

RESEARCH ARTICLE

Open Access

Genome analysis reveals three genomospecies in *Mycobacterium abscessus*

Mohamed Sassi¹ and Michel Drancourt^{1,2*}

Abstract

Background: *Mycobacterium abscessus* complex, the third most frequent mycobacterial complex responsible for community- and health care-associated infections in developed countries, comprises of *M. abscessus* subsp. *abscessus* and *M. abscessus* subsp. *bolletii* previously referred as *Mycobacterium bolletii* and *Mycobacterium massiliense*. The diversity of this group of opportunistic pathogens is poorly described.

Results: In-depth analysis of 14 published *M. abscessus* complex genomes found a pan-genome of 6,153 proteins and core-genome of 3,947 (64.1%) proteins, indicating a non-conservative genome. Analysing the average percentage of amino-acid sequence identity (from 94.19% to 98.58%) discriminates three main clusters C1, C2 and C3: C1 comprises strains belonging to *M. abscessus*, C2 comprises strains belonging to *M. massiliense* and C3 comprises strains belonging to *M. bolletii*; and two sub-clusters in clusters C2 and C3. The phylogenomic network confirms these three clusters. The genome length (from 4.8 to 5.51-Mb) varies from 5.07-Mb in C1, 4.89-Mb in C2A, 5.01-Mb in C2B and 5.28-Mb in C3. The mean number of prophage regions (from 0 to 7) is 2 in C1; 1.33 in C2A; 3.5 in C2B and five in C3. A total of 36 genes are uniquely present in C1, 15 in C2 and 15 in C3. These genes could be used for the detection and identification of organisms in each cluster. Further, the mean number of host-interaction factors (including PE, PPE, LpqH, MCE, Yrbe and type VII secretion system ESX3 and ESX4) varies from 70 in cluster C1, 80 in cluster C2A, 74 in cluster C2B and 93 in clusters C3A and C3B. No significant differences in antibiotic resistance genes were observed between clusters, in contrast to previously reported in-vitro patterns of drug resistance. They encode both penicillin-binding proteins targeted by β -lactam antibiotics and an Ambler class A β -lactamase for which inhibitors exist.

Conclusions: Our comparative analysis indicates that *M. abscessus* complex comprises three genomospecies, corresponding to *M. abscessus*, *M. bolletii*, and *M. massiliense*. The genomics data here reported indicate differences in virulence of medical interest; and suggest targets for the refined detection and identification of *M. abscessus*.

Keywords: *Mycobacterium abscessus*, *Mycobacterium bolletii*, *Mycobacterium massiliense*, *Araucaria*, *Mycobacteriophage*, Prophage

Background

The non-tuberculous mycobacterium *Mycobacterium abscessus* was long confused with *Mycobacterium chelonae* [1]. Other closely related species include *Mycobacterium salmoniphilum* [2], *Mycobacterium immunogenum* [3], *Mycobacterium massiliense* [4], *Mycobacterium bolletii* [5] and *Mycobacterium franklinii* [6] altogether forming the *Mycobacterium chelonae-abscessus* complex. This complex is the third most frequent mycobacterial complex

infecting humans in developed countries besides the *Mycobacterium tuberculosis* and *Mycobacterium avium* complexes [7,8]. Bibliometrics retrieving over 1,700 publications in the Medline database illustrates the fact that this complex is emerging, causing both sporadic cases and outbreaks of community-acquired and health-care associated infections [9]. Not only humans but also cats [10,11] and dolphins [12-14] are infected while fishes are uniquely infected by *M. salmoniphilum* [2,15].

Current nomenclature is that the species *M. abscessus* comprises two subspecies named *M. abscessus* subsp. *abscessus* and *M. abscessus* subsp. *bolletii*. Later taxon accommodates isolates previously referred as *M. bolletii* or *M. massiliense* [16]. This nomenclature however may

* Correspondence: michel.drancourt@univ-amu.fr

¹ Aix Marseille Université, URMITE, UMR63, CNRS 7278, IRD 198, Inserm 1095, Marseille, France

² Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Faculté de Médecine, 27, Boulevard Jean Moulin, Marseille cedex 5, France

obscure the true diversity of mycobacteria in this complex. While the 16S rRNA gene yields an identical sequence for *M. abscessus* and *M. bolletii*, it shares 99% sequence identity with *M. massiliense*. *RpoB* gene sequencing founded the description of recent species [17-19] but yielded further conflicting results [20-22]. Multilocus sequencing analysis [23] and multispacer sequence typing [24] differentiated *M. massiliense* from *M. bolletii*. In this report, the previous nomenclature *M. abscessus*, *M. bolletii* and *M. massiliense* forming the *M. abscessus* complex, has been retained for clarity.

The availability of 39 *M. abscessus*, 13 *M. massiliense* and two *M. bolletii* genomes in the National Center for BioInformatics (NCBI) genome database provides new opportunities to assess the diversity of this species. Here, we review 14 complete published *M. abscessus* complex genomes and compare them with the re-annotated *M. tuberculosis* H37Rv genome (Table 1) in order to in-depth analyse the diversity of *M. abscessus*.

Results and discussion

M. abscessus complex pan- and core-genome

M. abscessus complex genomes comprise one circular chromosome. In addition, *M. abscessus* ATCC 19977 contains one 23-kb plasmid identical to the *Mycobacterium*

marinum pMM23 plasmid, encoding mer operon and mercury reductase protein, which may confer resistance to organo-mercury compounds [25]. In order to normalize the predicted proteins and to minimize the differences of presence/absence of genes and length, coding sequences were predicted using prodigal software [26]. We identified a total of 70,309 protein-coding sequences which number varies from 4,651 to 5,079 in each genome (Table 2). The core-genome contains 57,172 protein sequences accounting for 64.15% of the pan-genome. This figure indicates a non-conservative genome contrary to that of *Mycobacterium tuberculosis*, a conservative-genome pathogen which core-genome accounts for 96.1% of the pan-genome [27]. Using orthoMCL [28] with a conservative parameter value of 50% sequence identity, we categorized these 70,309 proteins into 6,153 orthologous protein groups, including 3,947 core-genome groups and 55 strain-specific groups.

