Genomic Insights Into the Mycobacterium kansasii Complex: An Update

Tomasz Jagielski*, Paulina Borówka*, Zofia Bakula†, Jakub Lach‡, Blażej Marciniak‡, Anna Brzostek§, Jarosław Dziadek§, Mikołaj Dziurzyński§, Lian Pennings⁷, Jakko van Ingen⁷, Manca Žolnir-Dovč⁸ and Dominik Strapagiel⁡

‡ Department of Applied Microbiology, Faculty of Biology, Institute of Microbiology, University of Warsaw, Warsaw, Poland,
§ Biobank Lab, Department of Molecular Biophysics, Faculty of Biology and Environmental Protection, University of Łódź, Łódź, Poland,
† Department of Anthropology, Faculty of Biology and Environmental Protection, University of Łódź, Łódź, Poland,
⁷ BBMRI.pl Consortium, Wrocław, Poland,
⁸ Department of Bacterial Genetics, Faculty of Biology, Institute of Microbiology, University of Warsaw, Warsaw, Poland,
⁹ Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, Netherlands,
⁰ Laboratory for Mycobacteria, University Clinic of Respiratory and Allergic Diseases, Gołnik, Slovenia

Only very recently, has it been proposed that the hitherto existing Mycobacterium kansasii subtypes (I–VI) should be elevated, each, to a species rank. Consequently, the former M. kansasii subtypes have been denominated as Mycobacterium kansasii (former type I), Mycobacterium persicum (II), Mycobacterium pseudokansasii (III), Mycobacterium innocens (V), and Mycobacterium attenuatum (VI). The present work extends the recently published findings by using a three-pronged computational strategy, based on the alignment fraction-average nucleotide identity, genome-to-genome distance, and core-genome phylogeny, yet essentially independent and much larger sample, and thus delivers a more refined and complete picture of the M. kansasii complex. Furthermore, five canonical taxonomic markers were used, i.e., 16S rRNA, hsp65, rpoB, and tuf genes, as well as the 16S-23S rRNA intergenic spacer region (ITS). The three major methods produced highly concordant results, corroborating the view that each M. kansasii subtype does represent a distinct species. This work not only consolidates the position of five of the currently erected species, but also provides a description of the sixth one, i.e., Mycobacterium ostraviense sp. nov. to replace the former subtype IV. By showing a close genetic relatedness, a monophyletic origin, and overlapping phenotypes, our findings support the recognition of the M. kansasii complex (MKC), accommodating all M. kansasii-derived species and Mycobacterium gastri. None of the most commonly used taxonomic markers was shown to accurately distinguish all the MKC species. Likewise, no species-specific phenotypic characteristics were found allowing for species differentiation within the complex, except the non-photochromogenicity of M. gastri. To distinguish, most reliably, between the MKC species, and between M. kansasii and M. persicum in particular, whole-genome-based approaches should be applied. In the absence of clear differences in the distribution of the virulence-associated region of difference 1 genes among the M. kansasii-derived species, the pathogenic potential of each of these species can only be speculatively assessed based on their prevalence among the clinically relevant population. Large-scale molecular epidemiological studies...
are needed to provide a better understanding of the clinical significance and pathobiology of the MKC species. The results of the in vitro drug susceptibility profiling emphasize the priority of rifampicin administration in the treatment of MKC-induced infections, while undermining the use of ethambutol, due to a high resistance to this drug.

Keywords: Mycobacterium kansasii complex, Mycobacterium ostraviense sp. nov., non-tuberculous mycobacteria (NTM), whole genome sequencing, taxonomy

INTRODUCTION

Non-tuberculous mycobacteria (NTM) comprise all species of the Mycobacterium genus, except those aetiologically implicated in tuberculosis (TB) and leprosy, that is members of the M. tuberculosis complex and M. leprae or M. lepromatosis, respectively. More than 180 NTM species have been recognized to date (LPNSN database, 2019). This figure, however, will soon need to be revised, since several new species, on average, are continuously being added every year (Tortoli, 2014). With the increasing number of mycobacterial species descriptions, the number of reported infections, essentially due to NTM, is also growing on a global level. Not as much an enlarging spectrum of NTM species, but more a heightened clinical awareness, expanding population of vulnerable hosts, and advancement of diagnostic and surveillance services are responsible for this scenario (Sood and Parrish, 2017). Although many of the newly described NTM species are potentially pathogenic, having been isolated from clinically affected individuals, only less than a third has consistently been associated with significant health disorders in humans.

Mycobacterium kansasii is one of the most virulent and prevalent NTM pathogen in human medicine. It was first described by Buhler and Pollak in 1953 from a series of respiratory samples of patients with a TB-like pulmonary disease (Buhler and Pollak, 1953). The species was originally named a “yellow bacillus” to emphasize its brilliant yellow pigmentation on exposure to light, but amended thereupon to M. tuberculosis luciflavum by Middlebrook (Middlebrook, 1956) and M. luciflavum by Manten (Manten, 1957). The current species name (M. kansasii) was proposed by Hauduroy in 1955 and refers to where its first isolations were performed (Kansas City, USA) (Hauduroy, 1955). Since the early 1960s, M. kansasii infections have been among the very top of all NTM diseases reported worldwide. A remarkable upsurge in the incidence of M. kansasii infections was seen at the turn of 1980s and 1990s, with the burgeoning of the HIV/AIDS epidemic, and persisted until the early 1990s. A pioneering work by Ross et al. showed the existence of a genetic subspecies of M. kansasii by sequencing of the 5′ end of the 16S rRNA gene (Ross et al., 1992). Subsequent studies involving the amplification of the 16S-23S rRNA spacer region (Abed et al., 1995), PCR-restriction analysis (PRA) of the highly-conserved hsp65 gene (Pilikaytis et al., 1992; Telenti et al., 1993) and Southern blot hybridization with the major polymorphic tandem repeat (MPTR) (Hermans et al., 1992) or insertion sequence-like element, IS1652 (Yang et al., 1993) as a probe have confirmed M. kansasii as a genetically heterogeneous species. The genetic variability of M. kansasii was clearly demonstrated in a study of Picardeau et al. who divided the species into five subspecies or (sub-)types, based on the analysis of restriction fragment length polymorphisms (RFLPs) using the MPTR probe, pulsed-field gel electrophoresis (PFGE), amplified fragment length polymorphism (AFLP) analysis, and
PRA of the hsp65 gene (Picardeau et al., 1997). The validity of the five \textit{M. kansasii} subtypes was further corroborated by sequencing of the 16S-23S rRNA gene or internal transcribed spacer (ITS) region (Alcaide et al., 1997). Somewhat later, two novel types (VI and VII) have been described, according to their hsp65 restriction profiles and sequencing results of the 16S rRNA gene and the 16S-23S rRNA spacer (Richter et al., 1999; Taillard et al., 2003). Moreover, \textit{M. kansasii} isolates with an intermediate type I (I/II) and atypical type II (IIb) have been reported. Whereas, the former had type I-specific sequence of the hsp65 gene and type II-specific sequence of the spacer region, the latter displayed type II-specific spacer sequence and a unique hsp65 gene sequence (Iwamoto and Saito, 2005). The separateness of the \textit{M. kansasii} subspecies was further supported by polymorphisms at other genetic loci, including the RNA polymerase gene \textit{(rpoB)} and the translational elongation factor Tu \textit{(tuf)}, successfully applied for the differentiation between the subspecies I–VI (Kim et al., 2001; Bakula et al., 2016). Noteworthy, the inter-subspecies differences have been detected at the protein level. For each of the six (I–VI) \textit{M. kansasii} subspecies, specific matrix-assisted laser desorption-ionization time-of-flight (MALDI TOF) mass spectral profiles have recently been established (Murugaiyan et al., 2018).

Finally, the genetic diversity of \textit{M. kansasii} exists not only between the subspecies but also between strains of the same subspecies. This has been repeatedly evidenced upon AFLP, PFGE, repetitive unit (rep-)PCR profiling (Alcaide et al., 1997; Linuma et al., 1997; Picardeau et al., 1997; Gaafar et al., 2003; Zhang et al., 2004; Wu et al., 2009; Thomson et al., 2014; Kwenda et al., 2015; Bakula et al., 2018a) and more recently, by a newly developed typing method, based on the analysis of variable number of tandem repeat (VNTR) loci (Bakula et al., 2018a).

Across all genetic studies performed so far on \textit{M. kansasii}, a controversy has been growing over the taxonomic rank of the genetic variants of the species, best reflected by their terminology, which includes subspecies, subtypes, and genotypes. The purpose of this study was to resolve the phylogenetic and taxonomic structure of \textit{M. kansasii} by combining whole genome sequencing with traditional polyphasic classification approaches. The use of a polyphasic strategy, incorporating phylogenetic, biochemical, and chemotaxonomic criteria for resolving the taxonomic identity of mycobacterial species, especially within the NTM group, has been heavily advocated (Saini et al., 2009).

The issue of molecular taxonomy of \textit{M. kansasii} has been addressed in a very recent work by Tagini et al. (2019), which appeared shortly before the completion of our own draft. The present study extends the recently published findings by using a new independent sample and somewhat different methodology, and thus delivers a more refined and complete picture of the \textit{M. kansasii} complex.

**MATERIALS AND METHODS**

**Strains and Culture Conditions**

A total of 27 \textit{M. kansasii} strains were used in the study (Table 1). Included in this number were 21 single-patient, epidemiologically unrelated, clinical isolates and 6 environmental isolates. The strains were originally recovered before 2016 from 8 countries, i.e., the Netherlands \((n=8)\), Poland \((n=5)\), the Czech Republic \((n=4)\), Spain \((n=2)\), Belgium \((n=1)\), Germany \((n=3)\), South Korea \((n=3)\), and Italy \((n=1)\). Patients from whom the strains had been collected were classified as having, or not, a pulmonary NTM disease according to the criteria of the American Thoracic Society (ATS) (Griffith et al., 2007). The type strains of \textit{M. kansasii} (ATCC 12478\textsuperscript{T}) \cite{ATCC12478}, \textit{M. gastri} (DSM 43505\textsuperscript{T}), \textit{M. marinum} (DSM 44344\textsuperscript{T}), \textit{M. szulgai} (DSM 44166\textsuperscript{T}), \textit{M. conspicuum} (DSM 44163\textsuperscript{T}), \textit{M. riyadhense} (DSM 45176\textsuperscript{T}), and \textit{M. tuberculosis} H37Rv were also included in the study.

All strains were maintained as frozen stocks and cultured on Löwenstein-Jensen or Middlebrook 7H10 agar (Becton-Dickinson, Franklin Lakes, USA) medium, supplemented with oleic acid, albumin, dextrose, and catalase, and incubated at either 30 or 37°C.

**Species Identification and Genotyping**

The strains were identified as \textit{M. kansasii} by using high pressure liquid chromatography (HPLC) of cell wall mycolic acids, in accordance with the Centers for Disease Control and Prevention (CDC) guidelines (Butler et al., 1996) and by the GenoType Mycobacterium CM/AS assay (Hain Lifesciences, Nehren, Germany), according to the manufacturer’s instructions. For genotypic identification, total DNA was purified from solid bacterial cultures with a standard extraction method described previously (Santos et al., 1992). Genotyping of \textit{M. kansasii} strains was performed by PCR-RFLP analysis of the hsp65, \textit{rpoB}, and \textit{tuf} genes, as reported elsewhere (Telenti et al., 1993; Kim et al., 2001; Bakula et al., 2016).

**Biochemical Profiling**

The strains were evaluated for a panel of biochemical characteristics by conventional laboratory procedures (CLSI; Clinical and Laboratory Standards Institute, 2011). These comprised tests for niacin accumulation, nitrate and tellurite reduction, Tween 80 hydrolysis (10 days), catalase (thermostable and semi-quantitative), \(\beta\)-glucosidase, arylsulfatase (3 and 14 days), urease (5 days), and pyrazinamidase. Inhibition tests of tolerance to thiphen-2-carboxylic acid hydrazide (TCH), 5% sodium chloride, and acidic (pH 5.5) conditions were also carried out. In addition, cultural features, including colony morphology, photochromogenicity, the ability to grow on MacConkey agar without crystal violet, and at different temperatures (25, 35, and 45°C) were assessed. For each assay, each strain was tested in triplicate. Only if at least two replications produced identical results, the test was considered complete.

