International Journal of Antimicrobial Agents 58 (2021) 106401

Contents lists available at ScienceDirect

International Journal of Antimicrobial Agents

journal homepage: www.elsevier.com/locate/ijantimicag

Genomic-based surveillance reveals high ongoing transmission of multi-drug-resistant *Mycobacterium tuberculosis* in Southern Brazil

Richard Steiner Salvato[a,b,*], Ana Júlia Reis[c,†], Sun Hee Schiefelbein[d,†], Michael Andrés Abril Gómez[c], Stéphanie Steiner Salvato[b], Larissa Vitória da Silva[b], Elis Regina Dalla Costa[e], Gisela Unis[f], Claudia Fontoura Dias[f], Miguel Viveiros[g], Isabel Portugal[h], Andrea von Groll[c], Pedro Eduardo Almeida da Silva[c], Afrânio Lineu Kritski[e,‡], João Perdigão[h,‡], Maria Lucia Rosa Rossetti[a,d,‡]

[a] *Programa de Pós-graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil*
[b] *Centro de Desenvolvimento Científico e Tecnológico, Centro Estadual de Vigilância em Saúde, Secretaria Estadual da Saúde do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil*
[c] *Medical Microbiology Research Center, Faculdade de Medicina, Programa de Pós-graduação em Ciências da Saúde, Universidade Federal do Rio Grande, Rio Grande, Rio Grande do Sul, Brazil*
[d] *Programa de Pós-graduação em Biologia Celular e Molecular Aplicada à Saúde, Universidade Luterana do Brasil, Canoas, Rio Grande do Sul, Brazil*
[e] *Programa Acadêmico de Tuberculose, Faculdade de Medicina e complexo hospitalar HUCFF-IDT, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*
[f] *Hospital Sanatório Partenon, Porto Alegre, Rio Grande do Sul, Brazil*
[g] *Unidade de Microbiologia Médica, Global Health and Tropical Medicine, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisboa, Portugal*
[h] *iMed.ULisboa – Research Institute for Medicine, Faculdade de Farmácia, Universidade de Lisboa, Lisboa, Portugal*

ARTICLE INFO

Article history:
Received 8 April 2021
Accepted 3 July 2021

Editor: Jean-Marc Rolain

Keywords:
Tuberculosis
Drug resistance
Transmission
Whole-genome sequencing

ABSTRACT

Genomic-based surveillance on the occurrence of drug resistance and its transmission dynamics has emerged as a powerful tool for the control of tuberculosis (TB). A whole-genome sequencing approach, phenotypic testing and clinical-epidemiological investigation were used to undertake a retrospective population-based study on drug-resistant (DR)-TB in Rio Grande do Sul, the largest state in Southern Brazil. The analysis included 305 resistant *Mycobacterium tuberculosis* strains sampled statewide from 2011 to 2014, and covered 75.7% of all DR-TB cases identified in this period. Lineage 4 was found to be predominant (99.3%), with high sublineage-level diversity composed mainly of 4.3.4.2 [Latin American and Mediterranean (LAM)/RD174], 4.3.3 (LAM/RD115) and 4.1.2.1 (Haarlem/RD182) sublineages. Genomic diversity was also reflected in resistance of the variants to first-line drugs. A large number of distinct resistance-conferring mutations, including variants that have not been reported previously in any other setting worldwide, and 22 isoniazid-monoresistant strains with mutations described as disputed in the *rpoB* gene but causing rifampicin resistance generally missed by automated phenotypic tests as BACTEC MGIT. Using a cut-off of five single nucleotide polymorphisms, the estimated recent transmission rate was 55.1%, with 168 strains grouped into 28 genomic clusters. The most worrying fact concerns multi-drug-resistant (MDR) strains, of which 73.4% were clustered. Different resistance profiles and acquisition of novel mutations intraclusters revealed important amplification of resistance in the region. This study described the diversity of *M. tuberculosis* strains, the basis of drug resistance, and ongoing transmission dynamics across the largest state in Southern Brazil, stressing the urgent need for MDR-TB transmission control state-wide.

* Corresponding author. Address: Centro Estadual de Vigilância em Saúde, Secretaria Estadual da Saúde do Rio Grande do Sul, Av. Ipiranga 5400, CEP 90610-000, Porto Alegre, Rio Grande do Sul, Brazil. Tel.: +55 51 985767035.
E-mail address: richardsalvato@hotmail.com (R.S. Salvato).
† These authors contributed equally.
‡ These authors supervised this work equally.

https://doi.org/10.1016/j.ijantimicag.2021.106401
0924-8579/

1. Introduction

Resistance to anti-tuberculosis (TB) drugs is one of the main reasons why TB is still one of the leading infectious diseases. In 2019, TB accounted for 1.4 million deaths and 10 million new cases worldwide. Drug resistance increases the standard anti-TB treatment length to at least 1 year, and requires the use of different classes of drugs which are more expensive, less effective and have a lower cure rate: 26% for extensively drug-resistant (XDR) cases [1].

Annually, approximately 73,000 new cases of TB and 4500 deaths due to TB are notified in Brazil, resulting in an incidence rate of 35 cases/100,000 population (2019). From these cases, an important fraction is classified as multi-drug resistant (MDR)/rifampicin (RIF)-resistant: 1119 laboratory-confirmed MDR-TB/RIF-resistant cases were notified in 2017. Nevertheless, the distribution of the disease is heterogeneous across the country, and Rio Grande do Sul, the southernmost Brazilian state and the largest of the three states that compose the south region of Brazil, accounts for 54% of the cases of TB occurring in this region with an incidence rate of 40 cases/100,000 population [2]. Furthermore, Rio Grande do Sul has a high number of cases of MDR/RIF-resistant TB (109 cases in 2017) according to the data from the State Central Laboratory (not shown).