M. abscessus complex diversity

The average percentage of amino-acid sequence identity (AAI) of core proteins was determined as previously described [29]. The AAI values indicate that *M. abscessus* complex forms three main clusters: cluster 1 (C1) includes *M. abscessus* type strain and strains M93, 94, M152 and Go06; cluster 2 (C2) contains two subclusters: cluster 2A

Table 1 List of *Mycobacterium abscessus* genomes here studied

Clusters	Organism	Isolated from	Geography	BioProject
C1	<i>M. abscessus</i> Type strain	human knee infection	United States	PRJNA61613, PRJNA15691
C1	<i>M. abscessus</i> M93	sputum sample from a Malaysian patient presenting with a prolonged productive cough suggestive of a bacterial lower respiratory tract infection	Malaysia	PRJNA180393, PRJNA84203
C1	<i>M. abscessus</i> M94	sputum sample of a Malaysian patient with a persistent cough and fever and consolidation in the chest radiograph	Malaysia	PRJNA180394, PRJNA88149
C1	<i>M. abscessus</i> M152	acid-fast bacillus positive sputum of a Malaysian man	Malaysia	PRJNA159789
C1	<i>M. massiliense</i> strain GO 06	undergone knee joint surgery	Brazil	PRJNA170732, PRJNA168263
C2A	<i>M. massiliense</i> Type strain	sputum specimen from hemoptoic pneumonia	Marseille	PRJNA180742, PRJNA65215
C2A	<i>M. massiliense</i> M18	lymph node biopsy specimen from a Malaysian patient suspected of having tuberculous cervical lymphadenitis	Malaysia	PRJNA89593
C2A	<i>M. massiliense</i> M154	bronchoalveolar lavage fluid of a Malaysian patient presenting with lower respiratory tract infection	Malaysia	PRJNA89603
C2B	<i>M. abscessus</i> 47 J26	sputum sample from a patient with Cystic fibrosis	England	PRJNA179981, PRJNA73255
C2B	<i>M. abscessus</i> M115	sputum from a Malaysian patient presenting with persistent cough and loss of body weight suggestive of pulmonary tuberculosis	Malaysia	PRJNA89601
C2B	<i>M. abscessus</i> M139	sputum sample of a 26-year-old Nepalese male presenting with hemoptysis	Nepal	PRJNA159701
C2B	<i>M. abscessus</i> M172	sputum isolate from a Malaysian patient	Malaysia	PRJNA89599
C3A	<i>M. bolletii</i> Type strain	respiratory tract specimen collected in woman with hemoptoic pneumonia	Marseille	PRJNA180015, PRJNA73695
C3B	<i>M. abscessus</i> M24	the bronchoalveolar lavage fluid of a Malaysian patient	Malaysia	PRJNA89595

Table 2 *Mycobacterium abscessus* core genome and unique genes

Clusters	Organism	CDS	Unique core genome	Total genes
C1	<i>M. abscessus</i>	-	36	180
	<i>M. abscessus</i> T	4954	-	-
	strain GO 06	4944	-	-
	M93	4733	11	11
	M94	4841	10	10
C2A	<i>M. massiliense</i>	-	15	107
	<i>M. massiliense</i> T	4962	3	3
	M18	4663	8	8
	M154	4651	-	-
	M172	5079	20	20
C2B	47 J26	4766	-	-
	M115	4802	4	4
	M139	4754	4	4
	M172	5079	20	20
	M152	4762	-	-
C3A	<i>M. bolletii</i>	-	15	30
	<i>M. bolletii</i> T	4733	9	9
C3B	M24	4960	23	23
<i>M. abscessus</i> core genome		-	3,947	57,172

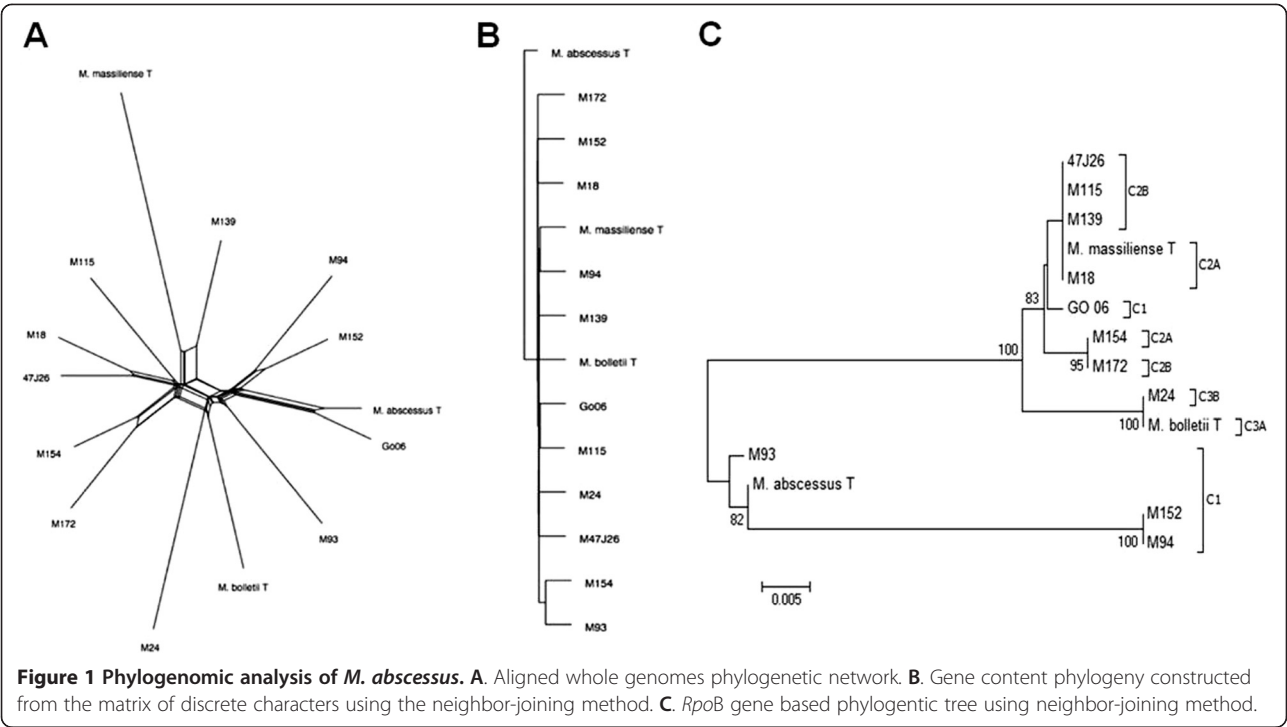
(C2A) includes *M. massiliense* type strain and strains M154 and M18; cluster 2B (C2B) includes strains 47 J26, M115, M172 and M139; cluster 3 (C3) includes two sub-clusters: cluster 3A (C3A) includes *M. bolletii* type strain and cluster 3B (C3B) includes *M. bolletii* strain M24 (Table 3).

