

Whole genome sequence of *Mycobacterium kansasii* isolates of the genotype 1 from Brazilian patients with pulmonary disease demonstrates considerable heterogeneity

Edson Machado¹, Sidra Ezidio Gonçalves Vasconcellos², Camillo Cerdeira^{1,3}, Lia Lima Gomes², Ricardo Junqueira³, Luciana Distasio de Carvalho⁴, Jesus Pais Ramos⁴, Paulo Redner⁴, Carlos Eduardo Dias Campos⁴, Paulo Cesar de Souza Caldas⁴, Ana Paula Chaves Sobral Gomes⁴, Telma Goldenberg⁴, Fatima Fandinho Montes⁴, Fernanda Carvalho de Queiroz Mello⁵, Vinicius de Oliveira Mussi⁶, Elena Lasunskiaia⁶, Dick van Soolingen⁷, Antonio Basílio de Miranda³, Leen Rigouts⁸, Bouke C de Jong⁸, Conor J Meehan⁸, Marcos Catanho¹, Philip N Suffys^{2,8/+}

¹Fundação Oswaldo Cruz-Fiocruz, Instituto Oswaldo Cruz, Laboratório de Genômica Funcional e Bioinformática, Rio de Janeiro, RJ, Brasil

²Fundação Oswaldo Cruz-Fiocruz, Instituto Oswaldo Cruz, Laboratório de Biologia Molecular Aplicada a Micobactérias, Rio de Janeiro, RJ, Brasil

³Fundação Oswaldo Cruz-Fiocruz, Instituto Oswaldo Cruz, Laboratório de Biologia Computacional e Sistemas, Rio de Janeiro, RJ, Brasil

⁴Fundação Oswaldo Cruz-Fiocruz, Escola Nacional de Saúde Pública, Centro de Referência Professor Hélio Fraga, Laboratório de Referência Nacional para Tuberculose, Rio de Janeiro, RJ, Brasil

⁵Universidade Federal do Rio de Janeiro, Instituto de Doenças do Tórax, Rio de Janeiro, RJ, Brasil

⁶Universidade Estadual do Norte Fluminense Darcy Ribeiro, Laboratório de Biologia do Reconhecer, Campos dos Goytacazes, RJ, Brasil

⁷National Institute for Public Health and the Environment, Bilthoven, the Netherlands

⁸Institute of Tropical Medicine, Unit of Mycobacteriology, Antwerp, Belgium

Mycobacterium kansasii is an opportunistic pathogen and one of the most commonly encountered species in individuals with lung disease. We here report the complete genome sequence of 12 clinical isolates of *M. kansasii* from patients with pulmonary disease in Brazil.

Key words: *Mycobacterium kansasii* - genomes - lung disease - genotype I - diversity - drug resistance

Nontuberculous *Mycobacterium* (NTM) species are widespread in the (man-made) environment and some species cause opportunistic infections in humans. *Mycobacterium kansasii*, frequently isolated from tap water, is a slow-growing photochromogenic NTM and a pathogen that is commonly isolated from patients with pre-existing lung disease, similar to other NTM clinical species such as *M. abscessus*, *M. avium* complex, *M. malmoense* and *M. xenopi*.⁽¹⁾ In the USA and South America *M. kansasii* is the second most isolated NTM after *M. avium* complex^(2,3) while in Rio de Janeiro, it is the most frequent NTM to cause pulmonary disease.⁽⁴⁾ Besides chronic bronchopulmonary disease, *M. kansasii* causes other clinical manifestations such as lymphadenitis⁽⁵⁾, skin and soft tissue infection⁽⁶⁾, tenosynovitis, arthritis, osteomyelitis and disseminated infection in patients co-infected with human immunodeficiency virus.⁽⁷⁾

Seven major genotypes (I to VII) of *M. kansasii* have been described and while human isolates are mainly of types I and II (with type II mainly HIV-related)^(8,9), en-

vironmental isolates are mostly of the other subtypes.⁽¹⁰⁾ Type I has been described as a heterogeneous group, incompletely characterized on the genomic level.^(11,12) Recently, the genome sequence of the ATCC strain 12478 Hauduroy isolated in Kansas in 1955 has been compared with that of the *M. tuberculosis* strain H37Rv⁽¹³⁾, while other *M. kansasii* genomes are available from environmental⁽¹⁴⁾, human⁽¹⁵⁾ and simian⁽¹⁶⁾ sources.

