Adékambi *et al.*, 2004, 2006b) were first described as novel species closely related to *M. abscessus* on the basis of *rpoB* 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

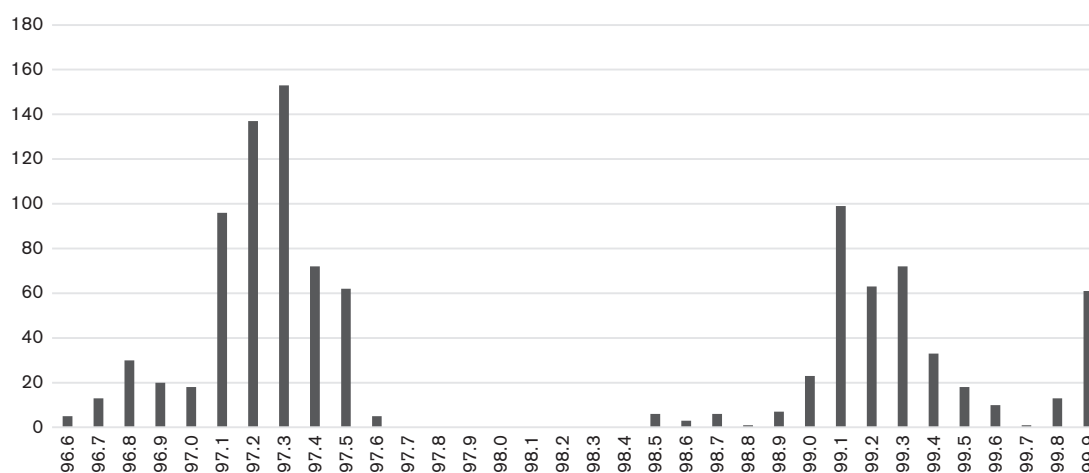


Fig. 1. Distribution of 1035 ANI percentages. The bimodal distribution separates genome–genome comparisons between members of the same subspecies (>98.4) from those between members of different subspecies (<97.7 and >96.5).

sequencing. In addition, we sequenced the type strains of *M. abscessus* subsp. *bolletii* (CCUG 50184^T) and of the former species *M. massiliense* (CCUG 48898^T). The bioinformatic analysis also included the complete sequences, downloaded from GenBank, of *M. abscessus* subsp. *abscessus* ATCC 19977^T (NC_010397) and of the strains M24 and M115 (NZ_AJLY00000000, NZ_AJLZ00000000) reported previously as possible novel subspecies of *M. abscessus* (Sassi & Drancourt, 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 tested 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 (46×45/2) for both GGD and ANI. Based on GGD data, a neighbour-joining tree, rooted using the genome of *Mycobacterium tuberculosis* H37Rv^T 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^T 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

	1	2	3
1. <i>M. abscessus</i> subsp. <i>abscessus</i>	98.8–100		
2. <i>M. abscessus</i> subsp. <i>bolletii</i>	96.9–97.6	98.4–99.9	
3. <i>M. abscessus</i> subsp. <i>massiliense</i>	96.8–97.6	96.6–97.2	98.8–100