M. abscessus complex proteomes were further aligned using Mauve software [30] to infer phylogeny using the Neighbor-Net algorithm in the package SplitsTree4 [31]. The phylogenomic network confirms the three clusters C1, C2 and C3 (Figure 1A). A phylogenomic tree based on gene content (i.e., the presence or absence of orthologs) (Figure 1B) organizes *M. abscessus* differently from the whole genome concatenated tree (Figure 1A) or even the phylogenetic tree based on *rpoB* gene sequence (Figure 1C). Phylogenomic analysis indicates that the *M. abscessus* gene repertoires have different evolutionary histories and suggests that differential gene loss and lateral gene acquisition are playing important roles in the evolution of some *M. abscessus* strains. Notably, the situation of strain Go06 is confusing, as it presents 98.4% AAI with *M. abscessus* type strain in C1 (Figure 1A) whereas its *rpoB* gene sequence and single nucleotide polymorphisms analysis are indicative of *M. massiliense* [8,32]. Our analyses indicate that strain Go06 have an ambiguous classification as a chimera between *M. abscessus* and *M. massiliense* and is the only example compatible with a lateral transfer of *rpoB* gene.

Functional analysis using Clusters of Orthologous Groups database (COG) [33] and BLASTP indicates unique genes in C1, encoding hypothetical proteins, proteins implicated in transcription, energy production and transport, carbohydrate metabolism and transport, lipid metabolism, nucleotide metabolism and transport, amino-acid metabolism and transport, post-translational modification and inorganic ion transport and metabolism (Table 2, Figure 2). Within C2, unique genes encode hypothetical proteins, proteins implicated in transcription and lipid metabolism. Within C3,

Table 3 Average nucleotide identity and characteristics of *Mycobacterium abscessus* genomes

Clusters	Strains	Genome length Mb	Genome GC%	AAI Vs <i>M. abscessus</i> T	AAI Vs <i>M. bolletii</i> T	AAI Vs <i>massiliense</i> T
<i>M. abscessus</i>	C1	<i>M. abscessus</i> T	5,09	62,7	100,00	94,74
		M93	5,08	64,2	97,30	94,76
		M94	5,1	64,2	97,56	94,79
		M152	4,91	64,1	98,59	95,73
		strain GO 06	5,07	64,2	98,35	95,64
<i>M. massiliense</i>	C2A	<i>M. massiliense</i> T	5,2	64,2	95,56	100,00
		M18	4,89	64,2	96,66	97,57
		M154	4,8	64,1	96,14	97,26
	C2B	M115	4,98	64,1	96,16	96,92
		M172	5,2	64,2	95,30	96,17
		M47 J26	4,87	64,1	96,23	96,93
		M139	5,05	64,1	95,94	96,88
<i>M. bolletii</i>	C3A	<i>M. bolletii</i> T	5,05	64,2	94,51	95,33
	C3B	M24	5,51	64,2	94,91	94,20

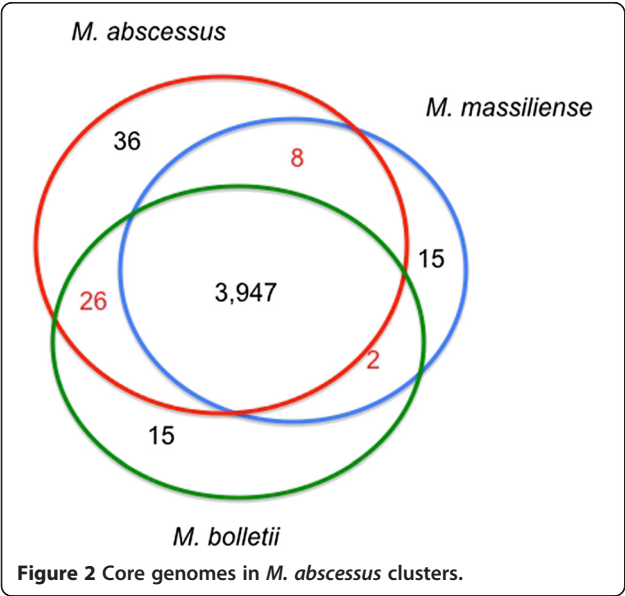


unique genes encode hypothetical proteins, proteins implicated in amino-acid metabolism and transport and translation. These unique genes could base a refined identification of the three genomospecies. However, we could not exclude that these unique genes could be due to a coding sequence, which arose de novo, to HGT or gene loss for the other subspecies. In the case of absence of a gene, this could also be due to a real loss or to an assembly artefact.

C1 strains have been isolated from American and Malaysian patients suffering knee infection and lower

respiratory infection, respectively (Table 1). C2A strains were isolated from Malaysian and French patients suffering severe, respiratory tract infections. C2B strains were isolated from Nepalese, Malaysian and English patients suffering respiratory tract infections, including cystis fibrosis and pulmonary tuberculosis patients. C3A and C3B strains were exclusively isolated from patients suffering respiratory tract infections, in France and Malaysia, respectively. Therefore, clusters specify the clinical form and geographical origin of the infection.

Altogether, genomics analyses revealed a more heterogeneous structure of *M. abscessus* complex than the one currently suggested by the nomenclature, which recognizes only two subspecies within *M. abscessus* [16]. It has been proposed that two genomes exhibiting AAI >96% belong to the same species [34,35]. Therefore, AAI analysis indicates that *M. abscessus* is in fact comprising of three genomospecies, corresponding to previous nomenclature of *M. abscessus* (C1), *M. massiliense* (C2) and *M. bolletii* (C3). Using an AAI <97% threshold would further determine two subspecies in *M. massiliense* (C2A and C2B) and in *M. bolletii* (C3A and C3B). Recent whole genome sequencing analyses of clinical isolates in Great Britain also clearly distinguished three clusters in agreement with the three here reported [8]. All these data support reevaluating the taxonomy of *M. abscessus* complex, to recognize three genomospecies *M. abscessus* (C1), *M. bolletii* (C2), and *M. massiliense* (C3); and four unnamed subspecies C2A, C2B; C3A, C3B.



M. abscessus prophagome

M. abscessus median GC% content is 64.2%, ranging from 62.7% (*M. abscessus* ATCC 19977) to 64.2% (strain Go 06). The GC% is not characteristic of the clusters as the median GC% content of C1, C2A and C3 is 64.2%, close to the median 64.1% GC% content in C2B.