Here, we report the genome sequence of 12 clinical *M. kansasii* isolates belonging to genotype I, as determined by *hsp65* sequencing. The strains were isolated from human Brazilian patients with pulmonary disease and possibly earlier tuberculosis. Nine isolates were from sputum, two from bronchoalveolar lavage (7287 and 10742) and one of unknown origin (1580). Isolates were from residents of the state of Rio de Janeiro (n = 7), Pernambuco (n = 3), Rio Grande do Sul (n = 1), and Santa Catarina (n = 1). Genomic DNA libraries were constructed using the Nextera XT DNA library kit and whole-genome shotgun sequencing was performed on the Illumina HiSeq2500 platform, generating paired end reads of 2 × 100 bp. Genomes were *de novo* assembled using SPAdes software version 3.11.1⁽¹⁷⁾, and annotated with RAST⁽¹⁸⁾ (Table). No evidence for the presence of plasmids was observed.

To evaluate the overall similarity of the genomes of the isolates from Brazil with that of the isolate ATCC 12478, we also compared the 12 genomes with the ATCC reference genome using the BLAST Ring Image Generator (BRIG) program⁽¹⁹⁾ (Fig. 1).

doi: 10.1590/0074-02760180085

Financial support: CNPq (scholarships 207422/2014-1, 500769/2014-1, 311554/2013-0; grants 407624/2012-0, 459100/2014-9).

MC and PNS contributed equally to this work.

+ Corresponding author: psuffys@gmail.com

Received 15 February 2018

Accepted 5 June 2018



TABLE
Genome features, predicted genes, and GenBank accession numbers of
Mycobacterium kansasii strains isolated in Brazil from patients with pulmonary disease

Isolate	Accession	Location	Type (hsp65)	Reads	Genome size (bp)	Contigs	Coverage (x)	Genes
BR3657	PQOL00000000	RJ	1B ^d	18,094,408	6,374,282	294	80.90	5,965t
BR6498 ^a	PQOM00000000	RS	1A	9,038,224	6,225,838	317	37.37	5,804
BR7287	PQON00000000	RJ	1A	19,684,890	6,282,326	274	79.59	5,842
BR6884	PQOO00000000	RJ	1A	21,305,254	6,411,468	280	125.91	5,980
BR6849	PQOP00000000	SC	1A	16,399,304	6,464,047	265	78.89	6,023
BR10953 ^b	PQOQ00000000	RJ	1A	14,895,938	6,284,879	297	52.02	5,850
BR10742 ^b	PQOR00000000	RJ	1A	14,634,490	6,289,619	282	53.58	5,864
BR1580	PQOS00000000	RJ	1B	20,820,372	6,307,487	226	143.20	5,913
BR4404	PQOT00000000	RJ	1B	4,686,134	6,282,923	466	31.43	5,880
BR8835 ^a	PQOU00000000	PE	1B	16,906,236	6,228,381	321	53.45	5,775
BR8837 ^c	PQOV00000000	PE	1B	13,311,444	6,213,268	247	75.41	5,773
BR8839 ^c	PQOW00000000	PE	1B	9,434,892	6,178,862	270	52.48	5,727

a: resistant to rifampicin; *b* and *c* are pairs of isolates obtained from a single patient each. Isolates were from patients residents of Rio Grande do Sul (RS), Rio de Janeiro (RJ), Santa Catarina (SC) and Pernambuco (PE); *d*: 1B differs from the *hsp65* sequence 1A of the ATCC 12478 strain by a SNP A to G at genome position 39635339 (GenBank KJ186611.1).

We observed many deletions spread in the genomes that were either shared by all or part of the isolates. Three isolates (1580, 3657 and 4404) presented a region of difference of about 27 kb causing the loss of five helicases, a restriction endonuclease and two hypothetical proteins. We also observed multiple deletions shared by the isolates 1580, 3657, 4404, 8835, 8837 and 8839 (Fig. 1).