Previous studies on circulating *Mycobacterium tuberculosis* in the south region of Brazil have shown predominance of Lineage 4 strains, but with important diversity at the sublineage level with predominance of 4.3.3 [Latin American and Mediterranean (LAM)], 4.3.4.2 (LAM) and 4.1.2.1 (Haarlem) sublineages [3,4]. Moreover, strains rarely found in other regions, such as SIT863 [5], have been observed in recent studies, coupled with ongoing transmission of highly resistant strains [4,6]. However, previous studies in the region only undertook limited sampling, mainly from MDR strains. To obtain a broader understanding of the DR-TB scenario and aiming to inform DR-TB control efforts in the region, a representative sample of DR *M. tuberculosis* strains circulating in Rio Grande do Sul state between 2011 and 2014 was characterized in terms of their diversity and genomic similarity to estimate ongoing transmission of *M. tuberculosis*.

2. Materials and methods

2.1. Study population

A retrospective population-based study on DR-TB was conducted in Rio Grande do Sul state, Southern Brazil, between 2011 and 2014. Available *M. tuberculosis* clinical strains were collected at the State Central Laboratory (LACEN-RS). LACEN-RS is the reference laboratory in charge of drug susceptibility testing (DST) from TB cases notified state-wide. The study included 305 clinical strains of *M. tuberculosis* that presented resistance to at least one of the following first-line anti-TB drugs: isoniazid (INH), RIF, ethambutol (EMB) and streptomycin (STR).

Rio Grande do Sul, the largest state in Southern Brazil, has an estimated population of 11.3 million people [7], and ranks among the high-burden TB states in Brazil with a high rate of MDR-TB. In 2017, 5031 new TB cases [2] (40 cases/100,000 population) and 90 cases of diagnosed MDR (data from LACEN-RS) were notified from Rio Grande do Sul. Over the 4-year study period, 403 DR-TB samples were identified at LACEN-RS. The study sample includes 305 of these 403 notified DR-TB samples (75.7%). Clinical and epidemiological data from enrolled individuals were obtained from the Brazilian National System for Notifiable Diseases (SINAN) and the Brazilian Special Tuberculosis Treatment Information System (SITE-TB). A confirmed epidemiological link (epi-link) was considered when two individuals lived in the same neighbourhood or spent time at the same prison for any time in the same year.

2.2. Drug susceptibility testing

DST results were obtained for clinical samples from LACEN-RS, where testing was performed using a liquid BACTEC MGIT 960 SIRE Kit for the BACTEC Mycobacteria Growth Indicator Tube 960 (MGIT 960) system (Becton Dickinson Diagnostic Systems, Sparks, MD, USA). Susceptibility evaluation was conducted for the following first-line anti-TB drugs: RIF (breakpoint 1.0 mg/L), INH (breakpoint 0.1 mg/L), EMB (breakpoint 5.0 mg/L) and STR (breakpoint 1.0 mg/L) in accordance with the manufacturer's instructions.

2.3. DNA extraction and whole-genome sequencing

M. tuberculosis genomic DNA was extracted using the cetyltrimethylammonium bromide method, as described by van Embden et al. [8] in a Biosafety Level (BSL) 2 laboratory with BSL-3 safety equipment and work practices. The genomic DNA from the 305 studied *M. tuberculosis* strains was subjected to next-generation sequencing to access its whole-genome sequence. Paired-end sequencing (2 × 150 bp) was performed on an Illumina NextSeq machine using either a 300 cycle v2 mid-output or high-output kit (Illumina, Code FC-404-2003 or Code FC-404-2004), using the standard Illumina procedure as described previously [4].

2.4. Bioinformatic analysis of whole-genome sequencing data

Raw FASTQ files were trimmed to remove adapter sequences and low-quality reads using Trimmomatic [9]. The read data quality was assessed in fastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/), and then mapped against the reference genome of *M. tuberculosis* H37Rv (GenBank Accession NC000962.3) using the BWA-MEM algorithm [10]. The quality of the resulting mapped BAM files was checked using Qualimap [11]. SAMtools/BCFtools and GATK tools were used for variant calling of single nucleotide polymorphisms (SNPs) and small indels, as described previously [12,13]. Both variant sets called by each tool were combined, and only the concordant set between both callers was retained for downstream analysis. SNP sites with an excess of 10% missing calls were removed from the analysis [14], as well as SNP positions within PE/PPE genes. Variants occurring in drug-resistance-associated genes were retained for increased resolution at the micro-evolutionary level. Retained variants were converted to a FASTA format file and then used to generate an alignment composed of 28,974 high-quality SNP sites.

The final [whole-genome SNP (wgSNP)] alignment was submitted to the JmodelTest tool [15] for best-fit nucleotide substitution model selection under Akaike Information Criterion. Generalized Time-Reversible model with estimated proportion of invariable sites was selected to reconstruct a maximum-likelihood phylogenetic tree using the RAxML tool [16], applying the bootstrap branch support metric, and the resulting tree was annotated in the Interactive Tree of Life online tool [17]. A cut-off of five SNPs was used to delineate genomic clusters [18] among the wgSNP alignment using the *ape* package and *hclust* function implemented in R. Recent transmission among *M. tuberculosis* genomes was estimated from the ratio between the number of clustered strains and the total number of included strains. To analyse the geographical distribution of DR-TB cases, patient addresses were plotted in a map using the online tool Microreact [19]. Pairwise geographical distances between patients were determined using the *Imap* package implemented in R, and the Mann-Whitney *U*-test was applied.