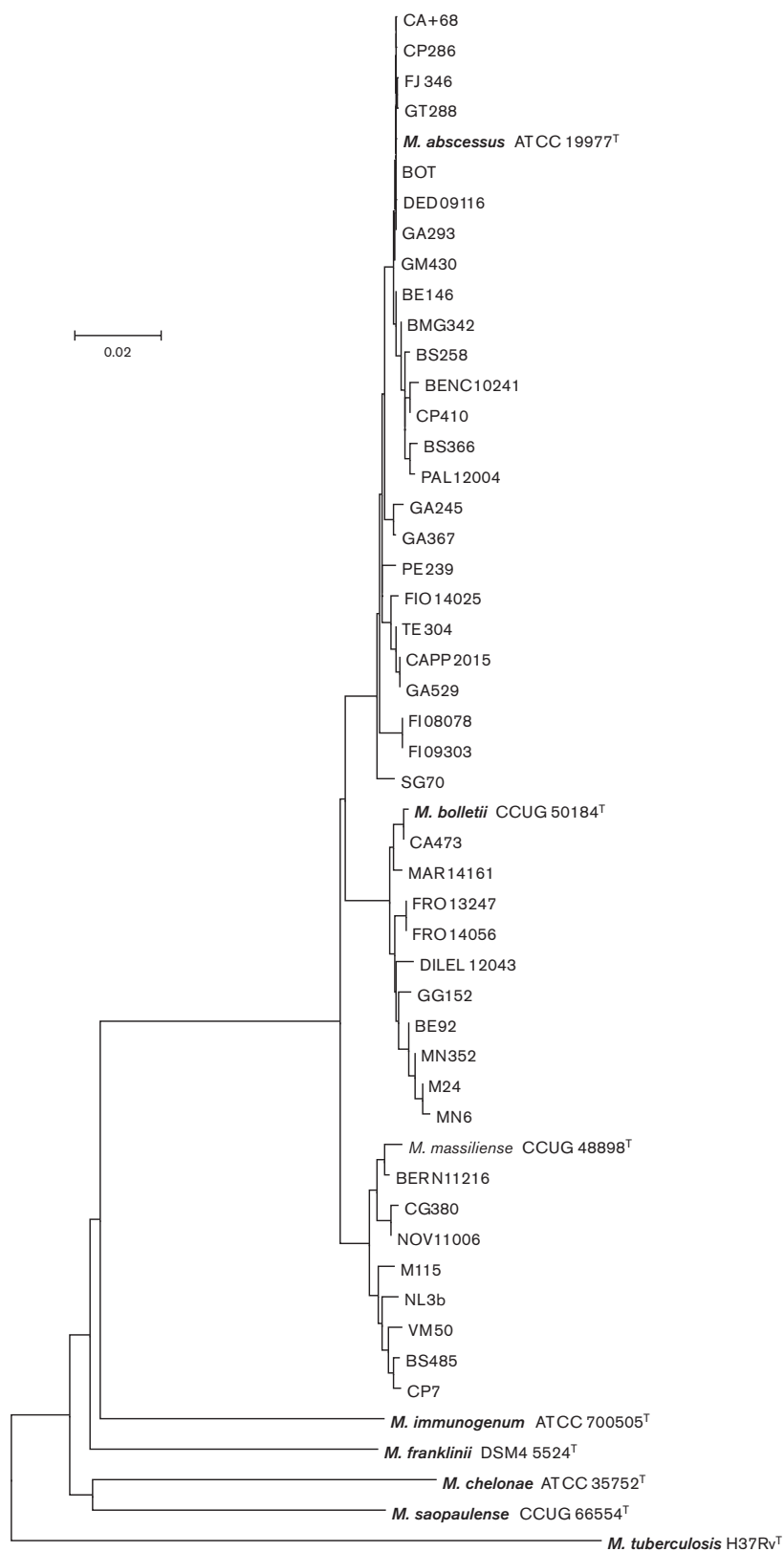


Fig. 2. Neighbour-joining tree inferred from the distance matrix of GGDs. *M. tuberculosis* H37Rv^T was used as an outgroup. Bar, 0.02 difference in GGD value.

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* subsp. *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* subsp. *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* subsp. *bolletii* and *M. abscessus* subsp. *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* subsp. *abscessus* and *M. abscessus* subsp. *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 (Macheras *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

	1	2	3
1. <i>M. abscessus</i> subsp. <i>abscessus</i>	0.0001–0.0107		
2. <i>M. abscessus</i> subsp. <i>bolletii</i>	0.0248–0.0319	0.0009–0.0169	
3. <i>M. abscessus</i> subsp. <i>massiliense</i>	0.0237–0.0308	0.0290–0.0339	0.0001–0.0100