In addition, a reference-based single-nucleotide polymorphism (SNP) calling was performed against the genome of the *M. kansasii* ATCC 12478 strain using both Snippy (version 3.2; <https://github.com/tseemann/snippy>) and the wgSNP module of BioNumerics v7.6 (Applied Maths, Sint-Martens-Latem, Belgium). Most isolates presented homogeneously distributed SNPs over the entire genome, except for isolates 6849 and 6498 which clearly presented regions of higher SNP frequency (data not shown). As demonstrated by a Neighbour-Joining phylogenetic tree based on 5,607,341 sites and 1000 bootstrap replicates (Fig. 2), three groups were observed when compared to the genome of the ATCC strain, one presenting less than 100 SNP ($n = 4$) (middle part), a second with either less than 5,000 SNPs ($n = 1$) and more than 10,000 SNPs ($n = 1$) (lower part) and a third group with more than 10,000 SNPs ($n = 6$) (upper part). We observed an association between SNP-based grouping and phylogeny on basis of large deletions, that led us suggest that the third group has evolved from the common ancestor.

From two patients, we performed whole genome sequencing of two isolates each; the pair of isolates from a patient from Rio de Janeiro (10742 and 10953) had 0 SNPs difference; the second pair from a patient from Recife (8837 and 8839) presented 277 mutations (data not shown).

Additionally, two isolates (6498 and 8835) presented a Minimal Inhibitory Concentration of $> 1 \mu\text{g/mL}$ for rifampicin *in vitro* and both presented a non-synonymous SNP, localized respectively at position 4267612 (C to G)

and 4267613 (A to C), causing an amino acid change in codon 411 of *rpoB* of respectively Gln to Glu and of Gln to Pro. We as such confirm the scarce existing data on the association between resistance against rifampicin and presence of point mutations in *M. kansasii rpoB* gene. Indeed, point mutations in *rpoB* gene are the main reason for resistance against rifampicin and this has been confirmed in several bacteria. Also, molecular determinants of drug resistance in *M. kansasii* has been recently suggested by Bakuła et al.⁽²⁰⁾. Among the present isolates, we observed considerable differences of *in vitro* virulence in macrophages as measured by growth and induction of necrosis and of cytokine production (data not shown).

In conclusion, these are preliminary data on the variability within the *M. kansasii* genotype 1 in Brazil and, even on the basis of this small sample number, we evidenced three genetic groups that are separated by the presence of large number of SNPs throughout the genome and by the pattern of large deletions. One grouping of strains from Rio de Janeiro is highly similar to the ATCC strain Hauduroy (< 100 SNPs) that was isolated more than 70 years ago in the US. A second group of human isolates from Rio de Janeiro and Pernambuco presented over 10,000 SNPs when compared to the ATCC strain. We also observed an association between genotypes and geographic origin of isolates, separating those from the states of Rio de Janeiro, Pernambuco and Santa Catarina. Our observation of considerable differences of *in vitro* and *in vivo* virulence and their differences on the genome level is under investigation.

ACKNOWLEDGEMENTS

To the platforms at Fiocruz (Plataformas de Bioinformática RPT04-A/RJ and Sequenciamento de Ácidos Nucléicos de Nova Geração RPT01-J/RJ). To Applied Maths for offering a one and a half year free access to the Bionumerics software.