The *fbp*C103 polymorphism (G->A at codon 103) used to differentiate LAM strains from non-LAM strains and genomic variants

underlying drug resistance were identified from VCF files cited previously. The command-line version of TB-Profiler (v2.8.13) [20] was used to determine the *M. tuberculosis* SNP-based type, and SpoTyping (v2.0) [21] was used for prediction of in-silico spoligotypes. The clade and shared international type (SIT) of each spoligotyping pattern were assigned using the SITVIT2 online database [22].

2.5. Resolving discordant phenotypic and whole-genome sequencing results

Discordant results were resolved by repeating the DST on the MGIT 960 system and performing minimum inhibitory concentration (MIC) analyses with the resazurin microtitre assay method, as described previously [23,24], using the breakpoint concentrations of ≤0.25mg/L and ≤0.5 mg/L for INH and RIF, respectively.

3. Results

3.1. Study population and clinical characteristics

From 2011 to 2014, 403 cases of DR-TB were detected statewide and further classified as MDR (n=240), resistant to INH alone (n=122), resistant to RIF alone (n=6), resistant to EMB alone (n=3), resistant to STR alone (n=3), and polyresistant (n=32). Of these 403 DR-TB notified cases, 305 *M. tuberculosis* resistant strains from unique patients were available (75.7% of total from study period) with the following DR profiles: 169 MDR, 111 resistant to INH alone, 19 polyresistant, five resistant to RIF alone, and one resistant to STR alone.

Among the 305 included patients, 76.1% were male and the median age was 40 years (range 13–84 years). For 34.4% of the patients, it was their first TB diagnosis (primary resistance), 29.8% were starting a new treatment period for a TB relapse, 25.3% were starting treatment following loss to follow-up, and 9.2% were starting treatment after failure of a previous treatment. Information was not available for four individuals (1.3%) (according to data obtained from SINAN and SITE-TB). Regarding the main comorbidities and risk factors, 28.5% were infected with human immunodeficiency virus; 8.9% had diabetes mellitus; and 33.1%, 19.7% and 18.4% were alcohol, tobacco and illicit drug users, respectively. In addition, 37 (12.13%) individuals were prison inmates at the time of sample collection. Information related to treatment outcome was obtainable for 254 (83.3%) cases. A favourable outcome (cure or treatment completion) was reported for 49.9% (104/254) of those cases; 35.8% (91/254) of the individuals were lost to follow-up, the current treatment failed for 6.7% (17/254) of the individuals, and 16.5% (42/254) of the individuals died during treatment (see online supplementary material for detailed characteristics of individuals included).

3.2. Drug resistance and associated mutations

From the 111 INH-monoresistant clinical strains, 95 (85.6%) carried a well-described mutation related to resistance. Among the 169 phenotypically MDR strains, only one did not have any known mutations underlying INH and RIF resistance. For 43 samples that presented discordant results for INH and RIF resistance between MGIT testing and the whole-genome sequencing (WGS) prediction, MGIT- and MIC-based DST were performed for the two drugs. This revealed that 15 INH-monoresistant strains on MGIT-SIRE did not have associated resistance mutations, and were, in fact, susceptible. Therefore, for the purposes of evaluation of WGS-based drug-resistance prediction, the susceptibility profile was changed. For the remaining strains, the phenotypic resistance profile remained the same.

Additionally, well-described disputed mutations in the *rpoB* gene causing RIF resistance, generally missed by automated phenotypic DST methods [25], were identified in 22 (22.9%) of the 96 remaining INH-monoresistant strains: D435Y (8/22), L452P (5/22), H445Y (4/22), L430P (2/22), H445G (1/22), H445N (1/22) and D435Y+S431T (1/22). Thus, for the purpose of evaluation of WGS-based drug-resistance prediction, those strains were considered to be RIF resistant based on previous reports of the phenotypic resistance caused by these mutations, and the inability of MGIT to detect it [25]. Concerning clinical evolution for these 22 patients, 17 of them received MDR-TB regimens after a review of their clinical status (non-response or treatment failure to first-line drugs, including RIF). For the remaining five patients, the treatment information was not available, but two of them died from TB. Finally, the overall concordance on WGS-based resistance prediction was 99.3% for INH, 98.7% for RIF, 82.3% for EMB and 86.9% for STR, considering MGIT-based testing as the gold standard (Table 1).

The most common variants sustaining INH resistance were found in the *katG* gene, with the distinct occurrence of some mutations among different resistance profiles (Ser315Thr mainly in MDR and Ser315Asn mainly in INH-monoresistant). In total, 84.5% (240/284) of INH-resistant strains had some related resistance mutation at *katG*, and among these, 53 (18.7%) had additional -15 C>T or T>C -8 changes at the *inhA* promoter. Thirty of 284 (10.6%) INH-resistant strains only carried the *inhA*-promoter -15 C>T mutation, 11/284 (3.9%) only had mutations in the *oxyR-ahpC* regulatory region, and one had the Ser94Ala codon change in *inhA* gene (Figure 1). Regarding RIF resistance, 173 (99.4%) of the 174 RIF-resistant isolates carried mutations in the *rpoB* gene. Amino acid changes in codon 450 were the most frequent, occurring in 73.6% of the resistant isolates. An uncommon insertion (*rpoB* 435 QNNP > QNNPQNNP) [6] was identified in 16/174 isolates (9.2%), and two isolates (1.1%) had deletion of two amino acids (*rpoB* 435 FMD>F) (Figure 1). Disputed mutations occurring in the *rpoB* gene were found in 22 RIF-susceptible (according to MGIT) strains: 14 of them were INH-monoresistant strains, seven were resistant to INH and STR, and one was susceptible to first-line drugs.