**Drug Susceptibility Testing**

To test antimicrobial susceptibility, minimal inhibitory concentrations (MICs) for 13 drugs, including isoniazid (INH), rifampicin (RMP),streptomycin (STR), ethambutol (EMB), clarithromycin (CLR), amikacin (AMK), rifabutin (RFB), moxifloxacin (MFX), ciprofloxacin (CFX), co-trimoxazole (SXT), linezolid (LZD), doxycycline (DOX), and ethionamide (ETO), were determined by the microdilution method, using the Sensititre\textsuperscript{®} SLOMYCO plates (TREK Diagnostic Systems, Cleveland, USA), following the Clinical and Laboratory...
### TABLE 1 | Epidemiological, microbiological, and clinical characteristics of M. kansasii strains under the study.

| No. | ID  | Strain no. | Collected at | Collection date | Geographic location | Host disease | Isolation source | PCR-RFLP | PCR-sequencing | GenBank no. |
|-----|-----|------------|--------------|-----------------|--------------------|--------------|-----------------|----------|----------------|-------------|
|     |     |            |              |                 |                    |              |                 | hsp65    | rpoB | tuf | hsp65 | rpoB | tuf | ITS |
| 1   | 5MK | NLA001000521 | RUMC 2010 | Nijmegen, Netherlands | NTM disease | BAL | I | I | I | ND | ND | 95% | MWKYO1 |
| 2   | 6MK | ATCC25221 | FB 1968 | Borstel, Germany | NTM disease | sputum | I | I | I | I | I | ND | CP019885 |
| 3   | 1MK | NLA001000927 | RUMC 2010 | Nijmegen, Netherlands | NTM disease | sputum | I | I | I | I | I | ND | CP019883 |
| 4   | 4MK | NLA001000449 | RUMC 2010 | Nijmegen, Netherlands | NTM disease | sputum | I | I | I | I | I | ND | CP019884 |
| 5   | 10MK | 6200 WMU 2008 | Warsaw, Poland | NTM disease | sputum | I | I | I | I | I | ND | CP019886 |
| 6   | 9MK | 7728 WMU 2009 | Warsaw, Poland | NTM disease | BW | I | I | I | I | I | ND | CP019888 |
| 7   | 11MK | 7744 WMU 2009 | Warsaw, Poland | NTM disease | BW | I | I | I | I | I | ND | CP019887 |
| 8   | 5JD | 1010001495 RUMC | ND | Czech Republic | NA | water | I | I | I | I | I | ND | LWCL01 |
| 9   | K4  | K4 | SMC 2008 | Ulisan, South Korea | NTM disease | sputum | I | I | I | I | I | ND | MWKX01 |
| 10  | K14 | K14 | SMC 2011 | Seoul, South Korea | NTM disease | sputum | I | I | I | I | I | ND | NKQW01 |
| 11  | K19 | K19 | SMC 2012 | Seoul, South Korea | none | sputum | I | I | I | I | I | ND | NKQX01 |
| 12  | 12MK | II 2193 WMU 2011 | Warsaw, Poland | none | BW | II | II | II | 97% | ND | ND | 95% | LWCM01 |
| 13  | 3MK | II 2100011128 RUMC 2010 | Nijmegen, the Netherlands | none | BAL | II | II | 96% | ND | ND | 95% | MWQA01 |
| 14  | 7MK | II B11073207 RUMC 2011 | Nijmegen, Netherlands | none | sputum | II | II | 97% | ND | ND | 95% | MWKZ01 |
| 15  | 8MK | II B11063838 RUMC 2011 | Nijmegen, Netherlands | none | BAL | II | II | 96% | ND | ND | 95% | MWKV01 |
| 16  | 3B/6JD | II 1010001469 RUMC | ND | Italy | NA | water | II | II | 97% | ND | ND | 95% | NKRA01 |
| 17  | H47 | II H47 | HB 1999 | Bilbao, Spain | NTM disease | sputum | II | II | 97% | ND | ND | 95% | NKQZ01 |
| 18  | H48 | II H48 | HB 1999 | Bilbao, Spain | none | sputum | II | II | 97% | ND | ND | 95% | NKRA01 |
| 19  | 14,15 | III 14,15 IPH 2015 | Šumperk, Czech Rep. | none | sputum | III | III | 96% | ND | ND | 95% | NKRO01 |
| 20  | 174,15 | III 174,15 IPH 2015 | Karviná, Czech Rep. | NTM disease | sputum | III | III | 96% | ND | ND | 95% | NKRD01 |
| 21  | 3JD | III 1010001468 RUMC | ND | Belgium | NA | soil | III | III | 96% | ND | ND | 91% | LWJ001 |
| 22  | 2JD | IV 1010001458 RUMC | ND | Germany | NA | tap water | III | III | 96% | ND | ND | 91% | LWGD01 |
| 23  | 241/15 | IV 241/15 IPH 2015 | Karviná, Czech Rep. | none | sputum | IV | IV | 96% | ND | ND | 91% | LWGD01 |
| 24  | 1JD | V 1010001454 RUMC | ND | Germany | NA | tap water | V | V | V | 95% | ND | 91% | LWGD01 |
| 25  | 4JD | V 1010001493 RUMC | ND | The Netherlands | NA | water | V | V | V | 95% | ND | 91% | LWGD01 |
| 26  | 49,15 | V 49,11 HID 2011 | Warsaw, Poland | none | BW | V | V | V | ND | ND | ND | NKRC01 |
| 27  | 2MK VI NLA001001166 RUMC 2010 | Nijmegen, Netherlands | none | sputum | VI | VI | VI | 96% | ND | ND | 91% | MWKW01 |

---

a According to the WGS-based (MiS1 method) grouping; M. kansasii strains NLA001000521, K14, and K19 had originally been described as types I/II (‘’) and IIb (**), based on the PCR-RFLP/PCR-sequencing analysis of the hsp65, rpoB, and/or tuf genes.

b RUMC, Radboud University Medical Center, Nijmegen, the Netherlands; FB, Forschungszentrum Borstel, Germany; WMU, Department of Internal Medicine, Pulmonology, and Allergology, Warsaw Medical University, Warsaw, Poland; SMC, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea; HB, Hospital de Basurto, Bilbao, Spain; IPH, Institute of Public Health in Ostrava, Czech Republic; HID, Hospital of Infectious Diseases in Warsaw, Poland.

c BAL, bronchoalveolar lavage; BW, bronchial washing.

M. kansasii type was established based on at least 99% similarity of a query sequence with the respective one of the M. kansasii ATCC12478 (type I) reference strain. For other than type I strains, similarity of the query sequence with the respective sequence of the M. kansasii ATCC12478 reference strain was indicated.

NA, not applicable; ND, no data.
TABLE 2 | Genome features of \textit{M. kansasii} strains under the study.

| No. | ID | Type$^a$ | Strain no. | Sequencing platform | Genome size (bp) | Coverage | GenBank no. |
|-----|----|----------|------------|---------------------|------------------|----------|-------------|
| 1   | 5MK | I$^*$    | NLA001000521 | MiSeq               | 6 649 816        | 44.6     | MWKY01      |
| 2   | 6MK | I        | ATCC25221    | MiSeq               | 6 482 452        | 36       | CP019885    |
| 3   | 1MK | I        | NLA001000927 | MiSeq               | 6 421 275        | 35       | CP019883    |
| 4   | 4MK | I        | NLA001000449 | MiSeq               | 6 440 784        | 39       | CP019884    |
| 5   | 10MK| I        | 6200         | MiSeq               | 6 421 364        | 39       | CP019886    |
| 6   | 5MK | I        | 7728         | MiSeq               | 6 463 923        | 33       | CP019888    |
| 7   | 11MK| I        | 7744         | MiSeq               | 6 434 062        | 45.8     | CP019887    |
| 8   | 5JD | I        | 100001496    | NextSeq 500         | 6 358 240        | 28       | LWCL01      |
| 9   | K4  | I        | K4           | NextSeq 500         | 6 503 804        | 58       | NKGW01      |
| 10  | K14 | I$^**$   | K14          | NextSeq 500         | 6 640 863        | 60       | NKQY01      |
| 11  | K19 | I$^**$   | K19          | NextSeq 500         | 6 652 978        | 60       | NKQX01      |
| 12  | 12MK| II       | 2193         | MiSeq               | 6 254 980        | 33       | MWQA01      |
| 13  | 3MK | II       | NLA001001128 | MiSeq               | 6 251 123        | 46       | MWKX01      |
| 14  | 7MK | II       | B11073207    | MiSeq               | 6 123 476        | 34.6     | MWKZ01      |
| 15  | 8MK | II       | B11063838    | MiSeq               | 6 126 434        | 44       | MWKV01      |
| 16  | 3B/6JD | II    | 100001469    | NextSeq 500         | 6 266 032        | 34       | LWCM01      |
| 17  | H47 | II       | H47          | NextSeq 500         | 6 238 170        | 53       | NKRA01      |
| 18  | H48 | II       | H48          | NextSeq 500         | 6 217 650        | 29       | NKRZ01      |
| 19  | 14_15| III      | 14_15        | HiSeq2500           | 6 363 138        | 293      | NKRB01      |
| 20  | 174_15 | III   | 174_15       | HiSeq2500           | 6 342 319        | 173      | NKRD01      |
| 21  | 3JD | III      | 100001488    | NextSeq 500         | 6 141 835        | 24       | LWJC01      |
| 22  | 2JD | IV       | 100001458    | NextSeq 500         | 6 027 332        | 37       | LWKJ01      |
| 23  | 241/15| IV     | 241/15       | HiSeq2500           | 6 093 288        | 147      | NKRE01      |
| 24  | 1JD | V        | 100001454    | NextSeq 500         | 6 171 688        | 25       | LWCH01      |
| 25  | 4JD | V        | 100001493    | NextSeq 500         | 5 627 124        | 29       | LWCK01      |
| 26  | 49_11| V       | 49_11        | HiSeq2500           | 6 634 420        | 280      | NKRFC01     |
| 27  | 2MK | VI       | NLA001001166 | MiSeq               | 6 443 486        | 33       | MWKW01      |

$^a$According to the WGS-based (MiSI method) grouping; \textit{M. kansasii} strains NLA001000521, K14, and K19 had originally been described as types \textit{I/II} ($^*$) and \textit{IIb} ($^**$), based on the PCR-RFLP/PCR-sequencing analysis of the hsp65, rpoB, and/or tuf genes.

Standards Institute (CLSI) recommendations (CLSI; Clinical and Laboratory Standards Institute, 2011).

**Genome Sequencing and Assembly**

For the whole-genome sequencing, chromosomal DNA from all 27 \textit{M. kansasii} strains under the study was extracted by mechanical cell (100 mg pellet) disruption by using zirconia ceramic beads in a FastPrep-24 instrument (MP Biomedicals, Valiant Co., Yantai, China) and further extracted chemically followed with a DNAzol reagent (Invitrogen, Carlsbad, USA).

DNA concentration, its purity and integrity was determined using a Qubit high-sensitivity (HS) assay kit (ThermoFisher, Waltham, USA).

Paired-end libraries were prepared from 1 ng of high-quality genomic DNA with the Nextera XT DNA sample preparation kit according to the manufacturer's instructions (Illumina Inc., San Diego, USA). The libraries were sequenced on a HiSeq 2500 or a NextSeq 500 instrument (Illumina, San Diego, USA) at a read length of 2 $\times$ 150 bp. The quality of reads before and after pre-processing was assessed using FastQC (v0.11.5) (Andrews, 2010). The raw reads were trimmed with TrimGalore/ (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), and \textit{de novo} assembled with the SPAdes Genome Assembler ver. 3.10.0 (Nurk et al., 2013). Scaffold-level assembly was tested using Quast (QUality ASsessment Tool) (Gurevich et al., 2013) to generate basic genome statistics.

In addition to genomes of 27 \textit{M. kansasii} strains, sequenced in this study, genomes of other 53 \textit{Mycobacterium} sp. strains were analyzed. The genomic sequences of these strains were retrieved from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/) and their appropriate accession numbers were provided in Supplementary Table 6.

**Genomic Data Availability**

The assembled genomes were deposited under NCBI Bio-Project accession numbers: PRJNA374853 and PRJNA317047. The genomes were deposited in the GenBank under accession numbers provided in Table 1 and Supplementary Table 6.