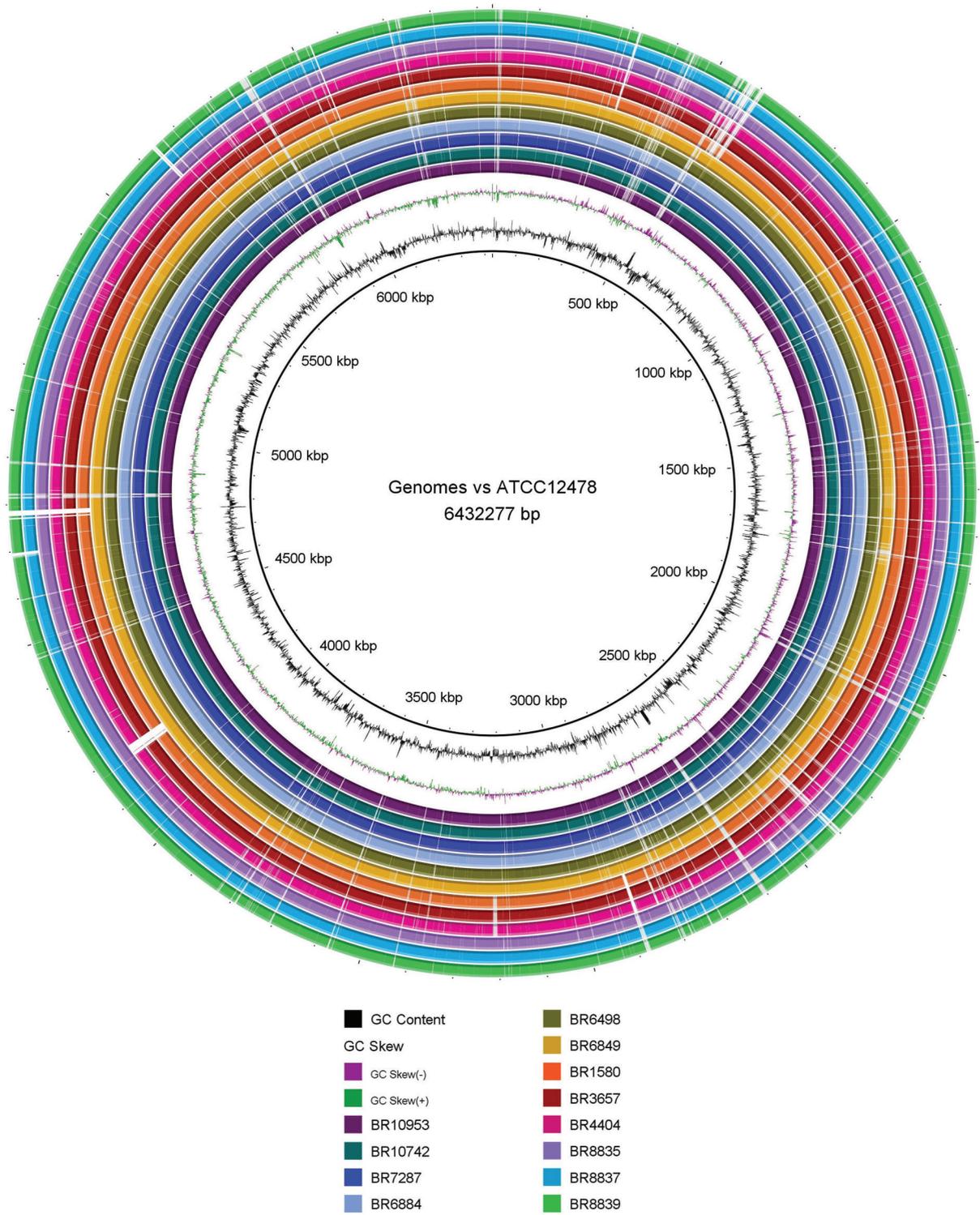


Fig. 1: BLAST ring image of the 12 Brazilian genomes against the ATCC 12478 reference genome.