Mutations associated with EMB resistance were found in 91.7% (11/12) of the resistant isolates, all carried on the Met306Val amino acid change in the *embB* gene. However, 18.1% (53/293) of the EMB-susceptible strains also had resistance-associated mutations. Among 47 STR-resistant strains, 28 (59.6%) had related resistance mutations identified on the *rrs* (11/47), *gidB* (11/47) and *rpsL* (7/47) genes (Figure 1). Second-line drug-resistance-associated variants were found in nine strains, all of which carried amino acid substitutions at codon 94 on the *gyrA* gene (Asp94His, Asp94Gly, Asp94Asn, Asp94Tyr and Asp94Ala), related to fluoroquinolone resistance. One of these nine strains also had a mutation on the *rrs* gene (1401 A>G), which is a known marker of aminoglycoside resistance (see online supplementary material for complete resistance-associated variant profiles). For one of these nine patients with pre/XDR-TB, the outcome was cure following a change in the anti-TB regimen. In the other eight individuals, an unfavourable outcome was seen: four had treatment failure, three were lost to follow-up (including the one case of XDR-TB), and one died.

3.3. Genetic diversity of M. tuberculosis

The two typing methods implemented – in-silico spoligotyping and SNP-based typing – assigned 303 (99.3%) clinical strains to *M. tuberculosis* Lineage 4, one strain to Lineage 1 (EAI5) and one strain to *M. bovis* (BOV1) species. SNP-based classification showed predominance of the 4.3.4.2 (LAM) sublineage in 63 (20.7%) strains, 4.3.3 (LAM) in 61 (20%) strains, and 4.1.2.1 (Haarlem) in 54 (17.7%)

Table 1
Agreement between phenotypic testing and genome-based drug resistance prediction with sensitivity and specificity values for whole-genome sequencing (WGS)-based resistance detection.

Drug	DST	Total	WGS (R)	WGS (S)	Sensitivity	Specificity	PVP	PVN
INH	R	284	282	2[a]	0.99	1.00	1.00	0.91
	S	21	0	21				
RIF	R	174	172	1[a]	0.99	0.97	0.98	0.99
	S	131	25[b]	106				
EMB	R	12	11	1	0.92	0.82	0.17	1.00
	S	293	53	240				
STR	R	47	28	19	0.60	0.92	0.57	0.93
	S	258	21	237				

DST, drug-susceptibility testing; R, resistant; S, susceptible; INH, isoniazid; RIF, rifampicin; EMB, ethambutol; STR, streptomycin; PVP, predictive value for a positive test; PVN, predictive value for a negative test.

[a]Mycobacteria Growth Indicator Tube (MGIT) and minimum inhibitory concentration (MIC) determination were performed to confirm phenotypic resistance.

[b] Disputed mutations in the *rpoB* gene were identified in 22 strains. Due to the well-established basis of these mutations, RIF-susceptible samples on MGIT presenting disputed mutations were considered to be phenotypically resistant. The remaining three RIF-susceptible samples on MGIT, carried on the *rpoB* gene and highly related to resistance variants, were submitted to MGIT to confirm phenotypic susceptibility.

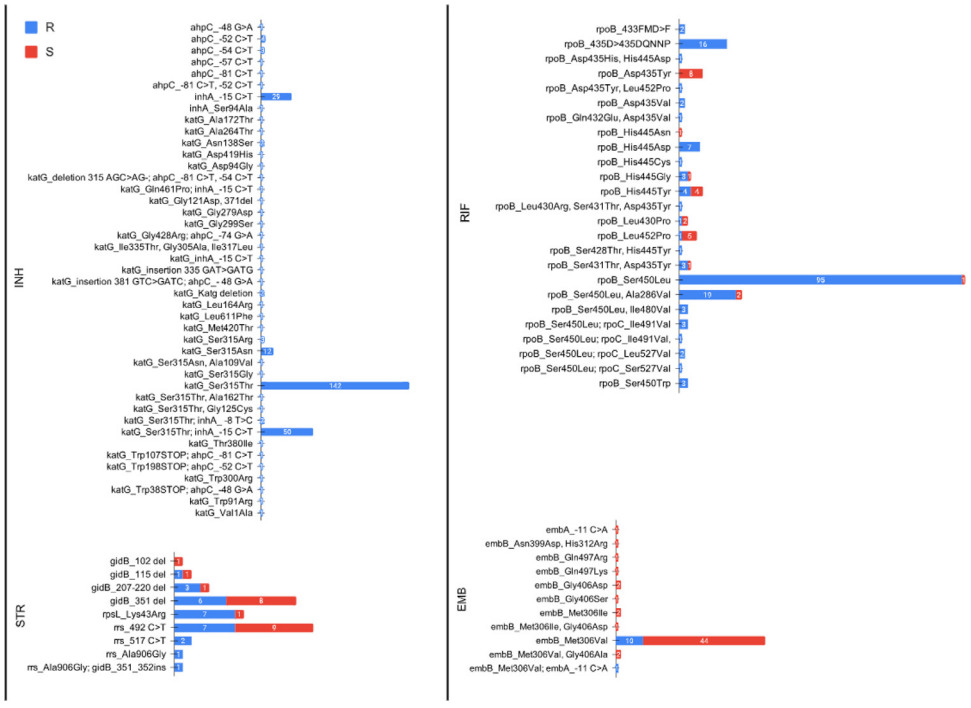

Figure 1. Mutation patterns underlying first-line drug resistance among the 305 drug-resistant *Mycobacterium tuberculosis* strains.