However, there is a significant 14.7% variation in the genome length from 4.8-Mb (*M. abscessus* M154) to 5.51-Mb (*M. abscessus* M24) with a median of 5.07-Mb. The median of genome size is 5.07-Mb in C1, 4.89-Mb in C2A, 5.01-Mb in C2B and 5.28-Mb in C3. Differences in the genome size correlate with the number of prophage regions which are detected in 13/14 *M. abscessus* genomes (Figure 3): *M. abscessus* M154 (*M. massiliense* C2A) has the smallest genome encoding no prophage whereas *M. bolletii* M24 (C3) has the largest genome encoding seven prophage regions (Figure 3). Prophage regions comprise up to 5% of the genome length in *M. abscessus* M172. The number of prophage regions in other genomes is diverse, ranging from one to six regions encoding putative genes in the subsystem of phages, prophages, transposable elements and plasmids, which might contribute to species diversity [36]. The mean number of prophage regions is 2 in C1, 1.33 in C2A, 3.5 in C2B and 5 in C3. This observation confirms the particularity of C3: *M. bolletii* CIP108541^T contains a 13-kb and a 63-kb prophage whereas *M. bolletii* strain M24 contains seven prophage regions including one 17-kb region homologous to the *M. bolletii* CIP108541^T 13-kb region and a 27-kb region homologous to the *M. massiliense* CCUG 48898 50-kb region [37,38] (Table 4). *M. abscessus* genomes encode putative phage-related genes necessary

for phage replication as well as phage-tail protein, phage endolysin, capsid proteins (major protein and scaffold proteins) and phage tape measure protein. Both ends of this region encode putative phage integrases. *M. abscessus* genomes encode small prophage-like regions. However, only *M. bolletii* has been reported to produce a mycobacteriophage that we named Araucaria after we recently resolved its electron microscopy 3D structure [39]. *M. abscessus* M94 genome harbours one particular pseudo-tRNA spanning the region 51,150-57,394 in contig 33, which is not observed in the other *M. abscessus* genomes [40]. Phages have been reported to increase virulence of their host and encode antimicrobial resistance genes [41]. In *M. abscessus* however, no such genes were identified but phages could be targeted for the differentiation between the three *M. abscessus* genomospecies.

M. abscessus complex resistome

As all mycobacteria, *M. abscessus* complex is embedded into a hydrophobic cell wall barrier to hydrophilic antibiotics. Accordingly, *M. abscessus* is multidrug resistant organisms exhibiting different drug resistance [42-44]. *M. abscessus* genomes encode many proteins potentially involved in drug-efflux systems, including members of the major facilitator family, ABC transporters and MmpL proteins; Small Multidrug Resistance-family, a family of lipophilic drug efflux proteins [45]; and a multidrug resistance stp protein similar to *M. tuberculosis* involved in spectinomycin and tetracycline resistance [46]. *M. abscessus*, *M. bolletii* and *M. massiliense* were reported to be *in-vitro* susceptible to amikacin; however, comparison with the *M.*

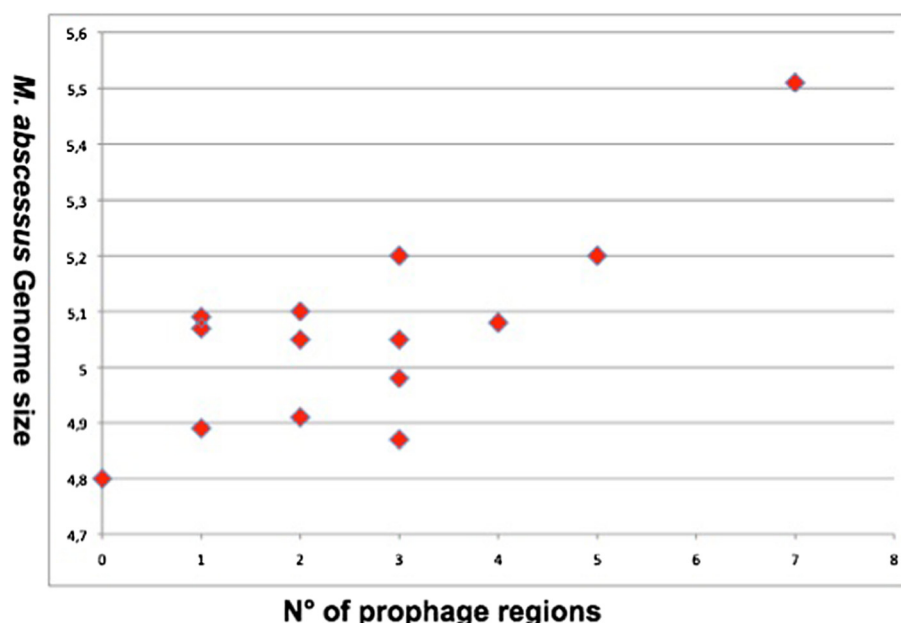


Figure 3 Correlation between *Mycobacterium abscessus* genomes size (y axis) and the number of prophages (x axis).

Table 4 *Mycobacterium abscessus* prophages

Clusters		Strain	Region	Region_length	CDS	Region_position	
<i>M. abscessus</i>	C1	<i>M. abscessus T</i>	1	81 Kb	110	1754551-1835095	
		<i>M93</i>	1	16.4 Kb	33	197463-213867	
			2	38 Kb	51	232006-270072	
			3	53 Kb	70	1762720-1815780	
			4	20.2 Kb	26	1820768-1841058	
		<i>M94</i>	1	58.3 Kb	84	1039523-1097850	
			2	79.4 Kb	99	4959719-5039151	
		<i>M152</i>	1	48.9 Kb	53	1897722-1946683	
			2	34.9 Kb	44	4784847-4819818	
		<i>Go06</i>	1	65 Kb	44	1768158-1833157	
<i>M. massiliense</i>	C2A		<i>M. massiliense T</i>	1	12.5 Kb	21	1600973-1613514
		2	31.3 Kb	33	1620002-1651385		
		3	50.4 Kb	69	3907205-3957680		
		<i>M18</i>	1	62.8 Kb	67	4702725-4765592	
		<i>M154</i>	0	0	0	0	
	C2B	<i>M115</i>	1	11.6 Kb	10	1416841-1428481	
			2	77.1 Kb	102	1624644-1701770	
			3	55.3 Kb	79	3356346-3411651	
		<i>M172</i>	1	55.1 Kb	74	502478-557677	
			2	50.7 Kb	50	546109-596832	
			3	59 Kb	67	1934186-1993225	
			4	31.1 Kb	33	2050376-2081567	
			5	39.4 Kb	45	3711805-3751246	
			6	19.6 Kb	40	3753466-3773078	
		<i>M47J26</i>	1	39.9 Kb	48	1066714-1106668	
			2	12.4 Kb	16	3596408-3608873	
			3	41.4 Kb	42	3823414-3864899	
			<i>M139</i>	1	35.9 Kb	43	2906235-2942215
		2		12.5 Kb	17	5033731-5046263	
<i>M. boletii</i>	C3A	<i>M. boletii T</i>	1	41.6 Kb	47	1684736-1726377	
			2	20.9 Kb	38	1727918-1748849	
			3	12.4 Kb	16	3641720-3654182	
	C3B	<i>M24</i>	1	37.1 Kb	51	560940-598047	
			2	37 Kb	37	1680197-1717263	
			3	17 Kb	21	3830340-3847343	
			4	18.1 Kb	34	5051771-5069955	
			5	26 Kb	35	5155113-5181190	
			6	19.2 Kb	26	5213195-5232444	
			7	26.5 Kb	33	5312024-5338593	

tuberculosis H37Rv resistome and the antibiotic resistance databases indicate that *M. abscessus* encodes an aminoglycoside 29-N-acetyltransferase and aminoglycoside phosphotransferases involved in resistance to aminoglycosides. Also, genetic analyses disclosed 16S rRNA gene mutations

conferring aminoglycoside resistance [4,5,47]. Indeed, the presence of a single rRNA operon in all of the *M. abscessus* genomes favours the occurrence of dominant mutations conferring resistance to aminoglycosides and macrolides. *M. abscessus* genomes encode a rifampin