**Annotation and Comparative Genome- and Gene-Scale Analyses**

Gene identification and annotation were achieved using the DFAST pipeline v1.0.5 with default settings (Tanizawa et al., 2017).
The genomic relatedness between the strains analyzed was assessed using the MiSI (Microbial Species Identifier) method, based on a combination of genome-wide average nucleotide identity (gANI) and alignment fraction (AF) of orthologous genes (Varghese et al., 2015). The gANI and AF values of 96.5 and 0.6, respectively, were assumed as cutoffs for species delimitation.

As a second method to evaluate the genomic relatedness, the Genome-to-Genome Distance Calculator (GGDC at http://ggdc.dsmz.de) was used. This algorithm was designed to replace standard DNA–DNA hybridization (DDH) by calculating DNA–DNA relatedness (Meier-Kolthoff et al., 2013a). The genome-to-genome distance (GGD) value of 0.0258 was identified as a maximum threshold to assign a genome pair to the same species.

The pairwise ANI and AF values were determined from using ANICalculator (v1.0) (Yoon et al., 2017).

Full-length sequences of the 16S rRNA (1537 bp; MH794239.1), hsp65 (664 bp; NC_000962.3:528608–530230), rpoB (3439 bp; C_000962.3:759807–763325), tuft (1191 bp; NC_000962.3:784821–786011) genes, and the 16S-23S rRNA intergenic spacer region (ITS) (358 bp; AL123456.3:1473359–1473716) were extracted from the whole-genome sequence of the M. tuberculosis reference strain H37Rv, using CLC Genomic Workbench 10 (Qiagen, Valencia, USA), and were used, each, as a reference for alignments with the respective sequences of other mycobacterial species. Sequences from the same loci of the draft genomes analyzed were identified using the blastn algorithm. Due to a significant variation of the sought sequences, whenever necessary, they were extracted manually or downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/).

Identification of RD1-14 Genes

To establish the presence of genes within 14 regions of deletions RD1-14, DIFFIND software with -c 0.7 (sequence identity threshold) and -s 0.5 (length difference cut-off) parameters was used (Marciniak et al., 2017). Identification of RD1-14 genes was based on the alignment of amino acid sequences of predicted genes from the analyzed genomes to the reference database. As references were used amino acid sequences produced by RD1-14 genes from the M. tuberculosis reference strain H37Rv (Accession no.: NC_000962.3) and listed by Brosch et al. (2002).

Phylogenetic Analysis

Core-genome single-ortholog tree was built as described previously (Tagini et al., 2019). Briefly, OrthoFinder (ver. 2.3.4) and MAFFT (ver. 7.310) were used to create core-genome concatamers for the analyzed genomes (Katoch and Standley, 2013; Emms and Kelly, 2019), which were submitted to FastTree (version 2.1.11 Double precision) for maximum-likelihood core-genome phylogeny calculations (Price et al., 2010). The tree was midpoint rooted in FigTree version 1.4.4 (https://github.com/rambaut/figtree).

For single-gene phylogenies, the respective sequences of each genetic locus were subjected to multiple alignment in MEGA X software (ClustalW algorithm) (Kumar et al., 2018). The resulting fragments were further trimmed to remove unnecessary gaps or regions, to a final length of 1,537, 644, 3,439, 1,180, and 277 bp for 16S RNA, hsp65, rpoB, tuft, and 16S-23S rRNA ITS region, respectively. The so prepared sequences were used for evolutionary distance calculation according to the Jukes-Cantor model (Jukes and Cantor, 1969). Phylogenetic trees for each target locus sequences were built using the neighbor-joining method and midpoint rooted with the MEGA X software (Saitou and Nei, 1987). Tree topologies were evaluated by bootstrap analysis based on 1,000 replications (Felsenstein, 1985).

Furthermore, a concatenated tree was constructed based on the 16sRNA, rpoB, and hsp65 gene sequences combined. Sequence concatenation was performed with Sequence Matrix 1.8 software (Vaidya et al., 2011), further corrected in Geneious Prime (Kearse et al., 2012), and plotted in MEGA X. Sequence similarity matrices were generated using Bioedit ver. 7.0.5 (IDENTIFY matrice) (Hall, 1999).

RESULTS AND DISCUSSION

Sequencing of 27 M. kansasii genomes yielded a mean coverage of 66.4x per genome. The average number of contigs per genome was 129 (±229) corresponding to an average N50 score of 4933 kb (±2,605 kb).

The genome sizes and GC contents ranged from 5.6 to 6.6 Mbp (avg. 6.2 Mbp ± 0.3 Mbp) and from 65.86 to 66.38 (avg. 66.14 ± 0.10), respectively. These values were consistent with the data reported for previously assembled M. kansasii genomes (Table 2).

The whole-genome-level relatedness among the 27 M. kansasii strains under the study was assessed with three species identification-relevant parameters, namely the alignment fraction (AF) of orthologous genes, the average nucleotide identity (ANI), and the genome-to-genome distance (GGD) (Tables 3 and 4). The AF/ANI metrics were also computed for genomic sequences of another 32 M. kansasii strains and 24 Mycobacterium sp. strains, representing 5 NTM species and 5 M. tuberculosis complex species, all extracted from NCBI databases (Supplementary Table 6). Whereas, the GGD analysis was performed on 69 genomes in total, including 59 M. kansasii genomes and 10 genomes representing single NTM (other than M. kansasii) and M. tuberculosis complex species (Supplementary Table 6).

The AF values for all M. kansasii strains ranged between 0.6 and 1.0. The lowest AF value recorded for two strains, members of the same M. kansasii genotype was 0.69 (type 1, range 0.69–1.0). Within all other types, the AF values were 0.85 or higher.

All M. kansasii genotypes showed AF values equal to or below 0.75 with strains of other Mycobacterium species, except for M. gastri and M. persicum, which yielded AF values as high as 0.86 with M. kansasii type IV and V (range, 0.79–0.86) or 0.99 with M. kansasii type II (range, 0.89–0.99), respectively (Table 3A).

Pairwise ANI values for any M. kansasii strains affiliated with different types never exceeded the 95–96% threshold (range, 88.6–94.9%), commonly used as a boundary of species delineation (Richter and Rosselló-Móra, 2009; Varghese et al., 2015). Whereas, the ANI values for strains of the same M. kansasii genotype were always higher than 98.8%. The ANI values between M. kansasii and other Mycobacterium
### TABLE 3 | Pairwise alignment fractions (A) and average nucleotide identities (B) for *M. kansasii* subtypes and outgroup species.

| Subtype/group | M. kansasii I | M. kansasii II | M. kansasii III | M. kansasii IV | M. kansasii V | M. tuberculosis | M. aficanum | M. caprae | M. microti | M. conspicuum | M. gaster | M. marinum | M. sydvaranea | M. szulgai |
|---------------|--------------|---------------|---------------|---------------|--------------|---------------|----------|---------|---------|-------------|---------|----------|-------------|---------|
| I             | 0.89–1.0     | 0.89–1.0      | 0.89–1.0      | 0.89–1.0      | 0.89–1.0     | 0.89–1.0      | 0.89–1.0 | 0.89–1.0 | 0.89–1.0 | 0.89–1.0    | 0.89–1.0 | 0.89–1.0 |
| M. kansasii II | 0.65–0.9     | 0.65–0.9      | 0.65–0.9      | 0.65–0.9      | 0.65–0.9     | 0.65–0.9      | 0.65–0.9 | 0.65–0.9 | 0.65–0.9 | 0.65–0.9    | 0.65–0.9 | 0.65–0.9 |
| M. persicum    | 0.89–1.0     | 0.89–1.0      | 0.89–1.0      | 0.89–1.0      | 0.89–1.0     | 0.89–1.0      | 0.89–1.0 | 0.89–1.0 | 0.89–1.0 | 0.89–1.0    | 0.89–1.0 | 0.89–1.0 |
| M. kansasii III | 0.64–0.87   | 0.64–0.87     | 0.64–0.87     | 0.64–0.87     | 0.64–0.87    | 0.64–0.87     | 0.64–0.87 | 0.64–0.87 | 0.64–0.87 | 0.64–0.87   | 0.64–0.87 | 0.64–0.87 |
| M. kansasii IV | 0.64–0.87   | 0.64–0.87     | 0.64–0.87     | 0.64–0.87     | 0.64–0.87    | 0.64–0.87     | 0.64–0.87 | 0.64–0.87 | 0.64–0.87 | 0.64–0.87   | 0.64–0.87 | 0.64–0.87 |
| M. kansasii V | 0.64–0.87   | 0.64–0.87     | 0.64–0.87     | 0.64–0.87     | 0.64–0.87    | 0.64–0.87     | 0.64–0.87 | 0.64–0.87 | 0.64–0.87 | 0.64–0.87   | 0.64–0.87 | 0.64–0.87 |
| MTBC          | 0.64–0.87   | 0.64–0.87     | 0.64–0.87     | 0.64–0.87     | 0.64–0.87    | 0.64–0.87     | 0.64–0.87 | 0.64–0.87 | 0.64–0.87 | 0.64–0.87   | 0.64–0.87 | 0.64–0.87 |
| M. tuberculosis | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |
| M. microti     | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |
| M. conspicuum  | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |
| M. gaster      | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |
| M. marinum     | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |
| M. sydvaranea  | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |
| M. szulgai     | 0.45–0.73   | 0.45–0.73     | 0.45–0.73     | 0.45–0.73     | 0.45–0.73    | 0.45–0.73     | 0.45–0.73 | 0.45–0.73 | 0.45–0.73 | 0.45–0.73   | 0.45–0.73 | 0.45–0.73 |

Each *M. kansasii* subtype was represented by a group of 2–31 strains (genomes). The *M. kansasii* II group included 7 strains (2193, NLA001001128, B11073207, B11063838, 1010001469, H47, and H48); the *M. persicum* included 4 strains [AFPC-000227 (T) (Jagieluk et al., 2017), MK4, MK15, and MK42 (Tagini et al., 2019)]. The *M. kansasii* II group corresponded to the groups of both *M. kansasii* II and *M. persicum* groups (11 strains total). The *M. kansasii* groups III, IV, and V included representatives of respective genomes sampled for this study as well as strains which had elsewhere been designated as *M. pseudotuberculosis* (MK21, MK35, and MK142), *M. innocens* (MK13), and *M. genitalium* (MK136, MK191), accordingly (Tagini et al., 2019). In addition, genomes of six *M. tuberculosis*, five *M. africanum*, five *M. bovis*, and single *M. caprae*, *M. microti*, *M. conspicium*, *M. gaster*, *M. marinum*, *M. rafinesquei*, and *M. szulgai* strains were analyzed. The MTBC contained genomes of all the aforesaid *M. tuberculosis*, *M. africanum*, *M. bovis, caprae*, and single *M. caprae*, *M. microti, conspicium*, *M. gaster, marinum, rafinesquei*, and *M. szulgai* strains (18 strains in total). See Supplementary Table 6. The *values were not given, since only one genome sequence per species was analyzed; AF/ANI values calculated within the groups of strains (genomes) are indicated in bold. The AF and gANI values of 0.6 and 96.5, respectively, were identified as minimum thresholds to assign a genome pair to the same species.
FIGURE 1 | Application of the Microbial Species Identifier (MiSI) method in classifying Mycobacterium kansasii subtype I–VI strains. Separation of different species into different cliques is depicted. Each color represents a species, while each circle—a genome of that species. Within cliques corresponding to M. persicum (formerly subtype II), M. pseudokansasii (III), M. innocens (V), and M. attenuatum (VI), genomes, which had been analyzed previously and first assigned to newly proposed species are shown in other than red color. Within the M. tuberculosis (MTB) complex clique, genomes of M. tuberculosis, M. africanum, M. bovis, M. caprae, and M. microti are depicted in yellow, orange, green, blue, and gray, respectively.

species were all below 95% (range, 80.4–94.9%) except between M. kansasii genotype II and M. persicum (range, 99.7–99.9%) and between M. kansasii genotype IV and M. gastri (95.2%) (Table 3B).

The AF and ANI metrics were also analyzed combinatorially, using the Microbial Species Identifier (MiSI) algorithm. The MiSI method sorts the analyzed genomes into species-like taxa or cliques, based on the AF and ANI species-level cut-off values set at 0.6 and 96.5%, respectively (Varghese et al., 2015). For a total of 83 mycobacterial genomes studied, 15 different cliques were configured. The results were pictorially summarized in Figure 1. The genomes of 59 M. kansasii strains were clearly divided into six cliques, each containing strains of a distinct M. kansasii subtype (I–VI) only. All the remaining NTM species had their genomes clustered within separate cliques, except that four genomes of strains classified as M. persicum were allocated in the M. kansasii genotype II clique. The genomes of 19 strains, representing five M. tuberculosis complex species were accommodated in a single cluster (clique).