strains. All 188 strains (61.6%) assigned to the 4.3 (LAM) sublineage carried the *fbpC*[103] polymorphism, a known LAM SNP marker; RD174 was found in 64 (34%) of these LAM strains and RD115 was found in 55 (29.3%). In-silico spoligotyping revealed 62 distinct patterns and 48 different SITs among the 305 clinical strains. The most common SITs were: 53/T1 in 40 (13.1%) strains, 93/LAM5 in 26 (8.5%) strains, 65/T1 in 24 (7.9%) strains and 42/LAM9 in 22 (7.2%) strains. The SITVIT2 database did not assign any SITs to 24 (7.9%) strains (Figure 2). SNP barcode-based typing was able to assign a sublineage to 12 strains that were previously unclassified or ill defined by spoligotyping. Disagreements between spoligotyping and SNP-based typing methods were related to the T spoligotype. The complete typing profile for each strain is presented in the online supplementary material (molecular sheet).

3.4. Whole-genome-sequencing-based M. tuberculosis transmission analysis

The final wgSNP alignment of the 305 *M. tuberculosis* strains included in this study resulted in 28,974 positions and evidenced an average distance of 498.68 SNPs between strains. Five SNPs was used as the threshold for genomic relatedness detection in *M. tuberculosis* genomes, which could reveal recent or ongoing transmission as proposed by Walker et al. [18]. From 305 strains, 168 (55.1%) were grouped into 28 genomic clusters. Among those, 22 epi-links were identified, involving 77/168 (45.8%) patients: 47 at community level, 30 in prison and two household contacts (see online supplementary material). Among the 169 patients with MDR-TB, 124 (73.4%) were grouped into genomic clusters. The three largest clusters identified (GC1, GC2 and GC3) harboured strains

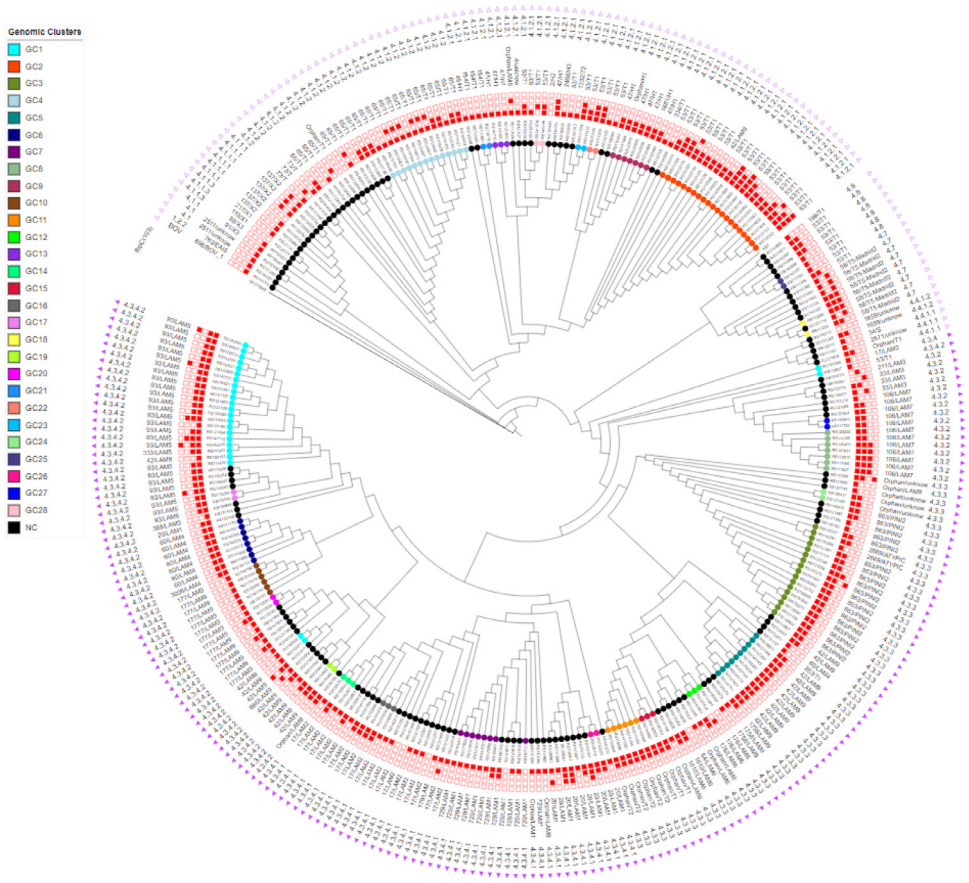

Figure 2. Phylogenetic tree reconstructed using the maximum likelihood approach from genome-wide alignment of the 305 *Mycobacterium tuberculosis* strains with 28,974 high-quality positions. The resulting tree was rooted on *Mycobacterium cannettii* and annotated in Itol [17]. The tips were coloured according to genomic clustering (see legend). The phenotypic drug susceptibility profile for first-line drugs is represented in the red squares (resistant, filled squares; susceptible, empty squares). Spoligotyping clade/shared international type and single-nucleotide-polymorphism-based typing are annotated, and the triangle indicates the presence or absence of fbpC103 polymorphism (presence, filled triangle; absence, empty triangle).

from the 4.3.4.2 (LAM), 4.1.2.1 (Haarlem) and 4.3.3 (LAM) sublineages, respectively. GC1 comprised 25 strains, 19 of which belonged to SIT 93/LAM5. GC2 had 21 strains in total, 20 of which belonged to SIT 53/T1. Of the 17 strains in GC3, 15 had a spoligotyping pattern assigned to SIT 863/PINI2 and 16 had a 12nt insertion at codon 435 of the *rpoB* gene (435QNNP > QNNPQNNP). Furthermore, 14 of the 22 strains harbouring disputed mutations, and susceptible according to the MGIT assay, were grouped into five different clusters, mainly GC4 (9/22).