ADP-ribosyl transferase and monooxygenases potentially involved in resistance to rifampin and tetracyclines. Moreover, *M. abscessus* genomes encode three *tet(M)* genes conferring resistance to tetracycline and doxycycline; the number of *tet(M)* genes was correlated to the resistance to cyclines in *Escherichia coli* [48]. However, *M. massiliense* was reported to be susceptible and *M. abscessus* and *M. bolletii* to be resistant to doxycycline [49]. *M. abscessus* genomes encode resistance to fusidic acid, glycopeptides, MLS (Macrolide-Lincosamide-StreptograminB), phenicols, rifampicin, sulphonamide and trimethoprim. Also, *M. abscessus* genomes encode FolP homologs conferring resistance to cotrimoxazole, homolog of UDP-N-acetylglucosamine 1-carboxyvinyltransferase, a MurA protein conferring resistance to fosfomycin and homologs of 23S rRNA methylases conferring resistance to macrolides. Also, *M. abscessus* genome encodes an *erm(41)* gene which mutations were reported to confer clarithromycin resistance [50]. *In-vitro* tests showed that *M. massiliense* clinical isolates could be distinguished from *M. abscessus* isolates for their susceptibility to ciprofloxacin [51] whereas *M. bolletii* isolates were reported to be resistant to all quinolones [52]. A mutation at codon 90 in *gyrA* gene was reported in clinical isolates of *M. abscessus* exhibiting high resistance to ciprofloxacin [51]. This observation contrasts with our genome analysis, which found no such mutations, suggesting that other mechanisms of resistance may be involved in high-level resistance to quinolones [52]. Accordingly, we found that *M. abscessus* mycobacteria encode *qepA2*, a plasmidic gene conferring quinolone resistance in gram-negative bacteria [53]. *M. abscessus* mycobacteria were reported to be *in-vitro* resistant to penicillin, amoxicillin, cefoxitin, ceftriaxone, cefotaxime and imipenem [4,5]. This contrasts with the fact that they encode Penicillin-binding proteins (PBPs), targets for β -lactam antibiotics (except for tabtoxinine- β -lactam, which inhibits glutamine synthetase), which are essential for peptidoglycan synthesis [54,55]. *M. abscessus* genomes encode an Ambler class A β -lactamase homologous to β -lactamases in gram-negative bacteria and to two β -lactamases in *M. tuberculosis*. β -lactamases inhibitors have not been evaluated against *M. abscessus sensu lato* mycobacteria.

Genome-based analysis of host-interactions

M. abscessus are ubiquitous environmental organisms in soil and water [9] where they may have to cope with amoeba: *M. chelonae*, *M. abscessus*, *M. massiliense* and *M. immunogenum* were reported to survive within *Acanthamoeba polyphaga* trophozoites and cysts [5]. Accordingly, our analyses indicate that *M. abscessus* genomes encode factors implicated in host interactions. The mean number of genes encoding proline-glutamate (PE), proline-proline glutamate (PPE), 10-kDa lipoprotein antigen precursor (LpqH), Mammalian Cell Entry

(MCE), oxidoreductase (Yrbe) and type VII secretion system is of 70 in C1, 80 in C2A, 74 in C2B and 93 in C3. In *M. abscessus*, rough colonies lack *mmpL4* (a gene required for glycopeptidolipid biosynthesis) and lost surface colonization, replication into human macrophages and stimulation of innate immune response; these observations suggested that glycopeptidolipid was a virulence factor [56-58]. Accordingly, glycopeptidolipids are required for sliding motility [59] and biofilm formation [60]. Glycopeptidolipids have also been suspected to inhibit phagocytosis of *M. avium* subsp. *avium* [61]. *M. abscessus* genomes encode MCE proteins similar to *M. tuberculosis* H37Rv. MCE operon promotes internalization of *M. tuberculosis* by mammalian cells [62] and initiates rapid induction of transcription of genes involved in substrate trafficking [63]. The number of *mce* operons which correlated with pathogenicity [64], varies from six in C2B to 13 in C3. In parallel, *M. abscessus* genomes encode 12 (C1) to 21 (C3A, C3B) copies of Yrbe proteins. As for secretion systems, recent evidences showed that mycobacteria evolved specialized type VII secretion systems to transport extracellular proteins across the cell wall [65]. Type VII secretion systems ESX-1 and ESX-5 are involved in cell-to-cell migration of *M. tuberculosis* [65,66]. In *M. abscessus*, our analyses indicate that ESX-3 and ESX-4 systems are conserved (Figure 4). However, *M. abscessus* M139 (C2B) lacks two proteins of the ESX-3 system and *M. abscessus* M93 (C1) lacks ESAT-6 like and CFP-10-like proteins secreted by the ESX-4 system. Interestingly, *M. abscessus* M18 (C2A) encodes ESAT-6 and CFP-10 proteins secreted by ESX-1 system. In addition, there are two or three PE and six (*M. massiliense*, *M. abscessus* M115 or *M. abscessus* 47 J26) to 12 (*M. bolletii* M24) PPE proteins, which are reported to be involved in the virulence of *M. tuberculosis* [67]. Our analyses further indicated that proteins related to phenazine biosynthesis, homogentisate catabolism, phenylacetic acid degradation and DNA degradation might have been transferred from *Actinobacteria* (e.g. *Rhodococcus* spp., *Streptomyces* spp.) and *Pseudomonas* (*Pseudomonas aeruginosa* and *Burkholderia cepacia*). Although distantly related, these bacteria share the same ecosystem as *M. abscessus* within cystic fibrosis microbiota.

Conclusions

Our in-depth genomic analyses indicate that *M. abscessus* has a non-conservative genome, suggesting the possibility of on-going transfer of additional genetic material. Unsurprisingly, *M. abscessus* has already acquired antibiotic resistance. Also, phages have mediated diversity and horizontal gene transfer which drove the rapid evolution of this complex. Indeed, gene transfers have driven the evolution of *M. abscessus* towards three different genomospecies *M. abscessus*, *M. massiliense* and *M. bolletii*; and the evolution of four different yet unnamed subspecies. Each

Availability of supporting data

The data set of Figure 1C supporting the results of this article is available in the TreeBase (<http://treebase.org/treebase-web/home.html>) repository, under the accession URL <http://purl.org/phylo/treebase/phyloids/study/TB2:S15632>.