The results of the AF/ANI calculations were fully corroborated by the GGD analysis, which serves as an in silico equivalent of the laboratory-based DNA-DNA hybridization (DDH) (Auch et al., 2010; Meier-Kolthoff et al., 2013b). Here, all strains contained within the M. kansasii genotypes shared enough sequence similarity to be considered as separate species (Table 4). The intra-genotype GGD values ranged from 0.0 to 0.0138, and thus fell under the recommended cut-off value of ≤0.0258, corresponding to a 70% DDH cut-off, for species demarcation (Meier-Kolthoff et al., 2013b). Much higher were the GGD values between strains representing different M. kansasii genotypes (range, 0.0601–0.124) or between any M. kansasii and any other Mycobacterium species (range, 0.0563–0.1971). The only exception was when comparing genomes of M. kansasii type II to any of M. persicum, for which the GGD values were between 0.0003 and 0.0044 (Table 4).

Another whole genome-level approach to clarify the taxonomic relationships between the six M. kansasii subtypes was the core-genome phylogenetic analysis. A dendrogram based on the concatenated 615,565-amino-acid sequences from 1,752 single-copy orthologous genes clearly separated all M. kansasii subtypes (Figure 2). All strains belonging to the same subtype formed distinct clades, supported by high bootstrap values (76.7–100%). Noteworthy, strains of M. persicum located in the same clade as the M. kansasii type II strains, whereas M. gastri branched sisterly to M. kansasii type IV. The topology of the tree supports the separation of M. kansasii subtypes, as
TABLE 4 | Genome-to-genome distance (GGD) of *M. kansasii* subtypes and other mycobacterial outgroup species.

| Subtype/group | M. kansasii I | M. kansasii II | M. kansasii III | M. kansasii IV | M. kansasii V | M. kansasii VI | MTBC | M. tuberculosis | M. africanum | M. bovis | M. caprae | M. microti | M. conspicuum | M. gastri | M. marinum | M. riyadhense | M. szulgai |
|---------------|--------------|---------------|----------------|---------------|--------------|---------------|-------|----------------|--------------|------------|-----------|-----------|--------------|---------|----------|---------------|-----------|
| M. kansasii I | 0.0–0.0138   | 0.0004–0.0045 | 0.0002–0.0045  | 0.0002–0.0045 | 0.0002–0.0045 | 0.0002–0.0045 |       |                |              |            |           |           |              |         |          |                |          |
| M. kansasii II| 0.0057–0.077 | 0.0002–0.0045 | 0.0002–0.0045  | 0.0002–0.0045 | 0.0002–0.0045 | 0.0002–0.0045 |       |                |              |            |           |           |              |         |          |                |          |
| M. persicum   | 0.0065–0.0766| 0.0002–0.0044 | 0.0003–0.0044  | 0.0002–0.0044 | 0.0002–0.0044 | 0.0002–0.0044 |       |                |              |            |           |           |              |         |          |                |          |
| M. kansasii III| 0.0707–0.0755| 0.0082–0.018    | 0.0084–0.0181  | 0.0082–0.0187  | 0.0001–0.006  | 0.0005–0.006  |       |                |              |            |           |           |              |         |          |                |          |
| M. kansasiv   | 0.0751–0.0796| 0.0069–0.0707  | 0.0701–0.0707  | 0.0699–0.0706  | 0.0826–0.0837 | 0.0025         |       |                |              |            |           |           |              |         |          |                |          |
| M. kansasii V | 0.0656–0.0688| 0.0601–0.0614  | 0.0601–0.0613  | 0.0602–0.0614  | 0.0715–0.0732  | 0.0647–0.0654  |       |                |              |            |           |           |              |         |          |                |          |
| M. kansasii VI| 0.1096–0.1038| 0.123–0.124     | 0.123–0.124    | 0.123–0.1239   | 0.106–0.1068   | 0.1214–0.1221  |       |                |              |            |           |           |              |         |          |                |          |
| MTBC          | 0.1891–0.1889| 0.1864–0.1879  | 0.1864–0.1879  | 0.1864–0.1879  | 0.1864–0.1879  | 0.1864–0.1879  |       |                |              |            |           |           |              |         |          |                |          |
| M. tuberculosis| 0.1817–0.1816| 0.1866–0.1876  | 0.1866–0.1876  | 0.1867–0.1874  | 0.1864–0.1864  | 0.1892–0.1894  |       |                |              |            |           |           |              |         |          |                |          |
| M. africanum  | 0.1861–0.1882| 0.1864–0.1875  | 0.1864–0.1875  | 0.1864–0.1871  | 0.1865–0.1866  | 0.1851–0.1859  |       |                |              |            |           |           |              |         |          |                |          |
| M. bovis      | 0.1867–0.1886| 0.1866–0.1876  | 0.1866–0.1876  | 0.1865–0.1873  | 0.1868–0.1869  | 0.1854–0.1862  |       |                |              |            |           |           |              |         |          |                |          |
| M. caprae     | 0.1867–0.1884| 0.1866–0.1877  | 0.1866–0.1877  | 0.1867–0.1874  | 0.1867–0.1869  | 0.1852–0.1862  |       |                |              |            |           |           |              |         |          |                |          |
| M. microti    | 0.1869–0.1889| 0.1867–0.1879  | 0.1867–0.1879  | 0.1867–0.1874  | 0.1867–0.1872  | 0.1850–0.1865  |       |                |              |            |           |           |              |         |          |                |          |
| M. conspicuum | 0.1925–0.1947| 0.1915–0.1933  | 0.1915–0.1932  | 0.1925–0.1939  | 0.1921–0.1922  | 0.1929–0.1922  |       |                |              |            |           |           |              |         |          |                |          |
| M. gastri     | 0.0943–0.0966| 0.0785–0.0797  | 0.0785–0.0797  | 0.0878–0.0893  | 0.0833–0.0826  | 0.0741–0.0728  |       |                |              |            |           |           |              |         |          |                |          |
| M. marinum    | 0.1876–0.1905| 0.1886–0.1892  | 0.1886–0.1882  | 0.1886–0.1891  | 0.1893–0.1902  | 0.1891–0.1891  |       |                |              |            |           |           |              |         |          |                |          |
| M. riyadhense | 0.1815–0.1881| 0.1872–0.1878  | 0.1872–0.1878  | 0.1872–0.1876  | 0.1849–0.1852  | 0.1843–0.1857  |       |                |              |            |           |           |              |         |          |                |          |
| M. szulgai    | 0.1886–0.1934| 0.1887–0.1931  | 0.1887–0.1931  | 0.1924–0.1931  | 0.1871–0.1922  | 0.1914–0.1914  |       |                |              |            |           |           |              |         |          |                |          |

Each *M. kansasii* subtype was represented by a group of 2–31 strains (genomes). The *M. kansasii* II group included 7 strains ([2193, NLA001001128, B11073207, B11063838, 1010001469, H47, and H48]); the *M. persicum* group included 4 strains (AFPC−000227 (T) (Shahraki et al., 2017), MK4, MK15, and MK42 (Tagini et al., 2019)); the *M. kansasii* II+ group incorporated strains of both *M. kansasii* II and *M. persicum* groups (11 strains in total). The *M. kansasii* groups III, V, and VI included strains of respective genotypes collected for this study as well as strains which had elsewhere been designated as *M. pseudokansasii* (MK21, MK35, and MK142), *M. innocens* (MK13), and *M. attenuatum* (MK41, MK136, MK191), accordingly (Tagini et al., 2019). In addition, genomes of single *M. tuberculosis* (MTBC), *M. africanum*, *M. bovis*, *M. caprae*, *M. microti*, *M. conspicuum*, *M. gastri*, *M. marinum*, *M. riyadhense*, and *M. szulgai* strains were analyzed. The MTBC contained genomes of all the aforementioned *M. tuberculosis*, *M. africanum*, *M. bovis*, *M. caprae*, and *M. microti* strains (5 strains in total). See Supplementary Table 6.

*The values were not given, since only one genome sequence per species was analyzed. GGDs calculated within the groups of strains (genomes) are indicated in bold. The GGD values lower than 0.0258 were used to assign a genome pair to the same species.*
FIGURE 2 | Maximum-likelihood phylogenetic tree based on the amino acid alignment of concatenated single-copy orthologous genes. Bar, number of amino acid substitutions per amino acid site. Node supports were computed using the Shimodaira-Hasegawa test.
FIGURE 3 | Phylogenetic tree based on 16S rRNA gene sequences, constructed using the Neighbor-Joining method. The bootstrap values were calculated from 1,000 replications. Bootstrap values are given at nodes. Strains of \textit{M. kansasii} type IV (241/15, 101001458) and \textit{M. gastri} are shown boxed in solid and dashed lines, respectively. Strains of \textit{M. kansasii} type I, as per WGS-based analysis, that clustered either within type I or II, upon single-gene or concatenated gene phylogenies are shown in bold. GenBank accession numbers for the sequences are parenthesized. Bar, 0.01 substitutions per nucleotide position.
distinct species, with *M. kansasii* type II being conspecific with *M. persicum*.

The taxonomic position of *M. kansasii* genotypes was also examined using phylogenetic inferences from five genetically conserved loci, including the canonical 16S rRNA gene, the ITS region, and three protein-coding genes, namely hsp65, *tuf*, and *rpoB*, all being widely applied as molecular markers for the classification of mycobacteria (Tortoli, 2014). Multi-alignment and phylogenetic analyses for each of the five loci was conducted on the sequences of all 59 *M. kansasii* strains and single strains of *M. tuberculosis* and five other NTM species (Supplementary Table 6).

Pairwise alignments of the 1,537-bp sequences of 16S rRNA gene from *M. kansasii* strains showed that they were highly similar or identical (99–100% sequence similarity), both within and between the subtypes (Supplementary Table 1). At the same time, strains of *M. kansasii* shared 98–98.6% similarity with *M. tuberculosis* and more than 98% similarity with other NTM species. The 16S rRNA gene sequences were identical between *M. persicum* and *M. kansasii* type II, and between *M. gastri* and *M. kansasii* types I and IV.

Comparisons of the 277-bp ITS sequences from *M. kansasii* strains showed at most 85 and 98% similarities with the corresponding sequences from *M. tuberculosis* and other NTM species, respectively. Only sequences from *M. kansasii* type II and *M. persicum* were almost identical, sharing 99.2–100% similarity (Supplementary Table 2). The ITS sequence similarities within and between different *M. kansasii* subtypes fell under relatively wide ranges, i.e., 92.1–100% and 81.2–100%, respectively. The highest inter-subtype similarity values (>95%) were observed between three *M. kansasii* type I strains (K14, K19, and NLA00100521) and strains of *M. kansasii* type II. The type II-specific ITS sequences of those three type I strains accounted for the high intra-type heterogeneity (92.1–100% sequence similarity).

Sequence analysis of the 644-bp hsp65 gene fragments showed similarities of 90.3–92.3% between *M. kansasii* and *M. tuberculosis*, and 90.9–97.9% between *M. kansasii* and other NTM species, except *M. persicum* which shared 99.5–100% similarity with *M. kansasii* type II (Supplementary Table 3). Alignments of the hsp65 gene sequences from members of different *M. kansasii* subtypes yielded similarities of <98%. The only exception were three strains (K4, K14, and K19) of *M. kansasii* type I sharing up to 99.8% similarity with *M. kansasii* type II strains.

The results of the partial *tuf* (1,180 bp) gene analysis were similar to those obtained with the hsp65 gene (Supplementary Table 4). The *tuf* gene sequence similarities between *M. kansasii* and *M. tuberculosis* were 93.4% at most, while those between *M. kansasii* and other NTM species were always below 98%, except that sequences of *M. kansasii* type II and *M. persicum* were identical. The similarity indexes calculated for *M. kansasii* of different subtypes did not exceed 98.3%, except that two strains of type I (K14 and K19) had the same *tuf* sequences as *M. kansasii* type II strains.

Finally, alignments of the partial *rpoB* (3,439 bp) gene sequences found all *M. kansasii* strains to share <89% sequence similarity with *M. tuberculosis* and <96% similarity with NTM species, but again not *M. persicum*, whose sequences were identical or nearly identical with those of *M. kansasii* type II (Supplementary Table 5). The level of the *rpoB* gene sequence similarity between members of different *M. kansasii* subtypes was consistently below 98%, excluding two type I strains (K14 and K19), which displayed high similarity or identity with type II strains.