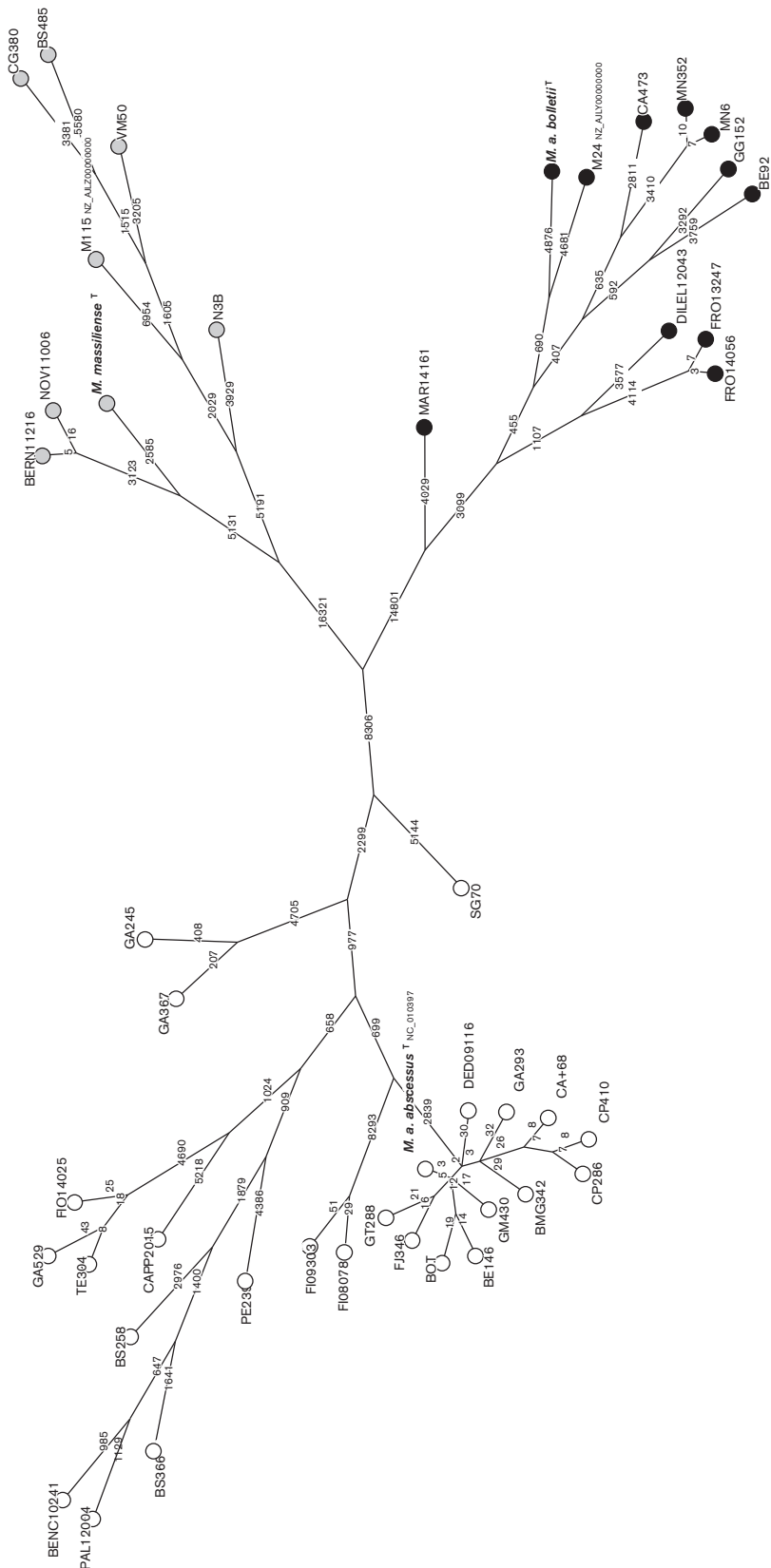


Fig. 3. Maximum-parsimony tree built from 97 162 SNP positions; logarithmic scale. White circles, *M. abscessus* subsp. *abscessus* strains; grey circles, *M. abscessus* subsp. *massiliense* strains; black circles, *M. abscessus* subsp. *bolletii* strains.

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^T); *M. abscessus* subsp. *massiliense* (for isolates with identity to CIP 108297^T and CCUG 48898^T); and *M. abscessus* subsp. *bolletii* (for isolates matching CIP 108541^T or CCUG 50184^T).

Emended description of *Mycobacterium abscessus* (Moore & Frerichs 1953 Kusunoki and Ezaki 1992 Leao *et al.* 2011

Mycobacterium abscessus (abs.ces'sus. L. gen. n. *abscessus* of an abscess).

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 5000 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^T (=TMC 1543^T=ATCC 19977^T=CCUG 20993^T=CIP 104536^T=DSM 44196^T).