Reviewer access URL: <http://purl.org/phylo/treebase/phyloids/study/TB2:S15632?x-access-code=6fa2ebc53b96e3ae412a8df19187ab41&format=html>.

The data sets of Figure 1A and B supporting the results of this article are included as the Additional file 1.

The data sets of Figure 4 supporting the results of this article are included as the Additional file 2.

Additional files

Additional file 1: A- Aligned *M. abscessus* genome matrix constructed using Mauve software to infer phylogeny using the Neighbor-Net algorithm in the package SplitsTree4. B- The matrix of binary discrete characters ('0' and '1' for absence and presence, respectively) constructed using the orthologous group data found by orthoMCL to infer phylogeny using the Neighbor-Net algorithm in the package SplitsTree4.

Additional file 2: The matrix used for Heatmap clusterisation of *Mycobacterium abscessus* type VII secretion system compared to *Mycobacterium tuberculosis* H37Rv.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MS carried out the molecular genetic studies, participated in the sequence alignment and drafted the manuscript. MD conceived of the study, and participated in its design and coordination and drafted the manuscript. Both authors read and approved the final manuscript.

Acknowledgments

MS is financially supported by the Infectiopôle Sud Foundation.

Received: 5 September 2013 Accepted: 30 April 2014

Published: 12 May 2014

References

- Kubica GP, Baess I, Gordon RE, Jenkins PA, Kwapinski JB, McDermont C, Pattyn SR, Saito H, Silcox V, Stanford JL, Takeya K, Tsukamura M: **A co-operative numerical analysis of rapidly growing mycobacteria.** *J Gen Microbiol* 1972, **73**:55–70.
- Ross AJ: *Mycobacterium salmoniphilum* sp. nov. from salmonid fishes. *Am Rev Respir Dis* 1960, **81**:241–250.
- Wilson RW, Steingrube VA, Böttger EC, Springer B, Brown-Elliott BA, Vincent V, Jost KC Jr, Zhang Y, Garcia MJ, Chiu SH, Onyi GO, Rossmore H, Nash DR, Wallace RJ Jr: *Mycobacterium immunogenum* sp. nov., a novel species related to *Mycobacterium abscessus* and associated with clinical disease, pseudo-outbreaks, and contaminated metalworking fluids: an international cooperative study on mycobacterial taxonomy. *Int J Syst Evol Microbiol* 2001, **51**:1751–1764.
- Adékambi T, Reynaud-Gaubert M, Greub G, Gevaudan MJ, La Scola B, Raoult D, Drancourt M: *Amoebal* coculture of "*Mycobacterium massiliense*" sp. nov. from the sputum of a patient with hemoptoic pneumonia. *J Clin Microbiol* 2004, **42**:5493–5501.
- 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.
- Simmon KE, Brown-Elliott BA, Ridge PG, Durtschi JD, Mann LB, Slechts ES, Steigerwalt AG, Moser BD, Whitney AM, Brown JM, Voelkerding KV, McGowan KL, Reilly AF, Kim TJ, Butler WR, Edelstein PH, Wallace RJ Jr, Petti CA: *Mycobacterium chelonae-abscessus* complex associated with sinopulmonary disease, Northeastern USA. *Emerg Infect Dis* 2011, **9**:1692–1700.
- Petrini B: *Mycobacterium abscessus*: an emerging rapid-growing potential pathogen. *APMIS* 2006, **5**:319–328.
- Bryant JM, Grogono DM, Greaves D, Foweraker J, Roddick I, Inns T, Reacher M, Aworth CSH, Curran MD, Harris SR, Peacock SJ, Parkhill J, Floto RA: **Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study.** *The Lancet* 2013, **381**:1551–1560.
- Medjahed H, Gaillard JL, Reyat JM: *Mycobacterium abscessus*: a new player in the mycobacterial field. *Trends Microbiol* 2010, **3**:117–123.
- Jassies-van der Lee A, Houwers DJ, Meertens N, van der Zanden AG, Willemsse T: **Localised pyogranulomatous dermatitis due to *Mycobacterium abscessus* in a cat: a case report.** *Vet J* 2009, **2**:304–306.
- Albini S, Mueller S, Bornand V, Gutzwiller ME, Burnand C, Hüsey D, Abriel C, Reitt K, Korczak BM, Miserez R: **Cutaneous atypical mycobacteriosis due to *Mycobacterium massiliense* in a cat. (Article in German).** *Schweiz Arch Tierheilkd* 2007, **12**:553–558.
- Whitaker BR, Hadfield CA, Clayton LA, Neiffer D, Stamper MA: *Mycobacterium abscessus* infection in an Atlantic bottlenose dolphin (*Tursiops truncatus*). In *Proceedings of the 36th Annual Conference of the International Association of Aquatic Animal Medicine*; 2012:49–51.
- Wienschmann A, Armien A, Harris NB, Brown-Elliott BA, Wallace RJ Jr, Rasmussen J, Willette M, Wolf T: **Disseminated panniculitis in a bottlenose dolphin (*Tursiops truncatus*) due to *Mycobacterium chelonae* infection.** *J Zoo Wildl Med* 2008, **39**:412–440.
- Clayton LA, Stamper MA, Whitaker BR, Hadfield CA, Simons B, Mankowski JL: *Mycobacterium abscessus* pneumonia in an Atlantic bottlenose dolphin (*Tursiops truncatus*). *J Zoo Wildl Med* 2012, **43**:961–965.
- Zerihun MA, Nilsen H, Hodneland S, Colquhoun DJ: *Mycobacterium salmoniphilum* infection in farmed Atlantic salmon, *Salmo salar* L. *J Fish Dis* 2011, **34**:769–781.
- Leao SC, Tortoli E, Euzéby JP, Garcia MJ: **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* 2011, **61**:2311–2313.
- Adékambi T, Colson P, Drancourt M: *rpoB*-based identification of nonpigmented and late-pigmenting rapidly growing mycobacteria. *J Clin Microbiol* 2003, **41**:5699–5708.
- Viana-Niero C, Lima KV, Lopes ML, Rabello MC, Marsola LR, Brillhante VC, Durham AM, Leao SC: **Molecular characterization of *Mycobacterium massiliense* and *Mycobacterium bolletii* in isolates collected from outbreaks of infections after laparoscopic surgeries and cosmetic procedures.** *J Clin Microbiol* 2008, **46**:850–855.
- Arnold C, Barrett A, Cross L, Magee JG: **The use of *rpoB* sequence analysis in the differentiation of *Mycobacterium abscessus* and *Mycobacterium chelonae*: a critical judgement in cystic fibrosis?** *Clin Microbiol Infect* 2012, **18**:E131–133.
- Leao SC, Tortoli E, Viana-Niero C, Ueki SY, Lima KV, Lopes ML, Yubero J, Menendez MC, Garcia MJ: **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* 2009, **47**:2691–2698.
- Macheras E, Roux AL, Ripoll F, Sivadon-Tardy V, Gutierrez C, Gaillard JL, Heym B: **Inaccuracy of single-target sequencing for discriminating species of the *Mycobacterium abscessus* group.** *J Clin Microbiol* 2009, **47**:2596–2600.
- Macheras E, Roux AL, Bastian S, Leão SC, Palaci M, Sivadon-Tardy V, Gutierrez C, Richter E, Rusch-Gerdes S, Pfyffer G, Bodmer T, Cambau E, Gaillard JL, Heym B: **Multilocus sequence analysis and *rpoB* sequencing of *Mycobacterium abscessus* (sensu lato) strains.** *J Clin Microbiol* 2011, **49**:491–499.
- Zelazny AM, Root JM, Shea YR, Colombo RE, Shampura IC, Stock F, Conlan S, McNulty S, Brown-Elliott BA, Wallace RJ Jr, Olivier KN, Holland SM, Sampaio EP: **Cohort study of molecular identification and typing of *Mycobacterium abscessus*, *Mycobacterium massiliense*, and *Mycobacterium bolletii*.** *J Clin Microbiol* 2009, **47**:1985–1995.
- Sassi M, Ben Kahla I, Drancourt M: *Mycobacterium abscessus* multispacer sequence typing. *BMC Microbiol* 2013, **13**:1–3.