To better illustrate the phylogenetic relatedness of *M. kansasii* subtypes, phylogenetic trees inferred from five individual loci were constructed (Figures 3–7). A separate tree was created using the concatenated 16S rRNA, hsp65, and *rpoB* genes (Figure 8), since such an approach is known to increase considerably discrimination and robustness of the dendrogram analysis (Devulder et al., 2005). In all but one dendrograms, all *M. kansasii* strains could be spread into six highly supported (bootstrap values ≥ 93%) clusters, according to their subtype affiliation (Figures 4–8). The 16S rRNA gene-based dendrogram was different in that it contained no cluster specific for *M. kansasii* type IV. Two type IV strains clustered together with *M. kansasii* type I strains (Figure 3). Noteworthy, in the same cluster the type strain of *M. gastri* was placed. A feature, which was apparent across all the trees was that strains of *M. kansasii* type II clustered along with *M. persicum*. Moreover, there were four *M. kansasii* type I strains that branched within that cluster. Two of these strains (K14 and K19) were always present in the *M. kansasii* type II–*M. persicum* cluster, whereas another two belonged to that cluster only in the trees based on the ITS region (strain no. NLA00100521, 16S rRNA gene (NLA00100521), and hsp65 gene (strain no. K4). Having the type I-specific hsp65 gene sequence (sequevar I) and type II-specific ITS sequence (sequevar II), strain no. NLA00100521 represents the so-called intermediate type I (I/II), considered a transitional form between environmental type II and human-adapted type I (Iwamoto and Saito, 2005). Whereas, strains K14 and K19 can be identified as atypical type II (Iib) due to their type II ITS sequence and unique, yet most similar to type II, hsp65 sequences (Iwamoto and Saito, 2005). Captivatingly, strain K4 had the same unique hsp65 gene sequence, with all the other sequences being characteristic of type I. Thus, the strain represents the so far unreported variant of *M. kansasii* type I, which can be tentatively designated as atypical type I (Iib).

Altogether, the results from the genome-wide comparisons demonstrated the six (I–VI) *M. kansasii* subtypes to represent distinct species. This was also the conclusion of the recent study by Tagini et al., who based their results upon ANI and GGD analysis of the genomes of 21 *M. kansasii* strains comprising all six subtypes (Tagini et al., 2019) (Twenty of these strains were used in the present work). Also, single- and multigene phylogenies, highly congruent between this and already published study, were indicative of species-level demarcations between *M. kansasii* subtypes. From these findings, Tagini et al. were first to propose new species designations, namely *M. pseudokansasii*, *M. innocens*, and *M. attenuatum*, replacing the former types III, V, and VI, respectively. The most prevalent type I was preserved under the ‘*M. kansasii*’ designation. Whereas, *M. kansasii* type II was found, as in our
FIGURE 4 | Phylogenetic tree based on ITS sequences constructed using the Neighbor-Joining method. The bootstrap values were calculated from 1,000 replications. Bootstrap values are given at nodes. Strains of M. kansasi type I, as per WGS-based analysis, that clustered either within type I or II, upon single-gene or concatenated gene phylogenies are shown in bold. GenBank accession numbers for the sequences are parenthesized. Bar, 0.01 substitutions per nucleotide position.
Phylogenetic tree based on hsp65 gene sequences constructed using the Neighbor-Joining method. The bootstrap values were calculated from 1,000 replications. Bootstrap values are given at nodes. Strains of *M. kansasii* type I, as per WGS-based analysis, that clustered either within type I or II, upon single-gene or concatenated gene phylogenies are shown in bold. GenBank accession numbers for the sequences are parenthesized. Bar, 0.01 substitutions per nucleotide position.

**FIGURE 5**

- **M. persicum** (M. kansasii II)
- **M. attenuatum** (M. kansasii VI)
- **M. pseudokansasii** (M. kansasii III)
- **M. ostraviense** (M. kansasii IV)
- **M. kansasii** (M. kansasii I)
- **M. innocens** (M. kansasii V)

---

Frontiers in Microbiology | www.frontiersin.org

January 2020 | Volume 10 | Article 2918
Phylogenetic tree based on tuf gene sequences constructed using the Neighbor-Joining method. The bootstrap values were calculated from 1,000 replications. Bootstrap values are given at nodes. Strains of *M. kansasii* type I, as per WGS-based analysis, that clustered either within type I or II, upon single-gene or concatenated gene phylogenies are shown in bold. GenBank accession numbers for the sequences are parenthesized. Bar, 0.01 substitutions per nucleotide position.
FIGURE 7 | Phylogenetic tree based on rpoB sequences constructed using the Neighbor-Joining method. The bootstrap values were calculated from 1,000 replications. Bootstrap values are given at nodes. Strains of *M. kansasi* type I, as per WGS-based analysis, that clustered either within type I or II, upon single-gene or concatenated gene phylogenies are shown in bold. GenBank accession numbers for the sequences are parenthesized. Bar, 0.01 substitutions per nucleotide position.
FIGURE 8 | Phylogenetic tree based on concatenated 16S rRNA, hsp65, and rpoB gene sequences constructed using the Neighbor-Joining method. The bootstrap values were calculated from 1,000 replications. Bootstrap values are given at nodes. Strains of *M. kansasii* type I, as per WGS-based analysis, that clustered either within type I or II, upon single-gene or concatenated gene phylogenies are shown in bold. GenBank accession numbers for the sequences are parenthesized. Bar, 0.01 substitutions per nucleotide position.
study, conspecific with *M. persicum*. Therefore, we share the view of assigning this name to all *M. kansasii* type II strains. In fact, the conspecificity of *M. persicum* and *M. kansasii* type II would have been disclosed upon the original description of the species (*M. persicum*), if the authors had included *M. kansasii* type II in the genome-based comparative analysis (Shahraki et al., 2017).

The two *M. kansasii* type IV strains, analyzed in our study, fully satisfied the genomic criteria for a separate species. This was also implied by our predecessors, but in the absence of any type strain, they could not formally establish the species. Here, we propose a new species name, *Mycobacterium ostraviense* sp. nov., to accommodate *M. kansasii* type IV strains, with a strain no. 241/15, as a type. The description of this new species is given at the end of the article.

The taxonomic status of *M. kansasii* type VII remains an enigma. Neither the strain nor its genomic sequence is available, precluding any relevant phylogenetic analyses. This type was reported only once (Taillard et al., 2003), and given the similarity of its *hsp65* RFLP banding patterns, which served as the only diagnostic means, to those of type III, it is plausible that type VII is a product of misidentification.

Since the mid-1990s, the PCR-RFLP analysis, based on single-copy, orthologous genes (*hsp65*, *rpoB*, and *tuf*) has been widely used for the identification of a plethora of NTM species, including *M. kansasii* and its subtypes (Alcaide et al., 1997; Devallois et al., 1997; Kim et al., 2001; da Silva Rocha et al., 2002; Santin and Alcaide, 2003; Zhang et al., 2004; Kwenda et al., 2015; Bakula et al., 2016). However, PCR-RFLP typing may not seldom produce misleading results. Single nucleotide polymorphisms can alter the recognition sites of the restriction enzymes and thus generate either patterns unidentifiable or corresponding to those of other species. Also, sequence analysis, even conducted on a combination of genes, may not resolve the species identity adequately. This is best illustrated in the already discussed strains of atypical type IIB, which despite sharing nearly 99% average nucleotide identity with type I, had their 16S rRNA, *hsp65*, and *rpoB* gene sequences, either individually or concatenated, more closely associated with type II. The type IIB strains would have been considered type II (*M. persicum*), if they had not been inspected at the whole genome-level. Thus, neither PCR-RFLP profiling nor single or multilocus sequencing allows unequivocal identification of *M. persicum*. A definite diagnosis should be supported by the genome-wide analysis. Still, the two type IIB strains, under this study, displayed the type I-specific *hsp65* RFLP patterns, upon digestion with HaeIII (but not with BstEII, which was type II-specific). This feature can be exploited for differentiation between types I and II, if whole-genome sequencing is not affordable. Nevertheless, a new, robust genetic marker allowing for fast and accurate identification of all *M. kansasii*-derived species (former types I-VI), bypassing the need for whole-genome analysis, would be of great benefit. For this, a more in-depth, comparative analysis of the genomes of more strains representing the six *M. kansasii*-derived species, and other NTM species, will have to be undertaken.

There has been a continuing debate on how the differences between *M. kansasii* subtypes translate into pathogenicity. The prevailing view is that only types I and II are true human pathogens, with the latter having been associated with immunodeficiency, and HIV infection in particular, whereas all the remaining types are considered non-pathogenic, and their sporadic isolation from clinical samples has been interpreted as colonization or environmental contamination (Tortoli et al., 1994; Taillard et al., 2003). Indeed, *M. kansasii* type I is the most commonly detected among clinical isolates and the predominant cause of *M. kansasii* disease worldwide (Alcaide et al., 1997; Kim et al., 2001; Gaafar et al., 2003; Santin and Alcaide, 2003; Taillard et al., 2003; Zhang et al., 2004; da Silva Telles et al., 2005; Shitrit et al., 2006; Thomson et al., 2014; Kwenda et al., 2015; Bakula et al., 2016). Infections attributable to *M. kansasii* type II are much rarer (Taillard et al., 2003; Zhang et al., 2004; Shitrit et al., 2006; Bakula et al., 2016), and those caused by other types are almost unreported in the literature. Still, types III, IV, and VI have been recognized among clinical isolates (Alcaide et al., 1997; Picardeau et al., 1997; Santin and Alcaide, 2003; Thomson et al., 2014; Kwenda et al., 2015) with types IV and VI implicated in the disease (Santin and Alcaide, 2003; Thomson et al., 2014). Due to the paucity of isolations of *M. kansasii* other than type I, their clinical relevance remains obscure. Some new light on this problem may be shed by the findings of an international, multicenter investigation, currently in progress, on the global distribution of *M. kansasii* subtypes. So far, we have documented nine confirmed cases of *M. kansasii* disease, etiologically linked to either types III, IV, V or VI (Jagielski et al., data unpublished). In this context, the newly proposed species names for types V (*M. innocens*) and VI (*M. attenuatum*) may not reflect the true phenotype of those bacteria.

Searching for the genetic determinants of pathogenicity, a key genomic region associated with *M. tuberculosis* virulence, known as “region of difference 1” (RD1) was interrogated across the genomes of *M. kansasii* subtypes for its functional integrity (i.e., presence of RD1 genes). The RD1 encodes a secretary apparatus (ESX-1 type VII secretion system) responsible for exporting two highly potent antigens and virulence factors—the 6-kDa early secreted antigenic target (ESAT-6) and the 10-kDa culture filtrate protein (CFP-10) (Berthet et al., 1998). These proteins, encoded by the same operon, play a critical role in modulation of the host immune response through inhibition of phagosome maturation, cytosolic translocation of mycobacteria or granuloma formation (MacGurn and Cox, 2007; van der Vel et al., 2007; Volkman et al., 2010). The *esat-6* (*esxa*) and *cfp-10* (*exxB*) genes have also been demonstrated in some NTM species, including *M. kansasii* (types I-V), *M. szulgit*, *M. marinum*, and *M. riyadhense* (van Ingen et al., 2009). Furthermore, the ESAT-6/CFP-10-mediated translocation of bacilli into the cytosol has been proven to occur in *M. kansasii* type I, but not in *M. kansasii* type V (Houben et al., 2012).