Emended description of *Mycobacterium abscessus* subsp. *abscessus* Leao *et al.* 2011

M. abscessus subsp. *abscessus* (abs.ces'sus. L. gen. n. *abscessus* of an abscess).

The description is as those given for *Mycobacterium abscessus* in 1972 (Kubica *et al.*, 1972) and 1992 (Kusunoki & Ezaki, 1992) and for *Mycobacterium abscessus* subsp. *abscessus* in 2011 (Leao *et al.*, 2011). The wild-type *erm*(41) gene confers inducible macrolide resistance which is in contrast absent when this gene presents a T28C substitution.

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

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

M. abscessus subsp. *bolletii* (bol.le'ti.i. N.L. gen. masc. *bolletii* of Bollet, to honour Claude Bollet, clinical microbiologist and taxonomist).

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^T (=CCUG 50184^T=CIP 108541^T=JCM 15297^T).

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

Mycobacterium abscessus subsp. *massiliense* (mas.si.li.en'se L. adj. n. *massiliense*, of the French city of Massilia, now Marseilles, France).

Basonym: *Mycobacterium massiliense*, Adékambi *et al.* 2006.

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^T (=CIP 108297^T=KCTC 19086^T=DSM 45103^T).

Acknowledgements

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

References

Adékambi, T., Reynaud-Gaubert, M., Greub, G., Gevaudan, M. J., La Scola, B., Raoult, D. & Drancourt, M. (2004). Amoebal coculture of