25. Ripoll F, Pasek S, Schenowitz C, Dossat C, Barbe V, Rottman M, Macheras E, Heym B, Herrmann JL, Daffé M, Brosch R, Risler JL, Gaillard JL: **Non-mycobacterial virulence genes in the genome of the emerging pathogen *Mycobacterium abscessus***. *PLoS One* 2009, **4**:e5660.
26. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ: **Prodigal: prokaryotic gene recognition and translation initiation site identification**. *BMC Bioinformatics* 2010, **11**:119.
27. Wozniak M, Wong L, Tiuryn J: **CAMBer: an approach to support comparative analysis of multiple bacterial strains**. *BMC Genomics* 2011, **12**(2):S6.
28. Li L, Stoeckert CJ Jr, Roos DS: **OrthoMCL: identification of ortholog groups for eukaryotic genomes**. *Genome Res* 2003, **13**:2178–2189.
29. Sentausta E, Fournier PE: **Advantages and limitations of genomics in prokaryotic taxonomy**. *Clin Microbiol Infect* 2013, **19**:790–795.
30. Darling AE, Mau B, Perna NT: **progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement**. *PLoS One* 2010, **5**:e11147.
31. Bryant D, Moulton V: **Neighbor-net: an agglomerative method for the construction of phylogenetic networks**. *Mol Biol Evol* 2004, **21**:255–265.
32. Raiol T, Ribeiro GM, Maranhão AQ, Bocca AL, Silva-Pereira I, Junqueira-Kipnis AP, Brígido Mde M, Kipnis A: **Complete genome sequence of *Mycobacterium massiliense***. *J Bacteriol* 2012, **194**:5455.
33. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA: **The COG database: an updated version includes eukaryotes**. *BMC Bioinformatics* 2003, **4**:41.
34. Konstantinidis KT, Ramette A, Tiedje JM: **The bacterial species definition in the genomic era**. *Philos Trans R Soc Lond B Biol Sci* 2006, **361**:1929–1940.
35. Konstantinidis KT, Tiedje JM: **Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead**. *Curr Opin Microbiol* 2007, **10**:504–509.
36. Canchaya C, Proux C, Fournou G, Bruttin A, Brüssow H: **Prophage genomics**. *Microbiol Mol Biol Rev* 2003, **67**:238–276.
37. Choi GE, Cho YJ, Koh WJ, Chun J, Cho SN, Shin SJ: **Draft genome sequence of *Mycobacterium abscessus* subsp. *bolletii* BD (T)**. *J Bacteriol* 2012, **194**:2756–2757.
38. Wong YL, Choo SW, Tan JL, Ong CS, Ng KP, Ngeow YF: **Draft genome sequence of *Mycobacterium bolletii* strain M24, a rapidly growing mycobacterium of contentious taxonomic status**. *J Bacteriol* 2012, **194**:4475.
39. Sassi M, Bebeacua C, Drancourt M, Cambillau C: **The first structure of a mycobacteriophage, *Araucaria***. *J Virol* 2013, Epub ahead of print.
40. Choo SW, Wong YL, Leong ML, Heydari H, Ong CS, Ng KP, Ngeow YF: **Analysis of the genome of *Mycobacterium abscessus* strain M94 reveals an uncommon cluster of tRNAs**. *J Bacteriol* 2012, **194**:5724.
41. Adékambi T, Drancourt M: ***Mycobacterium bolletii* respiratory infections**. *Emerg Infect Dis* 2009, **15**:302–305.
42. Iyer A, Barbour E, Azhar E, Qadri I, Chaudhary A, Abuzenadah A, Kumosani T, Damanhour G, Bahjri S, Al-Hejin A, Abdel Nour AM, Harakeh S: **Bacteriophages in *Escherichia coli* antimicrobial resistance**. *Adv Bioscience Biotechnol* 2013, **4**:469–476.
43. Kim HY, Kim BJ, Kook Y, Yun YJ, Shin JH, Kim BJ, Kook YH: ***Mycobacterium massiliense* is differentiated from *Mycobacterium abscessus* and *Mycobacterium bolletii* by erythromycin ribosome methyltransferase gene (*erm*) and clarithromycin susceptibility patterns**. *Microbiol Immunol* 2010, **54**:347–353.
44. Koh WJ, Jeon K, Lee NY, Kim BJ, Kook YH, Lee SH, Park YK, Kim CK, Shin SJ, Huitt GA, Daley CL, Kwon OJ: **Clinical significance of differentiation of *Mycobacterium massiliense* from *Mycobacterium abscessus***. *Am J Respir Crit Care Med* 2011, **183**:405–410.
45. Paulsen IT, Skurray RA, Tam R, Saier MH Jr, Turner RJ, Weiner JH, Goldberg EB, Grinius LL: **The SMR family: a novel family of multidrug efflux proteins involved with the efflux of lipophilic drugs**. *Mol Microbiol* 1996, **19**:1167–1175.
46. Ramón-García S, Martín C, De Rossi E, Ainsa JA: **Contribution of the Rv2333c efflux pump (the Stp protein) from *Mycobacterium tuberculosis* to intrinsic antibiotic resistance in *Mycobacterium bovis* BCG**. *J Antimicrob Chemother* 2007, **59**:544–547.
47. Nessar R, Cambau E, Reyat JM, Murray A, Gicquel B: ***Mycobacterium abscessus*: a new antibiotic nightmare**. *J Antimicrob Chemother* 2012, **67**:810–818.
48. Hu GZ, Pan YS, Wu H, Hu H, Xu R, Yuan L, Liu JH, Feng JK: **Prevalence of tetracycline resistance genes and identification of tet(M) in clinical isolates of *E. coli* from sick ducks in China**. *J Med Microbiol* 2013, **62**:851–858.