The largest genomic clusters were mainly composed of MDR strains. However, in nine genomic clusters, the strains presented distinct intracluster drug-resistance patterns. For example, GC1 was composed of 19 MDR strains and five INH-monoresistant strains, and GC4 had a heterogeneous composition (five polyresistant, four INH-monoresistant and four MDR strains). In addition, 13 GCs had intracluster acquisition of additional mutations conferring resistance to first- and second-line drugs (see online supplementary material). The distribution of clustered strains was also analysed in terms of the different cities in Rio Grande do Sul state, and the highest occurrence was found in cities with the highest incidence of TB (mainly the metropolitan area of Porto Alegre and the state capital) [26]. Figure 3 shows the geographical distribution of studied strains, revealing a higher concentration of studied cases and clustered strains (77.4%) in the metropolitan area of Porto Alegre (incidence rate 84.4 cases/100,000 population) [2]. In Porto Alegre, clustered cases were mainly observed in four districts (Santa Tereza, Rubem Berta, Mário Quintana and Sarandi) – disregarding the cases from the city's prison – accounting for 40.3% of clustered cases in the city. The pairwise geographical distance between patient residences within the five larger genomic clusters was statistically lower than observed in non-clustered cases ($P<0.05$).

4. Discussion

Rio Grande do Sul has the fourth highest TB mortality rate among Brazilian states, and recent studies in the region have shown important chains of transmission of DR-*M. tuberculosis* strains that impair disease control [2,4,27,28]. Data are lacking regarding the real incidence of DR-TB in Rio Grande do Sul, as in other Brazilian states, as these data are poorly covered in national reports. In this retrospective study, DR-TB cases identified at Rio Grande do Sul State Central Laboratory were reviewed to obtain the epidemiological scenario of resistance in the state.

Overall, the predicted genome-based resistance showed good concordance with phenotypic resistance testing, ranging from 99.3% and 98.7% resistance to INH and RIF, respectively, to 82.3% and 86.9% resistance to EMB and STR, respectively, in accordance with global data [29,30]. Interestingly, 88.7% of EMB-susceptible strains harbouring resistance-associated mutations were MDR, similar to the results of a previous study in Russia [31]. Lower sensitivity to the detection of STR resistance was found; this has been reported in previous studies with global strain sets, and could be associated with poor genomic predictive performance for STR resistance within Lineage 4 [32,33]. A relevant finding observed from molecular versus phenotypic comparison was the presence

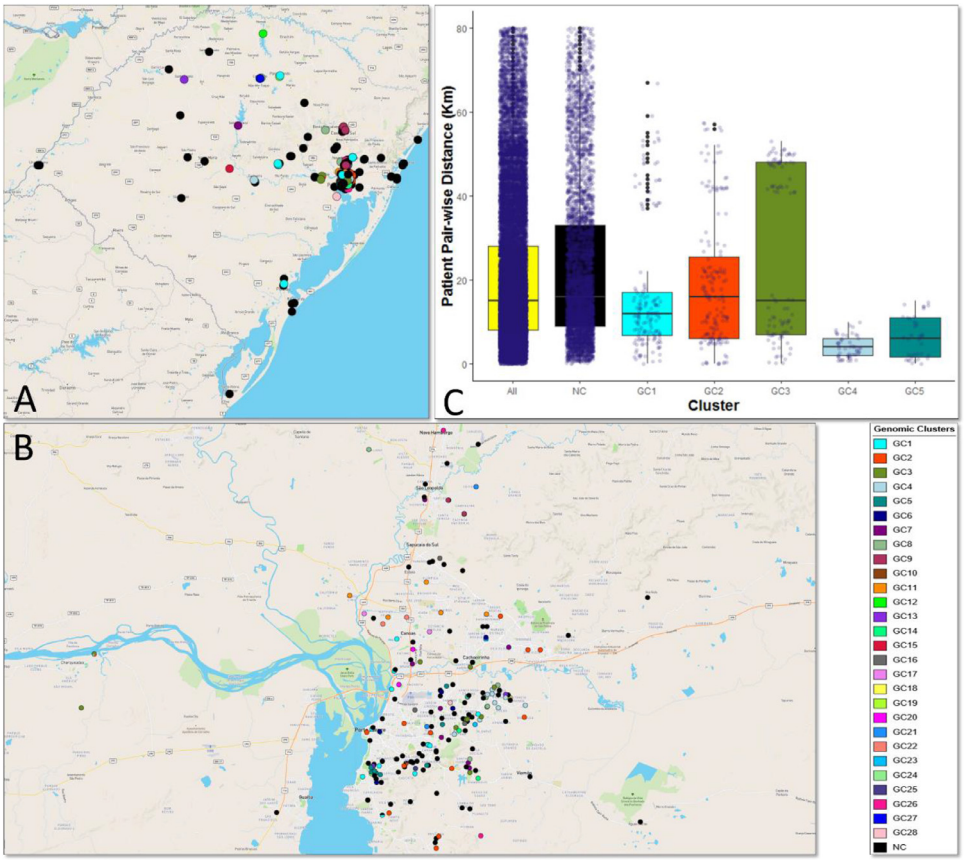

Figure 3. (A) Geographical distribution of the 305 drug-resistant tuberculosis cases in Rio Grande do Sul state. The colours of each marker represent the genomic cluster of each strain (see legend). (B) Zoomed image representing the Porto Alegre metropolitan region. The data were plotted in the map using Microreact [19]. An interactive version is available online at https://microreact.org/project/cZ1SD6nx9LqimJHhNoToG5. (C) Pairwise geographical distances of all patients, and non-clustered and intraclustered groups.