- '*Mycobacterium massiliense*' sp. nov. from the sputum of a patient with hemoptoic pneumonia. *J Clin Microbiol* **42**, 5493–5501.
- Adékambi, T., Berger, P., Raoult, D. & Drancourt, M. (2006a).** *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* **56**, 133–143.
- Adékambi, T., Reynaud-Gaubert, M., Greub, G., Gevaudan, M. J., La Scola, B., Raoult, D. & Drancourt, M. (2006b).** *Mycobacterium massiliense* sp. nov. In: List of new names and new combinations previously effectively, but not validly, published. List no. 111. *Int J Syst Evol Microbiol* **56**, 2025–2027.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S. & other authors (2012).** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455–477.
- Bastian, S., Veziris, N., Roux, A. L., Brossier, F., Gaillard, J. L., Jarlier, V. & Cambau, E. (2011).** Assessment of clarithromycin susceptibility in strains belonging to the *Mycobacterium abscessus* group by *erm(41)* and *rfl* sequencing. *Antimicrob Agents Chemother* **55**, 775–781.
- Blom, J., Jakobi, T., Doppmeier, D., Jaenicke, S., Kalinowski, J., Stoye, J. & Goesmann, A. (2011).** Exact and complete short-read alignment to microbial genomes using graphics processing unit programming. *Bioinformatics* **27**, 1351–1358.
- Brown-Elliott, B. A., Vasireddy, S., Vasireddy, R., Iakhiaeva, E., Howard, S. T., Nash, K., Parodi, N., Strong, A., Gee, M. & other authors (2015).** Utility of sequencing the *erm(41)* gene in isolates of *Mycobacterium abscessus* subsp. *abscessus* with low and intermediate clarithromycin MICs. *J Clin Microbiol* **53**, 1211–1215.
- Bryant, J. M., Grogono, D. M., Greaves, D., Foweraker, J., Roddick, I., Inns, T., Reacher, M., Haworth, C. S., Curran, M. D. & other authors (2013).** Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. *Lancet* **381**, 1551–1560.
- Cho, Y. J., Yi, H., Chun, J., Cho, S. N., Daley, C. L., Koh, W. J. & Shin, S. J. (2013).** The genome sequence of '*Mycobacterium massiliense*' strain CIP 108297 suggests the independent taxonomic status of the *Mycobacterium abscessus* complex at the subspecies level. *PLoS One* **8**, e81560.
- CLSI (2013).** Susceptibility Testing of Mycobacteria, Nocardiae, and Other Aerobic Actinomycetes; Approved Standard—Second Edition. M24–A2. Wayne, PA: N.C.C.L.S.
- Davidson, R. M., Hasan, N. A., Reynolds, P. R., Totten, S., Garcia, B., Levin, A., Ramamoorthy, P., Heifets, L., Daley, C. L. & other authors (2014).** Genome sequencing of *Mycobacterium abscessus* isolates from patients in the united states and comparisons to globally diverse clinical strains. *J Clin Microbiol* **52**, 3573–3582.
- Griffith, D. E., Aksamit, T., Brown-Elliott, B. A., Catanzaro, A., Daley, C., Gordin, F., Holland, S. M., Horsburgh, R., Huit, G. & other authors (2007).** An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am J Respir Crit Care Med* **175**, 367–416.
- Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. (2013).** QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075.
- Kim, H. Y., Kim, B. J., Kook, Y., Yun, Y. J., Shin, J. H., Kim, B. J. & Kook, Y. H. (2010).** *Mycobacterium massiliense* is differentiated from *Mycobacterium abscessus* and *Mycobacterium bolletii* by erythromycin ribosome methyltransferase gene (*erm*) and clarithromycin susceptibility patterns. *Microbiol Immunol* **54**, 347–353.
- Kim, M., Oh, H. S., Park, S. C. & Chun, J. (2014).** Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* **64**, 346–351.
- Koh, W. J., Jeon, K., Lee, N. Y., Kim, B. J., Kook, Y. H., Lee, S. H., Park, Y. K., Kim, C. K., Shin, S. J. & other authors (2011).** Clinical significance of differentiation of *Mycobacterium massiliense* from *Mycobacterium abscessus*. *Am J Respir Crit Care Med* **183**, 405–410.
- Konstantinidis, K. T. & Tiedje, J. M. (2005).** Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA* **102**, 2567–2572.
- Kubica, G. P., Baess, I., Gordon, R. E., Jenkins, P. A., Kwapinski, J. B., McDermont, C., Pattyn, S. R., Saito, H., Silcox, V. & other authors (1972).** A co-operative numerical analysis of rapidly growing mycobacteria. *J Gen Microbiol* **73**, 55–70.
- Kusunoki, S. & Ezaki, T. (1992).** 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 Bacteriol* **42**, 240–245.
- Leao, S. C., Tortoli, E., Viana-Niero, C., Ueki, S. Y., Lima, K. V., Lopes, M. L., Yubero, J., Menendez, M. C. & Garcia, M. J. (2009).** Characterization of mycobacteria from a major Brazilian outbreak suggests that revision of the taxonomic status of members of the *Mycobacterium chelonae*-*M. abscessus* group is needed. *J Clin Microbiol* **47**, 2691–2698.
- Leao, S. C., Tortoli, E., Euzéby, J. P. & Garcia, M. J. (2011).** Proposal that *Mycobacterium massiliense* and *Mycobacterium bolletii* be united and reclassified as *Mycobacterium abscessus* subsp. *bolletii* comb. nov., designation of *Mycobacterium abscessus* subsp. *abscessus* subsp. nov. and emended description of *Mycobacterium abscessus*. *Int J Syst Evol Microbiol* **61**, 2311–2313.
- Lee, I., Kim, Y. O., Park, S. C. & Chun, J. (2016).** OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* (Epub ahead of print).
- Macheras, E., Roux, A. L., Bastian, S., Leão, S. C., Palaci, M., Sivadon-Tardy, V., Gutierrez, C., Richter, E., Rüscher-Gerdes, S. & other authors (2011).** Multilocus sequence analysis and *rpoB* sequencing of *Mycobacterium abscessus* (*sensu lato*) strains. *J Clin Microbiol* **49**, 491–499.
- McNabb, A., Eisler, D., Adie, K., Amos, M., Rodrigues, M., Stephens, G., Black, W. A. & Isaac-Renton, J. (2004).** Assessment of partial sequencing of the 65-kilodalton heat shock protein gene (*hsp65*) for routine identification of *Mycobacterium* species isolated from clinical sources. *J Clin Microbiol* **42**, 3000–3011.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P. & Göker, M. (2013).** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 60.
- Moore, M. & Frerichs, J. B. (1953).** An unusual acid-fast infection of the knee with subcutaneous, abscess-like lesions of the gluteal region. *J Invest Dermatol* **20**, 133–169.
- Nash, K. A., Andini, N., Zhang, Y., Brown-Elliott, B. A. & Wallace, R. J. (2006).** Intrinsic macrolide resistance in rapidly growing mycobacteria. *Antimicrob Agents Chemother* **50**, 3476–3478.
- Nash, K. A., Brown-Elliott, B. A. & Wallace, R. J. (2009).** A novel gene, *erm(41)*, confers inducible macrolide resistance to clinical isolates of *Mycobacterium abscessus* but is absent from *Mycobacterium chelonae*. *Antimicrob Agents Chemother* **53**, 1367–1376.
- NCCLS (2003).** *Susceptibility Testing for Mycobacteria, Nocardiae and Other Aerobic Actinomycetes*; Approved Standard M24-A. Wayne, PA: NCCLS.
- Sassi, M. & Drancourt, M. (2014).** Genome analysis reveals three genomospecies in *Mycobacterium abscessus*. *BMC Genomics* **15**, 359.
- Seemann, T. (2014).** PROKKA: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069.