49. Duarte RS, Lourenço MC, Fonseca Lde S, Leão SC, Amorim Ede L, Rocha IL, Coelho FS, Viana-Niero C, Gomes KM, da Silva MG, Lorena NS, Pitombo MB, Ferreira RM, Garcia MH, de Oliveira GP, Lupi O, Vilaça BR, Serradas LR, Chebabo A, Marques EA, Teixeira LM, Dalcolmo M, Senna SG, Sampaio JL: **Epidemic of postsurgical infections caused by *Mycobacterium massiliense***. *J Clin Microbiol* 2009, **47**:2149–2155.
50. Maurer FP, Rüegger V, Ritter C, Bloemberg GV, Böttger EC: **Acquisition of clarithromycin resistance mutations in the 23S rRNA gene of *Mycobacterium abscessus* in the presence of inducible *erm*(41)**. *J Antimicrob Chemother* 2012, **67**:2606–2611.
51. Monego F, Duarte RS, Nakatani SM, Araújo WN, Riediger IN, Brockelt S, Souza V, Cataldo JI, Dias RC, Biondo AW: **Molecular identification and typing of *Mycobacterium massiliense* isolated from postsurgical infections in Brazil**. *Braz J Infect Dis* 2011, **15**:436–441.
52. de Moura VC, da Silva MG, Gomes KM, Coelho FS, Sampaio JL, Mello FC, Lourenço MC, Amorim Ede L, Duarte RS: **Phenotypic and molecular characterization of quinolone resistance in *Mycobacterium abscessus* subsp. *bolletii* recovered from postsurgical infections**. *J Med Microbiol* 2012, **61**:115–125.
53. Cattot V, Poirel L, Nordmann P: **Plasmid-mediated quinolone resistance pump QepA2 in an *Escherichia coli* isolate from France**. *Antimicrob Agents Chemother* 2008, **52**:3801–3804.
54. Basu J, Chattopadhyay R, Kundu M, Chakrabarti P: **Purification and partial characterization of a penicillin-binding protein from *Mycobacterium smegmatis***. *J Bacteriol* 1992, **174**:4829–4832.
55. Cayrou C, Henrissat B, Gouret P, Pontarotti P, Drancourt M: **Peptidoglycan: a post-genomic analysis**. *BMC Microbiol* 2012, **18**:294.
56. Catherinot E, Clarissou J, Etienne G, Ripoll F, Emile JF, Daffé M, Perronne C, Soudais C, Gaillard JL, Rottman M: **Hypervirulence of a rough variant of the *Mycobacterium abscessus* type strain**. *Infect Immun* 2007, **75**:1055–1058.
57. Nessar R, Reyat JM, Davidson LB, Byrd TF: **Deletion of the *mmpL4b* gene in the *Mycobacterium abscessus* glycopeptidolipid biosynthetic pathway results in loss of surface colonization capability, but enhanced ability to replicate in human macrophages and stimulate their innate immune response**. *Microbiology* 2011, **157**:1187–1195.
58. Davidson LB, Nessar R, Kempaiah P, Perkins DJ, Byrd TF: ***Mycobacterium abscessus* glycopeptidolipid prevents respiratory epithelial TLR2 signaling as measured by H β D2 gene expression and IL-8 release**. *PLoS One* 2011, **6**:e29148.
59. Recht J, Martínez A, Torello S, Kolter R: **Genetic analysis of sliding motility in *Mycobacterium smegmatis***. *J Bacteriol* 2000, **182**:4348–4351.
60. Recht J, Kolter R: **Glycopeptidolipid acetylation affects sliding motility and biofilm formation in *Mycobacterium smegmatis***. *J Bacteriol* 2001, **183**:5718–5724.
61. Villeneuve C, Etienne G, Abadie V, Montrozier H, Bordier C, Laval F, Daffe M, Maridonneau-Pariin I, Astarie-Dequeker C: **Surface-exposed glycopeptidolipids of *Mycobacterium smegmatis* specifically inhibit the phagocytosis of mycobacteria by human macrophages. Identification of a novel family of glycopeptidolipids**. *J Biol Chem* 2003, **278**:5129–5130.
62. El-Shazly S, Ahmad S, Mustafa AS, Al-Attiayh R, Krajci D: **Internalization by HeLa cells of latex beads coated with mammalian cell entry (Mce) proteins encoded by the *mce3* operon of *Mycobacterium tuberculosis***. *J Med Microbiol* 2007, **56**:1145–1151.
63. Stavrum R, Valvatne H, Stavrum AK, Riley LW, Ulvestad E, Jonassen I, Doherty TM, Grewal HM: ***Mycobacterium tuberculosis* Mce1 protein complex initiates rapid induction of transcription of genes involved in substrate trafficking**. *Genes Immun* 2012, **13**:496–502.
64. Ishikawa J, Yamashita A, Mikami Y, Hoshino Y, Kurita H, Hotta K, Shiba T, Hattori M: **The complete genomic sequence of *Nocardia farcinica* IFM 10152**. *Proc Natl Acad Sci U S A* 2004, **101**:14925–14930.
65. Abdallah AM, van Pittius Gey NC, Champion PA, Cox J, Lurink J, Vandenbroucke-Grauls CM, Appelmek BJ, Bitter W: **Type VII secretion-mycobacteria show the way**. *Nat Rev Microbiol* 2007, **5**:88–91.
66. Simeone R, Bottai D, Brosch R: **ESX/type VII secretion systems and their role in host-pathogen interaction**. *Curr Opin Microbiol* 2009, **12**:4–10.
67. Mukhopadhyay S, Balaji KN: **The PE and PPE proteins of *Mycobacterium tuberculosis***. *Tuberculosis (Edinb)* 2011, **91**:441–447.
68. Snel B, Bork P, Huynen MA: **Genome phylogeny based on gene content**. *Nat Genet* 1999, **21**:108–110.

doi:10.1186/1471-2164-15-359

Cite this article as: Sassi and Drancourt: Genome analysis reveals three genomospecies in *Mycobacterium abscessus*. *BMC Genomics* 2014 **15**:359.