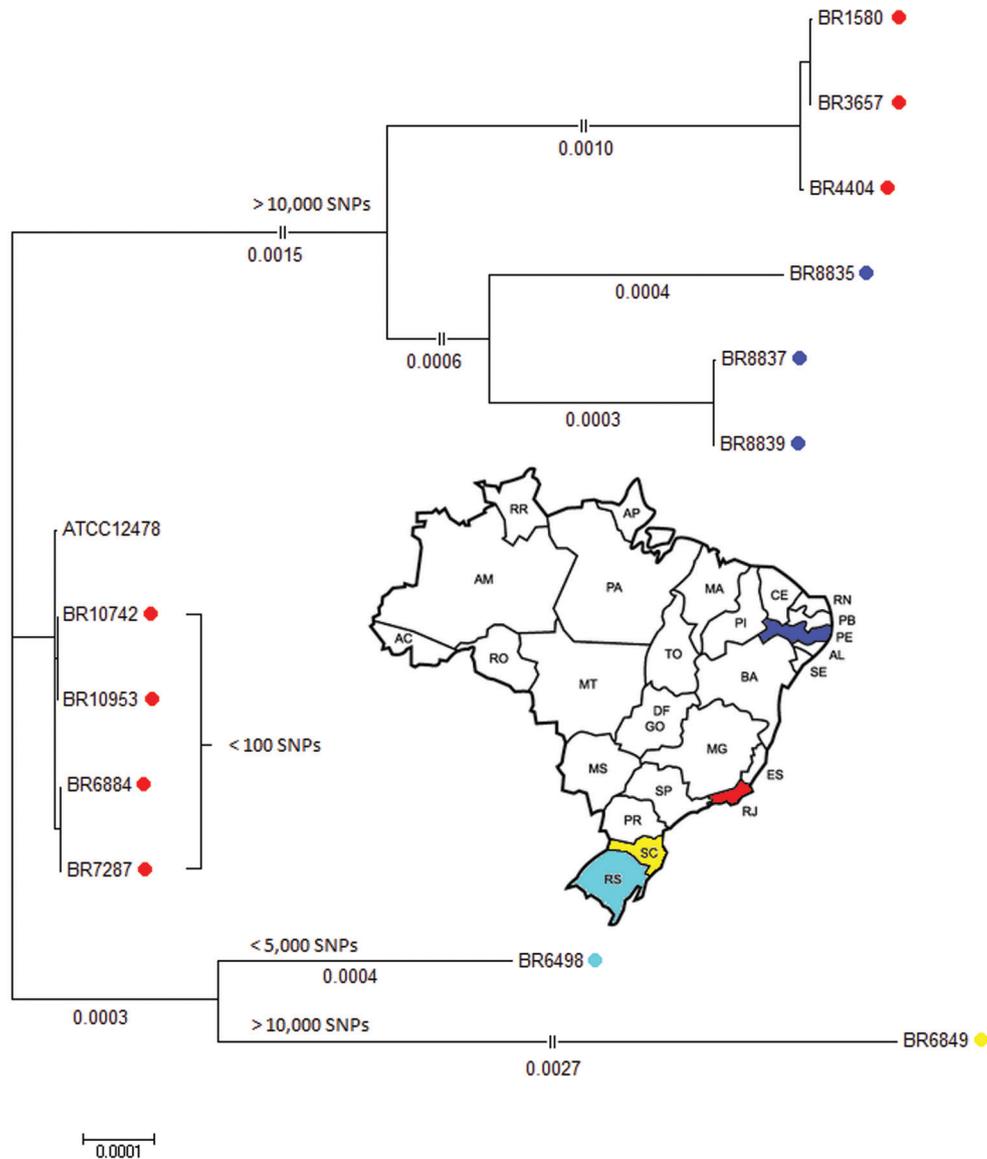


Fig. 2: Neighbor-Joining SNP-based tree of the 12 Brazilian isolates and the ATCC 12478 reference genome.

AUTHORS' CONTRIBUTION

PNS, EL, CM, LR, BCJ and MC conceived and designed the experiments; SEGV, LLG, RJ, LDC and VOM performed the experiments; EM, CC, SEGV, LLG, ABM and CM analysed the data; LDC, JPR, PR, CEDC, PCSC, APCSG, TG, FFM and FCQM were responsible for samples collection, strain isolation and identification; PNS, EM, EL, DS, SEGV, JPR, LR, BCJ and MC contributed for the manuscript writing.