of well-characterized disputed mutations in the *rpoB* gene [25] in 22 strains that were susceptible to RIF on MGIT. In these strains, disputed mutations were found in 21.9% (21/96) of phenotypically INH-monoresistant strains, suggesting potential underestimation of the proportion of MDR cases, and may result in incorrect anti-TB treatment. The fact that 17 patients infected with these strains received an MDR-TB regimen shows the failure of the first-line regimen despite being considered simple INH-monoresistant strains, and two deaths during treatment shows the clinical relevance of these mutations. In addition, the potential for dissemination of these strains is further stressed by the detection of isolates bearing these variants in genomic transmission clusters, particularly GC4. These novel findings in Brazil reveal the importance of tracking strains harbouring disputed mutations for accurate TB surveillance and control in the region.

As revealed by previous studies carried out in the region [4,27,28], the main genotypes associated with first-line drug resistance include high confidence mutations: *katG* S315T (in 50% of resistant strains), *katG* S315T + *inhA* -15 C>T (in 17.6% of resistant strains) and *inhA* -15 C>T (in 10.2% of resistant strains) causing resistance to INH. However, 3.9% (11/284) of INH-resistant strains only carried resistance-conferring mutations on the *ahpC* promoter (-48 G>A, -52 C>T, -54 C>T and -57 C>T), eight of which were resistant to INH alone. Despite limited evidence about the role of mutations in the *ahpC-oxyR* intergenic region conferring resistance to INH, several studies have found a potential association between these variants and INH resistance [29,34], but usually as a compensatory mechanism to mutations at *katG*, mainly non-315 mutations. Overexpression of the *ahpC* gene, producing an alkyl peroxidase that also protects the bacillus against the toxic effects of organic peroxides, may occur due to reduced activity of catalase peroxidase. In addition, this study found mutations in the *ahpC* gene in six strains, which also carried variants in the *katG* gene, all of which were non-315 mutations. Regarding the 11 strains that carried *ahpC* gene variants alone, no other molecular markers that could elucidate the mechanisms causing resistance were identified. In order to further understand the occurrence of resistance in these strains, there is a need to analyse other molecular mechanisms, such as the expression of efflux pumps.

RIF resistance was mainly caused by variants carrying the *rpoB* S450L (in 54.6%) and *rpoB* S450L + A286V (in 10.9%) mutations, and by the 12-nucleotide insertion at codon 435 of the *rpoB* gene (in 9.2%), stressing the significant spread of these highly RIF-resistant strains (MIC ≥32 mg/L) in the population [6]. In the same way, most common resistance-associated mutations were found among clinical strains that were phenotypically resistant to EMB (*embB* M306V in 83.3% of resistant strains) and STR (*rrs* 492 C>T in 14.9% and *rpsL* K43R in 14.9% of resistant strains), similar to results in global collections [14].

Mutations conferring resistance to second-line drugs were found in nine strains: six MDR strains had mutations conferring fluoroquinolone resistance in the *gyrA* gene, one polyresistant and two INH-monoresistant strains had associated resistance mutations in the *gyrA* gene, and one of the latter also had a nucleotide change (1401a>g) in the *rrs* gene that was related to resistance to injectable second-line drugs, characterizing an XDR-TB case (see online supplementary material). According to clinical data, only one of these nine pre/XDR-TB individuals had a favourable out-

come, and three of them were lost to follow-up, which may have contributed to the spread of this resistance. In this study, DST to second-line anti-TB drugs was not performed, precluding comparison with WGS-based predictions. However, a previous work showed good concordance among resistance-conferring mutations and phenotypic resistance to fluoroquinolones amid a subset of clinical strains from the same region [4].

As characterized previously by recent studies in the region, the population structure of *M. tuberculosis* is mainly composed of LAM and Haarlem strains [4,27]. The same predominance of LAM strains occurs in the bordering state of Santa Catarina [35], despite predominance of the 4.3.3 sublineage in the neighbouring state compared with predominance of the 4.3.4.2 sublineage observed in the present study. As well as being able to determine the sublineage of strains not classified by spoligotyping, SNP typing was also able to reclassify a large set of strains previously defined as T sublineage into more plausible sublineages based on their spoligotype pattern. From 93 strains that were assigned as T clade by SITVIT2, only 16 were classified as T clade by SNP-based typing; the remainder were mainly assigned to the 4.1.2.1/Haarlem sublineage (39.8%). This fact was evidenced by 12 strains carrying the fbpC103 variant, a known LAM marker, and by the position in the phylogenetic tree. Moreover, the ill-defined family T assigned by spoligotyping, especially the spoligotype SIT53, has been demonstrated previously as a common pattern for different sublineages in Lineage 4 [36].