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

Tan, J. L., Ngeow, Y. F. & Choo, S. W. (2015). Support from phylogenetic networks and subspecies signatures for separation of *Mycobacterium massiliense* from *Mycobacterium bolletii*. *J Clin Microbiol* **53**, 3042–3046.

Tettelin, H., Davidson, R. M., Agrawal, S., Aitken, M. L., Shallom, S., Hasan, N. A., Strong, M., de Moura, V. C., De Groot, M. A. & other authors (2014). High-level relatedness among *Mycobacterium abscessus*

subsp. *massiliense* strains from widely separated outbreaks. *Emerg Infect Dis* **20**, 364–371.

Wallace, R. J., Tanner, D., Brennan, P. J. & Brown, B. A. (1993). Clinical trial of clarithromycin for cutaneous (disseminated) infection due to *Mycobacterium chelonae*. *Ann Intern Med* **119**, 482–486.

Wayne, L. G., Stackebrandt, E., Kandler, O., Colwell, R. R., Krichevsky, M. I., Truper, H. G., Murray, R. G. E., Moore, W. E. C., Grimont, P. A. D. & other authors (1987). Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Evol Microbiol* **37**, 463–464.

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

Enrico Tortoli*, Thomas A. Kohl, Barbara A. Brown-Elliott, Alberto Trovato, Sylvia Cardoso Leão, Maria Jesus Garcia, Sruthi Vasireddy, Christine Y.

Turenne, David E. Griffith, Julie V. Philley, Rossella Baldan, Silvia Campana, Lisa Cariani, Carla Colombo, Giovanni Taccetti, Antonio Teri, Stefan

Niemann, Richard J. Wallace Jr., Daniela M. Cirillo

International Journal of Evolutionary and Systematic Microbiology

*Emerging Bacterial Pathogens Unit, San Raffaele Scientific Institute, Milan, Italy

e-mail: tortoli.enrico@hsr.it

TABLE S1. List of investigated *Mycobacterium abscessus* genomes.

Strain	Length (bp)	Contigs n.	GC %	N50 ^a	CDS n.	tRNA n.	tmRNA n.	rRNA n.	Subspecies	Source ^b	Accession n.
ATCC19977 ^T	5 067 172	1	64.13	5 067 172	4 955	49	1	3	abscessus	RS	NC_010397
BE92	4 791 487	34	64.18	330 599	4 645	50	1	3	bolletii	CF	ERS1157732
BE146	5 112 731	61	64.16	261 310	4 999	50	1	3	abscessus	CF	ERS1157731
BENC10241	5 115 825	61	64.26	195 248	5 039	48	1	3	abscessus	CF	ERS1157733
BERN11216 ^d	4 951 679	25	64.19	505 545	4 853	53	1	3	massiliense	CF	ERS1157734
BMG342	5 233.614	111	64.09	115 485	5 172	49	1	3	abscessus	CF	ERS1157708
BOT	5 074781	31	64.13	403 547	4 957	50	1	3	abscessus	CF	ERS1157735
BS258	4 974 102	55	64.21	221 567	4 859	48	1	3	abscessus	CF	ERS1157736
BS366	5 265 274	70	64.28	250 177	5 246	49	1	3	abscessus	CF	ERS1157737
BS485 ^d	5 022 204	38	64.11	385 152	4 964	48	1	3	massiliense	CF	ERS1157738
CA+68	5 122 709	57	64.22	229 039	5 015	49	1	3	abscessus	CF	ERS1157740
CA473	4 819 320	30	64.19	681 474	4 654	49	1	3	bolletii	CF	ERS1157739
CAPP2015	5 143 680	48	64.09	222 142	5 052	49	1	3	abscessus	CF	ERS1157741