Our analysis showed the presence of six RD1 genes (*rv3871*-5 and *rv3877*), including the ESAT-6 and CFP-10-coding genes and two other genes (*rv3871* and *rv3877*) coding for the essential components of the ESX-1 secretion system, in all types of *M. kansasii* and other NTM species (*Table 5*). Only single strains of *M. kansasii* type I, *M. marinum*, and *M. szulgit* were devoid of the *rv3872* gene coding for PE35, conjectured to play a role in the regulation of *exxB/A* expression (Brodin et al., 2006).
TABLE 5 | Distribution of region of difference 1 (RD1) genes among *M. kansasii* strains and other mycobacterial outgroup species.

| No. | Species/genotype ID | Strain no. | Locus** | GenBank no. |
|-----|---------------------|------------|---------|-------------|
| 1   | *M. kansasii* 5MK   | NLA001000521^D^ | Rv3879c | + | M. kansasii    |
| 2   | *M. kansasii* 6MK   | ATCC25221^D^ | Rv3879c | + | M. kansasii    |
| 3   | *M. kansasii* 1MK   | NLA001000927^D^ | Rv3879c | + | M. kansasii    |
| 4   | *M. kansasii* 4MK   | NLA001000449^D^ | Rv3879c | + | M. kansasii    |
| 5   | *M. kansasii* 10MK  | 6200^D^ | Rv3879c | + | M. kansasii    |
| 6   | *M. kansasii* 9MK   | 7728^D^ | Rv3879c | + | M. kansasii    |
| 7   | *M. kansasii* 11MK  | 7743^D^ | Rv3879c | + | M. kansasii    |
| 8   | *M. kansasii* 5JD   | 101001495^E^ | Rv3879c | + | M. kansasii    |
| 9   | *M. kansasii* K4    | K4^D^ | Rv3879c | + | M. kansasii    |
| 10  | *M. kansasii* K14   | K14^D^ | Rv3879c | + | M. kansasii    |
| 11  | *M. kansasii* K19   | K19^K^ | Rv3879c | + | M. kansasii    |
| 12  | *M. kansasii* –     | ATCC12478^D^ | Rv3879c | + | M. kansasii    |
| 13  | *M. kansasii* –     | 824^N^ | Rv3879c | + | M. kansasii    |
| 14  | *M. kansasii* –     | BR3657^D^ | Rv3879c | + | M. kansasii    |
| 15  | *M. kansasii* –     | BR6849^D^ | Rv3879c | + | M. kansasii    |
| 16  | *M. kansasii* –     | SMC1^E^ | Rv3879c | + | M. kansasii    |
| 17  | *M. kansasii* –     | BR8837^D^ | Rv3879c | + | M. kansasii    |
| 18  | *M. kansasii* –     | BR6496^D^ | Rv3879c | + | M. kansasii    |
| 19  | *M. kansasii* –     | 11_3813^R^ | Rv3879c | + | M. kansasii    |
| 20  | *M. kansasii* –     | BR6884^D^ | Rv3879c | + | M. kansasii    |
| 21  | *M. kansasii* –     | BR10742^D^ | Rv3879c | + | M. kansasii    |
| 22  | *M. kansasii* –     | BR1580^D^ | Rv3879c | + | M. kansasii    |
| 23  | *M. kansasii* –     | BR4404^D^ | Rv3879c | + | M. kansasii    |
| 24  | *M. kansasii* –     | BR8839^D^ | Rv3879c | + | M. kansasii    |
| 25  | *M. kansasii* –     | 662^N^ | Rv3879c | + | M. kansasii    |
| 26  | *M. kansasii* –     | BR7287^D^ | Rv3879c | + | M. kansasii    |
| 27  | *M. kansasii* –     | BR10695^D^ | Rv3879c | + | M. kansasii    |
| 28  | *M. kansasii* –     | BR8835^D^ | Rv3879c | + | M. kansasii    |
| 29  | *M. kansasii* –     | 11_3469^R^ | Rv3879c | + | M. kansasii    |
| 30  | *M. kansasii* –     | Mk40^J^ | Rv3879c | + | M. kansasii    |
| 31  | *M. kansasii* –     | MK5^U^ | Rv3879c | + | M. kansasii    |
| 32  | *M. kansasii* 12MK  | 2193^N^ | Rv3879c | + | M. kansasii    |
| 33  | *M. kansasii* 3MK   | NLA001001128^N^ | Rv3879c | + | M. kansasii    |
| 34  | *M. kansasii* 7MK   | B11073207^N^ | Rv3879c | + | M. kansasii    |
| 35  | *M. kansasii* 8MK   | B11063838^N^ | Rv3879c | + | M. kansasii    |
| 36  | *M. kansasii* 3B/6JD | 1010001499^F^ | Rv3879c | + | M. kansasii    |
| 37  | *M. kansasii* H47   | H47^D^ | Rv3879c | + | M. kansasii    |

(Continued)
| No. | Species/genotype ID | Strain no.\(^a\) | Locus\(^b\) | GenBank no. |
|-----|---------------------|------------------|-------------|-------------|
|     |                     |                  | Rv3879c | Rv3876 | Rv3878 | Rv3877 | Rv3873 | Rv3871 | Rv3874 | Rv3875 | Rv3872 |
| 38  | M. kansasii II      | H48              | –     | –     | –     | +     | +     | +     | +     | +     | +     | NKQZ01 |
| 39  | M. persicum         | –                | AFPC–000227\(^c\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHK01 |
| 40  | M. persicum         | –                | MK15\(^d\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHM01 |
| 41  | M. persicum         | –                | MK4\(^d\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHL01 |
| 42  | M. persicum         | –                | MK42\(^d\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHL01 |
| 43  | M. kansasii III     | 14_15            | 14_15\(^e\) | +     | +     | –     | –     | –     | +     | +     | +     | NKRB01 |
| 44  | M. kansasii III     | 174_15           | 174_15\(^d\) | –     | –     | –     | +     | +     | +     | +     | +     | NKRD01 |
| 45  | M. kansasii III     | 3JD              | 1010001468\(^f\) | +     | +     | –     | –     | –     | +     | +     | +     | LWcJ01 |
| 46  | M. kansasii III     | –                | 732\(^j\) | +     | +     | –     | –     | –     | +     | +     | +     | JANZ01 |
| 47  | M. pseudokansasii   | –                | MK142\(^d\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHU01 |
| 48  | M. pseudokansasii   | –                | MK21\(^j\) | –     | +     | –     | –     | –     | +     | +     | +     | UPHO01 |
| 49  | M. pseudokansasii   | –                | MK35\(^j\) | –     | +     | –     | –     | –     | +     | +     | +     | UPHN01 |
| 50  | M. kansasii IV      | 2JD              | 1010001458\(^f\) | +     | +     | –     | –     | –     | +     | +     | +     | LWcJ01 |
| 51  | M. kansasii IV      | 241/15           | 241/15\(^n\) | –     | +     | –     | –     | –     | +     | +     | +     | NKRE01 |
| 52  | M. kansasii V       | 1JD              | 1010001454\(^f\) | –     | +     | –     | –     | –     | +     | +     | +     | LWcH01 |
| 53  | M. kansasii V       | 4JD              | 1010001493\(^f\) | –     | +     | –     | –     | –     | +     | +     | +     | LWcK01 |
| 54  | M. kansasii V       | 49_11            | 49_11\(^n\) | –     | +     | –     | –     | –     | +     | +     | +     | NKRC01 |
| 55  | M. innocens         | –                | MK13\(^d\) | –     | +     | –     | –     | –     | +     | +     | +     | UPHQ01 |
| 56  | M. kansasii VI      | 2MK              | NLA001001166\(^n\) | –     | –     | –     | +     | +     | +     | +     | +     | MWkw01 |
| 57  | M. attenuatum       | –                | MK41\(^n\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHT01 |
| 58  | M. attenuatum       | –                | MK136\(^j\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHp01 |
| 59  | M. attenuatum       | –                | MK191\(^j\) | –     | –     | –     | +     | +     | +     | +     | +     | UPHS01 |
| 60  | M. tuberculosis     | –                | Beijing-like | +     | +     | –     | –     | –     | +     | +     | +     | CP017597 |
| 61  | M. africanum        | –                | GM041182 | +     | +     | –     | –     | –     | +     | +     | +     | NC_015758 |
| 62  | M. bovis            | –                | 09_1191 | +     | +     | –     | –     | –     | +     | +     | +     | JPP01 |
| 63  | M. caprae           | –                | Algaeu | +     | +     | +     | +     | +     | +     | +     | +     | CP016401 |
| 64  | M. microti          | –                | 12     | +     | +     | +     | –     | –     | –     | –     | –     | CP010333 |
| 65  | M. conspicuum       | –                | DSM44136 | –     | –     | –     | –     | –     | +     | +     | +     | LQO01.1 |
| 66  | M. gastri           | –                | DSM43505 | –     | +     | –     | –     | –     | +     | +     | +     | LQXo01 |
| 67  | M. marinum          | –                | E11    | –     | +     | –     | –     | –     | +     | +     | +     | H917972 |
| 68  | M. riyadhense       | –                | DSM45178 | –     | –     | –     | +     | +     | +     | +     | +     | LQpQ01 |
| 69  | M. szulgai          | –                | DSM44166 | +     | +     | –     | –     | –     | +     | +     | +     | LQpW01 |

\(^a\)The presence of a given gene was marked with a "+" and highlighted in gray; the superscript letters refer to isolation source: \(^d\)–clinical strains implicated in NTM disease; \(^n\)–clinical strains not implicated in NTM disease; \(^j\)–clinical strain from the rhesus macaque (Macaca mulatta); \(^i\)–clinical strains with unknown relation to NTM disease; \(^f\)–environmental strains. Strains whose genomes were sequenced in this study are marked with an asterisk (*).
The rv3876 gene was demonstrated in all *M. kansasii* types except for types II (*M. persicum*) and VI (*M. attenutatum*). Neither it was present in *M. szulgai* and *M. riyadhense*. The protein encoded by this gene is an ESX-1 secretion-associated protein EspI. It was shown that inactivation of the rv3876 gene did not impair secretion of ESAT-6 (Brodin et al., 2006). More recently, however, EspI was found to negatively regulate the ESX-1 secretion system in *M. tuberculosis*, in response to low cellular ATP levels. EspI was thus hypothesized to play a possible role during the latent phase of infection (Zhang et al., 2014).

The rv3879c gene, coding for another ESX-1 secretion-associated protein EspK, was variably distributed among *M. kansasii* strains. It was detected in half of the strains of types I and III, while being absent in all but one strains of types II and VI, and all type V strains. EspK seems not to be involved in virulence, since the rv3879c homolog deletion mutant of *M. bovis* was not attenuated in the guinea pig model (Inwald et al., 2003). Moreover, similar to rv3876, inactivation of the EspK-coding gene did not abolish ESAT-6 secretion (Brodin et al., 2006). Contrastingly, EspK of *M. marinum* was found crucial for the ESX-1-mediated secretion of EspB (Rv3881c), required for virulence and growth in macrophages (McLaughlin et al., 2007). It is thus conceivable that EspK may influence the pathogenicity also in *M. kansasii*.

Collectively, based on the distribution of RD1 genes, neither of the *M. kansasii* types could be categorized as being more or less pathogenic, given that all genes essential for the functioning of the ESX-1 secretion machinery were uniformly present in *M. kansasii* strains, and that the absence of certain genes was reported in both clinically relevant and neutral strains. To explore more in depth the genetic background of virulence in *M. kansasii*, more advanced, functional studies should be performed, with a focus not only on the RD1 genes but several other genes flanking that cluster, which together form the “extended RD1” region. Furthermore, we cannot exclude that the pathogenic and non-pathogenic *M. kansasii* strains differ in terms of expression of the RD1 genes or activity of their proteins. More RD1-targeted experimental investigations are required to validate such scenarios.

Analysis of other than RD1 regions of deletions (RD2-14) did not show any consistent (i.e., shared across all strains of a given species) species-specific pattern of RD genes (Supplementary Data Sheet 2). It was noteworthy, however, that the only two genes found in the RD3 locus were present in either of the two most pathogenic *M. kansasii* types I and II. The rv1577 gene occurred in more than 90% of *M. kansasii* type I and almost 73% of *M. kansasii* type II. Whereas, the rv1586 gene was demonstrated exclusively in *M. kansasii* type I, at a frequency of nearly 81%. Likewise, only *M. kansasii* types I and II harbored genes of the RD11 locus. The rv2651 gene was present in 90 and 73% of *M. kansasii* type I and II, respectively. The rv2646 was evidenced in slightly more than 60% of *M. kansasii* type I. Interestingly, both RD3 and RD11 represent phage inserts within the *M. tuberculosis* genome, and are thought to generate antigenic variation (Ahmed et al., 2007).

It has been canonically accepted, upon description of new species, to provide a detailed phenotypic characterization. However, in the era of genomic-based bacterial taxonomy, the significance of the phenotype has much eroded and the use of conventional biochemical testing has been increasingly abandoned. The algorithm for routine differential diagnostics of NTM species should obligatorily include only growth rate and pigment production (Tortoli et al., 2017).