REFERENCES

1. Wassilew N, Hoffmann H, Andrejak C, Lange C. Pulmonary disease caused by non-tuberculous mycobacteria. *Respiration*. 2016; 91(5): 386-402.
2. Hoefsloot W, van Ingen J, Andrejak C, Ängeby K, Bauriaud R, Bemer P, et al. The geographic diversity of nontuberculous mycobacteria isolated from pulmonary samples: An NTM-NET collaborative study. *Eur Respir J*. 2013; 42(6): 1604-13.
3. Kaur P, Fishman JA, Misdraji J, Varma MC, Kotton CN. Disseminated *Mycobacterium kansasii* infection with hepatic abscesses in a renal transplant recipient. *Transpl Infect Dis*. 2011; 13(5): 531-5.
4. de Mello KG, Mello FC, Borga L, Rolla V, Duarte RS, Sampaio EP, et al. Clinical and therapeutic features of pulmonary nontuberculous mycobacterial disease, Brazil, 1993-2011. *Emerg Infect Dis*. 2013; 19(3): 393-9.
5. Choi YG, Cho SY, Lee DG, Yim E, Joo H, Ryu S, et al. *Mycobacterium kansasii* pneumonia with mediastinal lymphadenitis in a patient with acute myeloid leukemia : successful treatment to stem cell transplantation. *Infect Chemother*. 2017; 49(1): 78-83.
6. Han SH, Kim KM, Chin BS, Choi SH, Lee HS, Kim MS, et al. Disseminated *Mycobacterium kansasii* infection associated with skin lesions: a case report and comprehensive review of the literature. *J Korean Med Sci*. 2010; 25(2): 304-8.
7. Menashe L, Kerr LD, Hermann G. *Mycobacterium kansasii* causing chronic monoarticular synovitis in a patient with HIV/AIDS. *J Radiol Case Rep*. 2015; 9(9): 26-35.

8. Santin M, Alcaide F, Benitez MA, Salazar A, Ardanuy C, Podzamczar D, et al. Incidence and molecular typing of *Mycobacterium kansasii* in a defined geographical area in Catalonia, Spain. *Epidemiol Infect.* 2004; 132(3): 425-32.
9. Taillard C, Greub G, Weber R, Pfyffer GE, Bodmer T, Zimmerli S, et al. Clinical implications of *Mycobacterium kansasii* species heterogeneity: Swiss National Survey. *J Clin Microbiol.* 2003; 41(3): 1240-4.
10. Li Y, Pang Y, Tong X, Zheng H, Zhao Y, Wang C. *Mycobacterium kansasii* subtype I is associated with clarithromycin resistance in China. *Front Microbiol.* 2016; 7: 2097.
11. Alcaide F, Richter I, Bernasconi C, Springer B, Hagenau C, Schulze-Röbbecke R, et al. Heterogeneity and clonality among isolates of *Mycobacterium kansasii*: implications for epidemiological and pathogenicity studies. *J Clin Microbiol.* 1997; 35(8): 1959-64.
12. Bakula Z, Safianowska A, Nowacka-Mazurek M, Bielecki J, Jagielski T. Short communication: subtyping of *Mycobacterium kansasii* by PCR-restriction enzyme analysis of the hsp65 gene. *Biomed Res Int.* 2013; 2013: 178725.
13. Wang J, McIntosh F, Radomski N, Dewar K, Simeone R, Enninga J, et al. Insights on the emergence of *Mycobacterium tuberculosis* from the analysis of *Mycobacterium kansasii*. *Genome Biol Evol.* 2015; 7(3): 856-70.
14. Strapagiel D, Borówka P, Marciniak B, Bakula Z, van Ingen J, Safianowska A, et al. Draft genome sequences of *Mycobacterium kansasii* strains 1010001454, 1010001458, 1010001468, 1010001493, 1010001495, and 1010001469, isolated from environmental sources. *Genome Announc.* 2016; 4(3): e00456-16.
15. Borówka P, Lach J, Bakula Z, van Ingen J, Safianowska A, Brzostek A, et al. Draft genome sequences of *Mycobacterium kansasii* clinical strains. *Genome Announc.* 2017; 5(22): e00406-17.
16. Panda A, Nagaraj S, Zhao X, Tettelin H, DeTolla LJ. Complete genome sequences of *Mycobacterium kansasii* strains isolated from *Rhesus* macaques. *Genome Announc.* 2017; 5(16): e00187-17.
17. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012; 19(5): 455-77.
18. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genomics.* 2008; 9: 75.
19. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011; 12: 402.
20. Bakula Z, Modrzejewska M, Penningsb L, Proboszcz M, Safianowskac A, Bieleckia J, et al. Drug susceptibility profiling and genetic determinants of drug resistance in *Mycobacterium kansasii*. *Antimicrob Agents Chemother.* 2018; 62(4): e01788-17.