CCUG48898 ^T	4 967 828	25	64.09	559 054	4 894	50	1	3	massiliense	RS	ERS1157727
CCUG50184 ^T	5 062 917	31	64.07	282 149	4 926	51	1	3	bolletii	RS	ERS1157726
CG380	4 830 164	37	64.31	210 243	4 796	47	1	3	massiliense	CF	ERS1157742
CP286	5 102 024	50	64,22	256 816	4 992	49	1	3	abscessus	CF	ERS1157743
CP410	5 103 166	57	64.22	228 997	4 985	49	1	3	abscessus	CF	ERS1157744
DED09116	5 304 297	45	63.99	296 827	5 194	49	1	3	abscessus	CF	ERS1157745
DILEL12043	5 005 184	25	64.05	396 471	4 861	50	1	3	bolletii	CF	ERS1157746
FI08078	4 793 119	68	64.08	223 390	4 652	50	1	3	abscessus	NCF	ERS1157747
FI09303	4 795 482	76	64.07	217 850	4 647	50	1	3	abscessus	NCF	ERS1157748
FIO14025	5 201 487	60	64.16	280 511	5 116	50	1	3	abscessus	CF	ERS1157749
FJ346	5 215 747	33	64.05	410 661	5 165	53	1	3	abscessus	CF	ERS1157709
FRO13247 ^b	4 772 749	18	64,24	553 115	4 609	50	1	3	bolletii	CF	ERS1157710
FRO14056 ^b	4 771 978	18	64.25	661 484	4 604	50	1	3	bolletii	CF	ERS1157711
GA245	4 905 113	34	64,27	273 060	4 747	50	1	3	abscessus	CF	ERS1157712
GA293	5 224 799	44	64.21	303 050	5 156	49	1	3	abscessus	CF	ERS1157713
GA367	5 827 082	1193	64.14	34 114	5 901	63	1	3	abscessus	CF	ERS1157714
GA529	5 447 900	102	64.15	124 497	5 456	50	1	3	abscessus	CF	ERS1157715
GG152	5 223 108	32	64,04	391 509	5 205	51	1	3	bolletii	CF	ERS1157716
GM430	5 123 747	41	64.12	261 307	5 019	49	1	3	abscessus	CF	ERS1157717
GT288	5 191 674	50	64.14	207 644	5 111	49	1	3	abscessus	CF	ERS1157718
M24	5 488 620	67	63.88	150 543	5 553	82	1	3	bolletii	NCF	NZ_AJLY00000000
M115 ^d	4 975 940	47	64.14	307 143	4 920	48	1	3	massiliense	NCF	NZ_AJLZ00000000
MAR14161 ^d	4 893 173	32	64,26	441 810	4 758	51	1	3	bolletii	CF	ERS1157719
MN6 ^{cd}	4 974 931	52	64.11	239 317	4 861	51	1	3	bolletii	CF	ERS1157721
MN352 ^{cd}	4 971 525	13	64.14	1 056 724	4 849	51	1	3	bolletii	CF	ERS1157720
N3B	5 043 958	68	64.28	262 846	4 977	51	1	3	massiliense	CF	ERS1157722
NOV11006 ^d	4 905 971	38	64,21	351 381	4 792	53	1	3	massiliense	CF	ERS1157723
PAL12004	5 467 637	80	64.01	218 469	5 510	102	1	3	abscessus	CF	ERS1157724
PE239	5 159 042	53	63.97	253 063	5 153	67	1	3	abscessus	CF	ERS1157725
SG70	4 998 036	17	64,15	355 913	4 949	49	1	3	abscessus	CF	ERS1157728
TE304	5 198 631	75	64,18	153 015	5 131	49	1	3	abscessus	CF	ERS1157729
VM50	5 061 787	152	64.28	112 875	4 973	49	1	4	massiliense	CF	ERS1157730

^a N50, 50% average contigs length; CF, cystic fibrosis patient; NCF, non cystic fibrosis patient; RS, reference strain;

^{b,c} isolates belonging to the same strain;

^d genomes presenting borderline GGD values.