For closely-related species, the diagnostic value of phenotyping is virtually negligible, as demonstrated also in this study (Table 6). All strains, irrespective of subtype (species), were almost invariably photochromogenic, niacin-negative, and grew at 25 and 37°C, but not at 45°C, in the presence of 5% (w/v) NaCl or on MacConkey agar without crystal violet. They were all resistant to thiophene-2-carboxylic acid hydrazide (TCH) and presented a strong catalase activity, but none exhibited pyrazinamidase activity or arylsulfatase at 3 days. Some variability was observed when testing for nitrate and tellurite reduction, Tween 80 hydrolysis, and urease activity. The only marked difference between the subtypes (species) was that strains of *M. ostraviense* (formerly *M. kansasii* type IV) were, unlike all other species, unable to reduce nitrate and that their catalase was heat-labile. Whether these features are stable within the species need to be verified on a larger set of strains. Interestingly, both these features are typical for *M. gastri* (Kent and Kubica, 1985), with which *M. ostraviense* shares the highest genetic similarity, as evidenced upon whole-genome analysis. *Mycobacterium gastri*, a casual resident of human stomach and only exceptionally pathogenic (Velayati et al., 2005), is easily distinguishable from all *M. kansasii*-derived species, as it is non-photochromogenic.

Finally, drug susceptibility profiles of 30 mycobacterial strains, including type strains of three newly established species, by Tagini et al. (2019), were compared within and between the species (former *M. kansasii* types) (Table 7). Shorty, all strains were susceptible to Rif, RFB, AMK, SXT, MFX, LZD, and CLR (except one CLR-resistant strain of *M. kansasii*). Of these drugs, only Rif and CLR showed interspecies differences in their activity, with the MICs for *M. kansasii* and *M. persicum* slightly higher than for other *M. kansasii*-derived species. More than 80% of strains were resistant to EMB. Among these, were all strains of *M. kansasii* (former type I), *M. persicum* (II), and *M. ostraviense* (IV). Single strains of *M. kansasii*, *M. pseudokansasii*, and *M. attenutatum* were resistant to CIP. The MICs of STR and DOX varied widely (<0.5–16 mg/L vs. 1–>16 mg/L), yet the highest values (16 vs. >16 mg/L) were observed only for strains of *M. kansasii*, *M. persicum*, and *M. attenutatum*. The INH and ETO MICs were low, and within relatively narrow ranges (<0.25–1 mg/L vs. <0.3–0.6 mg/L). These findings confirm the key observations from previous studies, on the susceptibility of *M. kansasii* strains to Rif and their high resistance to EMB (da Silva Telles et al., 2005; Wu et al., 2009; Shahraji et al., 2017; Bakula et al., 2018c). This, confronted with the ATS recommendation of a three-drug (INH-RIF-EMB) regimen for the treatment of *M. kansasii* disease, speaks for exclusion of EMB and its replacement with other potent drug, such as moxifloxacin.

In conclusion, the present paper updates and extends the findings of earlier investigation on the taxonomy of *M. kansasii*. 
TABLE 6 | Phenotypic characteristics of *M. kansasii* strains under the study.

| Type/Strain no. | Morphology | Pigmentation | Growth in/post 7 days at/°C | Tolerance of: | Niacin | Reduction of: | Arylsulfatase | Tween 80 hydrolysis | Catalase | Urease | PZA |
|----------------|------------|--------------|-----------------------------|--------------|--------|---------------|---------------|------------------|----------|--------|-----|
|                |            |              | 25°C | 37°C | 45°C | MCA | TCH | NaCl 5% | nitrate | tellurite | 3-day | 14-day | SQ | HS |
| *M. kansasii* |            |              |      |      |      |    |    |         |        |         |       |        |    |    |
| I (1)          |            |              |      |      |      |    |    |         |        |         |       |        |    |    |
| **NLA001000521** |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| **ATCC25221** |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| **NLA001000927** |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| **NLA001000449** |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| 6200          |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| 7728          |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| 7744          |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| 1012001495 |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| K4            |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| K14           |          |            |      |      |      |    |    |         |        |         |       |        |    |    |
| K19           |          |            |      |      |      |    |    |         |        |         |       |        |    |    |

All test reactions are given as: *"+" (positive), "-" (negative) or "v" (variable).* NT, not tested. Numbers in superscripts represent percentage of strains reacting as indicated.

*Results according to Kent and Kubica (1985); prior to subtype delineation.

**Results according to Shahraki et al. (2017).**

***Results according to Taghin et al. (2019).**

1. Colony morphology: S, smooth; R, rough; Y, yellow/bronze; W, white.

2. Photochromogenicity: N, non-photochromogenic; P, photochromogenic; S, scotochromogenic.

3. MCA, MacConkey agar without crystal violet.

4. TCH, thiopeine-2-carboxylic acid hydrazide.

5. After 10 days.

6. SQ, semi-quantitative catalase [mm]; HS, heat-stable catalase.

7. After 5 days; v, variable.

8. PZA, pyrazinamidase activity; after 4 days.

9. ATCC, American Type Culture Collection.

10. MI, Mycobacterium International.
### TABLE 7 | Drug susceptibility profiles of *M. kansasii* strains under the study.

| Type | Strain no.       | MIC [mg/L] of a |
|------|-----------------|----------------|
|      | RIF  | CLR  | INH  | EMB  | STR  | AMK  | SXT  | RFB  | MXF  | LZD  | CIP  | DOX  | ETO  |
| M. kansasii | I (11) | | | | | | | | | | | | |
| 174_15 | ATCC25221 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 7728 | NLA001000521 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 7744 | NLA0010000927 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 6000 | NLA001000449 | 0.25 | 0.25 | 2 | >16 | 8 | 2 | 0.12 | <0.25 | <0.12 | 2 | 2 | 16 | <0.3 |
| K4 | NLA001000002 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| K14 | NLA0010001458 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| K19 | NLA0010001495 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MIC50/90 | NLA0010001495 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MIC range | NLA0010001495 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| M. ksanssii | II (7) | | | | | | | | | | | | |
| 2193 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 14_15 | B11069383 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 174_15 | B110401468 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| H7 | B11069383 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| H8 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MIC50/90 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MIC range | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| M. pseudokanssii | III (4) | | | | | | | | | | | | |
| 14_15 | NLA001001128 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 7744 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 6000 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MK142 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MK50/90 | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| MIC range | B11037207 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| M. ostrawleriense | IV (2) | | | | | | | | | | | | |
| 1010001458 | 241/15 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 241/15 | 250/90 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| M. innocens | V (4) | | | | | | | | | | | | |
| 1010001454 | 241/15 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 7728 | NLA001000002 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 7744 | NLA0010000927 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 6000 | NLA001000449 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| M. attenualatum | VI (2) | | | | | | | | | | | | |
| 7728 | MK142 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 7744 | MK142 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| 6000 | MK142 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |
| TOTAL | MK50/90 | 0.25 | 0.12 | 2 | >16 | 8 | 2 | <0.12 | <0.25 | <0.12 | 2 | 2 | 4 | <0.3 |

Please note, that for some isolates those data were reported elsewhere (Blakou et al., 2018).

aAccording to the WGS-based (Msi) method grouping. Vertically are given the newly proposed names for each of the *M. kansasii* subtypes.

bRIF, rifampicin; CLR, clarithromycin; INH, isoniazid; EMB, ethambutol; STR, streptomycin; AMK, amikacin; SXT, co-trimoxazole; RFB, rifabutin; MXF, moxifloxacin; LZD, linezolid; CIP, ciprofloxacin; DOX, doxycycline; ETO, ethionamide.
Not only does it further substantiate the delineation of new species from the \textit{M. kansasii} group to replace the former subtypes I–VI, but consolidates the position of five of the so erected species, and provides a description of the sixth one, \textit{M. ostraviense}, a successor of the subtype IV. By showing a close genetic relatedness, a monophyletic origin, and overlapping phenotypes, our findings support the recognition of the \textit{M. kansasii} complex (MKC), accommodating all \textit{M. kansasii}-derived species and \textit{M. gastri}. Neither of the most commonly used taxonomic markers can accurately distinguish all the MKC species. Likewise, no species-specific phenotypic characteristics exist that would allow identification of the species, except the non-photochromogenicity of \textit{M. gastri}. In the context of the previously proposed polyphasic strategy in resolving species boundaries and their interrelatedness, FAME (fatty acid methyl ester) analysis, as an adjunct typing method, might be useful (Saini et al., 2009). However, preparatory techniques for FAME analysis are typically time-consuming, laborious, and material-intensive. Furthermore, chromatography-dedicated facilities require investment in instrumentation and training, and despite their services being offered by universities and other centres, they are often less accessible than sequencing facilities, even in the developing countries.

To distinguish, most reliably, between the MKC species, and between \textit{M. kansasii} and \textit{M. persicum} in particular, whole-genome-based approaches should be applied.

Since no clear differences in the repertoire of the virulence-associated RD1 genes have been observed among the \textit{M. kansasii}-derived species, the pathogenic capacity of each of these species can only be speculated based on their prevalence among the clinically relevant population. Large-scale molecular epidemiological studies are needed to gain a better understanding of the clinical significance and pathobiology of the MKC species.

### Description of \textit{Mycobacterium ostraviense} sp. nov. Jagielski and Ulmann

\textit{Mycobacterium ostraviense} [os.travi.en’se. N.L. neut. adj. ostraviense pertaining to Ostravia, the Latin name of Ostrava, a city in the north-east of the Czech Republic where one of the strain was isolated].

The species name refers to the former \textit{M. kansasii} subtype IV. Mature colonies, of rough surface and photochromogenic, are observed on Löwenstein-Jensen medium after more than 7 days of incubation at 37°C (Supplementary Figure 1). No growth occurs at 45°C and on the media containing TCH or 5% (w/v) NaCl. Similar to other MKC species, it tests positive for the semi-quantitative catalase (>45 mm) and 14-day arylsulfatase activity, while negative for niacin accumulation, tellurite reduction, and urease and pyrazinamidase activities. Unlike to other \textit{M. kansasii}-derived species but similar to \textit{M. gastri}, it does not reduce nitrates, nor it produces a heat-stable (68°C) catalase. Strongly resistant to ethambutol (>16 mg/L) but susceptible to amikacin, clarithromycin, \textit{co-trimoxazole}, linezolid, fluoroquinolones, and rifamycins.

The species has the same 16S rRNA sequences as \textit{M. kansasii} (former subtype I) and \textit{M. gastri}, yet it displays unique sequences at the \textit{hsp65}, \textit{tuf}, and \textit{rpoB} genes, and the ITS locus. At the genomic level, it is most closely related to \textit{M. gastri}, with pairwise ANI and GGD values of 95.2 and 0.056, respectively.

The type strain, 241/15\textsuperscript{T} was isolated from a sputum of a patient with no NTM disease, based in Karviná, near Ostrava, in the Moravian-Silesian Region of the Czech Republic. The type strain has been deposited in the Leibniz Institute German Collection of Microorganisms and Cell Cultures (DSMZ; Braunschweig, Germany) under the accession number DSM 110538.

### DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/Supplementary Material.

### ETHICS STATEMENT

Analyses were based on data that did not contain any sensitive personal information. Therefore, informed consent and ethical approval were not required in like with local legislation.

### AUTHOR CONTRIBUTIONS

TJ conceptualized and supervised the study, provided the funding, organized and integrated the data, and wrote the manuscript. PB, JL, and DS performed bioinformatic analyses including AF-ANI, GGD, and phylogenetic tree analysis. ZB performed culturing, subtyping, and phenotypic profiling of \textit{M. kansasii} strains. BM carried out analysis on regions of difference 1-14 (RD1-14) with a homemade script Diffind. AB carried out DNA isolations for whole-genome sequencing. JD analyzed the results on the distribution of the RD1-14 genes in \textit{M. kansasii} genomes. MD constructed the core-genome phylogenies. LP performed drug susceptibility testing. JL provided 13 \textit{M. kansasii} strains of subtypes I-V1 and critically reviewed the manuscript. MZ-D co-performed phenotypic assays.

### FUNDING

This work was supported by the National Centre for Research and Development grant (to: TJ) through the LIDER Programme under contract no. LIDER/044/457/L-4/12/NCBR/2013. DS, JL, and BM were supported by the Polish Ministry of Science and Higher Education grant under contract no. DIR/WK/2017/2018/01-1.

### ACKNOWLEDGMENTS

The authors are indebted to Aleksandra Safianowska (Department of Internal Medicine, Pulmonology, and Allergology, Warsaw Medical University, Warsaw, Poland), Won-Jung Kob (Division of Pulmonary and Critical Care Medicine; Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, South Korea), Vit Ulmann (Institute of Public Health, Ostrava, Czech Republic), Miren J. Unzaga Barañano (Clinical Microbiology and Infection Control, Basurto Hospital, Bilbao, Spain), Joanna Humiecka (Hospital of Infectious Diseases in Warsaw, Poland),...
and Katharina Kranzer (National Reference Laboratory for Mycobacteria, Research Center Borstel, Borstel, Germany) for their kind providing some of M. kansasii strains used in this study.

Special appreciation is extended to Fernando Alcaide and David Rodriguez-Temporal (Department of Pathology and Experimental Therapy, Universitat de Barcelona, Hospital de Llobregat, Spain) for their kind reviewing of the manuscript and valuable comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.02918/full#supplementary-material

Supplementary Tables 1–6 are published inside Supplementary Data Sheet 1.

REFERENCES

Abed, Y., Bollet, C., and de Micco, P. (1995). Demonstration of Mycobacterium kansasii species heterogeneity by the amplification of the 16S-23S spacer region. J. Med. Microbiol. 42, 112–114. doi: 10.1099/00222105-4-3-156

Ahmed, N., Saini, V., Raghuvanshi, S., Khurana, J. P., Tyagi, A. K., Tyagi, A. K., et al. (2007). Molecular analysis of a leprosy immunotherapeutic bacillus provides insights into Mycobacterium evolution. PLoS ONE 2:e968. doi: 10.1371/journal.pone.0000968

Alcaide, F., Richter, I., Bernasconi, C., Springer, B., Hagenau, C., Schulze-Jagielski, L., et al. (2019). Genomic Insights Into Mycobacterium kansasii. Genomic Insights Into Mycobacterium kansasii Complex

Auch, A. F., Kleink, H. P., and Göker, M. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Bakula, Z., Modrzejewska, M., Safianowska, A., Proboszcz, M., Bielecki, J., van Ingen, J., et al. (2018b). Clinical, radiological and molecular features of clinical isolates of Mycobacterium kansasii. J. Med. Microbiol. 54, 975–979. doi: 10.1099/jmm.0.65934-0

Bakula, Z., Modrzejewska, M., Safianowska, A., Proboszcz, M., Bielecki, J., van Ingen, J., et al. (2018a). Molecular typing of Mycobacterium kansasii using pulsed-field gel electrophoresis and a newly designed variable-number tandem repeat analysis. Sci. Rep. 8:4662. doi: 10.1038/s41598-018-21562-x

Bakula, Z., Kosciuch, J., Safianowska, A., Proboszcz, M., Bielecki, J., van Ingen, J., et al. (2018b). Clinical, radiological and molecular features of Mycobacterium kansasii pulmonary disease. Respir. Med. 139, 91–100.

doi:10.1016/j.rmed.2018.05.007

Bakula, Z., Modrzejewska, M., Pennings, L., Proboszcz, M., Safianowska, A., Bielecki, J., et al. (2018c). Drug susceptibility profiling and genetic determinants of drug resistance in Mycobacterium kansasii. Antimicrob Agents Chemother. C62, e01788-e01717. doi: 10.1128/AAC.01788-17

Bakula, Z., Modrzejewska, M., Safianowska, A., van Ingen, J., Proboszcz, M., Bielecki, J., et al. (2016). Proposal of a new method for subtyping of Mycobacterium kansasii based upon PCR restriction enzyme analysis of the tuf gene. Diagn. Microbiol. Infect. Dis. 84, 318–321. doi: 10.1016/j.diagmicrobio.2015.12.009

Berthelot, F. X., Rasmussen, P. B., Rosenkranz, L., Andersen, P., and Gicquel, B. (1998). A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). Microbiology 144, 3195–3203. doi: 10.1099/00221287-144-11-3195

Brodin, P., Majlessi, L., Marsollier, L., de Jonge, M. I., Bottai, D., Demangel, C., et al. (2006). Dissection of ESAT-6 system 1 of Mycobacterium tuberculosis and impact on immunogenicity and virulence. Infect. Immun. 74, 88–98. doi: 10.1128/IAI.74.1.88-98.2006

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., et al. (2002). A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proc. Natl. Acad. Sci. U.S.A. 99, 3684–3689. doi: 10.1073/pnas.052548299
length polymorphism technique for fingerprinting and differentiating of *Mycobacterium kansasii* type I isolates. *J. Clin. Microbiol.* 41, 3846–3850. doi: 10.1128/JCM.41.8.3846-3850.2003

G Griffith, D. E., Aksamit, T. A., Catanzaro, A., Daley, C., Gordin, F., et al. (2007). ATS Mycobacterial Diseases Subcommittee; American Thoracic Society; Infectious Disease Society of America. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am. J. Respir. Crit. Care Med.* 175, 367–416. doi: 10.1164/rccm.200604-0471ST

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Hall, T. A. (1999). BoEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Hauduroy, P. (1955). *Derniers Aspects du monde des Mycobactéries*. Paris: Masson et Cie, 73.

Hermans, P. W., van Soolingen, D., and van Embden, J. D. (1992). Différents Aspects du monde des Mycobactéries. *Hermans, P. W., van Soolingen, D., and van Embden, J. D.* (1992).

Inwald, J., Jahans, K., Hewinson, R. G., and Gordon, S. V. (2003). Inactivation of the *Mycobacterium bovis* bovis homologue of the polymorphic RD1 gene *rV3879c* (Mb3909c) does not affect virulence. *Cell. Microbiol.* 14, 1287–1298. doi: 10.1111/j.1462-5882.2012.01799.x

Iinuma, Y., Ichiyama, S., Hasegawa, Y., Shimokata, K., Kawahara, S., and Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kent, P. T., and Kubica, G. P. (1985). *Mammalian Protein Metabolism*, ed H. N. Munro (New York, NY: Academic Press), 21–132.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kent, P. T., and Kubica, G. P. (1985). *Mammalian Protein Metabolism*, ed H. N. Munro (New York, NY: Academic Press), 21–132.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kent, P. T., and Kubica, G. P. (1985). *Mammalian Protein Metabolism*, ed H. N. Munro (New York, NY: Academic Press), 21–132.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kent, P. T., and Kubica, G. P. (1985). *Mammalian Protein Metabolism*, ed H. N. Munro (New York, NY: Academic Press), 21–132.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kent, P. T., and Kubica, G. P. (1985). *Mammalian Protein Metabolism*, ed H. N. Munro (New York, NY: Academic Press), 21–132.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kent, P. T., and Kubica, G. P. (1985). *Mammalian Protein Metabolism*, ed H. N. Munro (New York, NY: Academic Press), 21–132.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
Ross, B. C., Jackson, K., Yang, M., Sievers, A., and Dwyer, B. (1992). Identification of a genetically distinct subspecies of *Mycobacterium kansasii*. *J. Clin. Microbiol.* 30, 2930–2933.

Saini, V., Raghuvanshi, S., Talwar, G. P., Ahmed, N., Khurana, J. P., Hasnain, S. E., et al. (2009). Polyphasic taxonomic analysis establishes *Mycobacterium indicus pranii* as a distinct species. *PLoS ONE* 4:e6263. doi: 10.1371/journal.pone.0006263

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Santin, M., and Alcaide, F. (2003). *Mycobacterium kansasii* disease among patients infected with human immunodeficiency virus type 1: improved prognosis in the era of highly active antiretroviral therapy. *Int. J. Tuberc. Lung Dis.* 7, 673–677.

Santos, A. R., De Miranda, A. B., Lima, L. M., Suffys, P. N., and Jagielski et al. Genomic Insights Into... (2015). Microbial species delineation using whole genome sequences. *Nucleic Acids Res.* 43, 6761–7671. doi: 10.1093/nar/gkv657

Sood, G., and Parrish, N. (2017). Outbreaks of nontuberculous mycobacteria. *Curr. Opin. Infect. Dis.* 30, 404–409. doi: 10.1097/QCD.0000000000000386

Tagini, F., Aeby, S., Bertelli, C., Droz, S., Casanova, C., Prod’Hom, G., et al. (2019). Phylogenomics reveal that *Mycobacterium kansasii* subtypes are species-level lineages. Description of *Mycobacterium pseudokansasii* sp. nov., *Mycobacterium innocens* sp. nov. and *Mycobacterium attenuatum* sp. nov. *Int. J. Syst. Evol. Microbiol.* 69, 1766–1770. doi: 10.1099/ijsem.0.001862

Tortoli, E., Fedrizzi, T., Meehan, C. J., Trovato, A., Grottola, A., Giacobazzi, E., et al. (2017). The new phylogeny of the genus *Mycobacterium*: the old and the news. *Infect. Genet. Evol.* 56, 19–25. doi: 10.1016/j.meegid.2017.10.013

Tortoli, E., Simonetti, M. T., Lacchini, C., Penati, V., and Urbano, P. (1994). Tentative evidence of AIDS-associated biotype of *Mycobacterium kansasii*. *J. Clin. Microbiol.* 32, 1779–1782.

Vaidya, G., Lohman, D. J., and Meier, R. (2011). SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27, 171–180. doi: 10.1111/j.1096-0031.2010.00329.x

van der Wel, N., Hava, D., Houben, D., Fliuista, D., van Zon, M., Pierson, J., et al. (2007). *M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* 129, 1287–1298. doi: 10.1016/j.cell.2007.05.059

van Ingen, J., de Zwaan, R., Dekhuijzen, R., Boeree, M., and van Sooeling, D. (2009). Region of difference 1 in nontuberculous *Mycobacterium* species adds a phylogenetic and taxonomic character. *J. Bacteriol.* 191, 5865–5867. doi: 10.1128/JB.00683-09

Varghese, N. J., Mukherjee, S., Ivanova, N., Konstantinidis, K. T., Mavrommatis, K., Kyrpides, N. C., et al. (2015). Microbial species delineation using whole genome sequences. *Nucelic Acids Res.* 43, 6761–6771. doi: 10.1093/nar/gkv657

Velayati, A. A., Booloosae, M. R., Farnia, P., Mohammad, F., Karam, M. B., Masjedi, M. R., et al. (2005). *Mycobacterium gastri* causing disseminated infection in children of same family. *Pediatr. Pulmonol.* 39, 284–287. doi: 10.1002/ppul.20186

Volkman, H. E., Pozos, T. C., Zheng, J., Davis, J. M., Rawls, J. F., and Ramakrishnan, L. (2010). Tuberculous granuloma induction via interaction of a bacterial secreted protein with host epithelium. *Science* 327, 466–469. doi: 10.1126/science.1179663

Witig, R. S., Fazal, B. A., Meza, R. M., Mushatt, D. M., Dejace, P. M., Greer, D. L., et al. (1995). Clinical manifestations and implications of coinfection with *Mycobacterium kansasii* and human immunodeficiency virus type 1. *Clin. Infect. Dis.* 21, 77–85. doi: 10.1093/clinids/21.1.77

Wu, T. S., Leu, H. C., Chu, C. H., Lee, M. H., Chiang, P. C., Wu, T. L., et al. (2009). Clinical manifestations, antibiotic susceptibility and molecular analysis of *Mycobacterium kansasii* isolates from a university hospital in Taiwan. *J. Antimicrob. Chemother.* 64, 511–514. doi: 10.1093/jac/dkp238

Yang, M., Ross, B. C., and Dwyer, B. (1993). Identification of an insertion sequence-like element in a subspecies of *Mycobacterium kansasii*. *J. Clin. Microbiol.* 31, 2074–2079.

Yoon, S. H., Ha, S. M., Lim, J. M., Kwon, S. J., and Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek*. 110, 1281–1286. doi: 10.1007/s10482-017-0844-4

Zhang, M., Chen, J. M., Sala, C., Rybniker, J., Dhar, N., and Cole, S. T. (2014). Espl regulates the ESX-1 secretion system in response to ATP levels in *Mycobacterium tuberculosis*. *Mol. Microbiol.* 93, 1057–1065. doi: 10.1111/mmi.12718

Zhang, Y., Mann, L. B., Wilson, R. W., Brwon-Elliott, B. A., Vincent, V., Linuma, Y., et al. (2004). Molecular analysis of *Mycobacterium kansasii* isolates from the United States. *J. Clin. Microbiol.* 42, 119–125. doi: 10.1128/JCM.42.1.119-125.2004

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Jagielski, Borowski, Bakula, Lach, Marciniak, Brzostek, Dziszynski, Penning, van Ingen, Zolnir-Dov et. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.