Overall, the estimated recent transmission rate of 55.1% among the 305 studied samples, which is higher than that registered in the bordering state of Santa Catarina [35] (mainly susceptible strains) and other regions with a high burden of TB [37] and MDR-TB [38], shows that important transmission chains are feeding the ongoing transmission in the region. However, the most worrying fact revealed in this study concerns MDR strains, 73.4% of which were grouped in genomic clusters. The three larger genomic clusters found in the present analysis consisted, almost exclusively, of MDR strains. GC1 was mainly constituted (24/25) by strains from individuals without a history of imprisonment, with 15 individuals involved in identified community epi-links, representing a large ongoing chain of transmission in the community. In GC2, 10/21 strains were isolated from prison inmates, and in GC3, 8/17 of the clustered strains came from prison inmates or individuals with a recent history of incarceration, along with a prison worker. Thus, GC3 represents an important active transmission chain involving the inmate population and the community [6], similar to that found in other regions in Brazil [39] and worldwide [40].

The larger ongoing transmission chains were found in Porto Alegre (capital and metropolitan area). In Porto Alegre, the four districts accounting for 40.3% of clustered cases in the city presented the lowest Human Development Index [7], reinforcing the need for more robust TB control measures in this region. The presence of MDR strains in genomic clusters composed principally of mono-/polyresistant profiles, along with intracluster acquisition of resistance-related mutations seen in multiple (mainly larger) clusters, indicates the de-novo emergence of MDR-TB in those clusters, leading to the amplification of resistance; this is an additional important area of concern for TB control.

5. Conclusions

These findings reveal important aspects of the molecular basis of drug resistance in *M. tuberculosis* strains circulating in Rio Grande do Sul, Brazil, showing the ability of molecular assays to detect drug resistance, and their importance to detect RIF resistance caused by disputed mutations, thus avoiding missing MDR cases. Multiple ongoing transmission events of DR *M. tuberculosis* strains were identified, mainly MDR strains, stressing the need for measures to interrupt *M. tuberculosis* transmission in the region and the need to improve TB control in prisons.

Acknowledgements

The authors wish to thank TGen, C-Path and ReSeqTB for supporting WGS, and the Brazilian Network of Tuberculosis Research for enabling this partnership. The authors would also like to thank the Mycobacteria Laboratory at LACEN-RS/CEVS/SES-RS for performing and providing information on drug susceptibility testing.

Funding: This study was funded, in part, by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (Grant No. 001) and supported by Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (Grant No. 17/1265–8 INCT-TB). JP is supported by Fundação para a Ciência e Tecnologia through Estímulo Individual ao Emprego Científico (CEECIND/00394/2017).

Competing interests: None declared.

Ethical approval: This project was approved by the Research Ethics Committee of the Fundação Estadual de Produção e Pesquisa em Saúde (Protocol No. 1.587.621 CAAE: 18269313.0.0000.5320).

Data availability: *M. tuberculosis* genome data were deposited in the NCBI BioProject database (IDs: PRJNA535343, PRJNA639713 and PRJNA692642). Individual accession numbers for genomes analysed in this study are given in the online supplementary material.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijantimicag.2021.106401.

References

[1] WHO Global tuberculosis report 2019. Geneva: World Health Organization; 2019.
[2] Secretaria de Vigilância em Saúde, Brasil Ministério da Saúde Boletim Epidemiológico. Brasília: Secretaria de Vigilância em Saúde; 2020.
[3] Perdigão J, Silva C, Diniz J, Pereira C, Machado D, Ramos J, et al. Clonal expansion across the seas as seen through CPLP-TB database: a joint effort in cataloguing Mycobacterium tuberculosis genetic diversity in Portuguese-speaking countries. Infect Genet Evol 2018;72:44–58. doi:10.1016/j.meegid.2018.03.011.
[4] Salvato RS, Costa ERD, Reis AJ, Schiefelbein SH, Halon ML, Barcellos RB, et al. First insights into circulating XDR and pre-XDR Mycobacterium tuberculosis in Southern Brazil. Infect Genet Evol 2020;78:104127. doi:10.1016/j.meegid.2019.104127.
[5] Dalla Costa ER, Vasconcelos SEG, Esteves LS, Gomes HM, Gomes LL, da Silva PA, et al. Multidrug-resistant Mycobacterium tuberculosis of the Latin American Mediterranean lineage, wrongly identified as Mycobacterium pinnipedii (Spoligotype International Type 863 [SIT863]), causing active tuberculosis in South Brazil. J Clin Microbiol 2015;53:3805–11.
[6] Rossetti ML, Almeida da Silva PE, Salvato RS, Reis AJ, Schiefelbein SH, von Groll A, et al. A highly rifampicin resistant Mycobacterium tuberculosis strain emerging in Southern Brazil. Tuberculosis 2020;125:102015.
[7] IBGE. Available at: https://cidades.ibge.gov.br/brasil/panorama [accessed 9 December 2019] 2021.
[8] Van Embden JDA, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol 1993;31:406–9.
[9] Bolger AM, Lohse M, Usadel B. Genome analysis Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.
[10] Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25:1754–60.
[11] Cruz LM, Garcı F, Okonechnikov K, Tarazona S, Dopazo J, Meyer TF, et al. Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics 2012;28:2678–9.
[12] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–303.
[13] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–9.
[14] Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K, Ali S, et al. Genome-wide analysis of multi- and extensively drug-resistant Mycobacterium tuberculosis. Nat Genet 2018;50:307–16.
[15] Posada D. jModelTest: phylogenetic model averaging. Mol Biol Evol 2008;25:1253–6.

[16] Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 2014;30:1312–13.
[17] Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 2019;47:W256–9.
[18] Walker TM, Ip CLC, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. Lancet Infect Dis 2013;13:137–46.
[19] Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. Microb Genomics 2016;2:e000093.
[20] Coll F, McNerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill–Cawthorne G, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. Genome Med 2015;7:51.
[21] Xia E, Teo Y-Y, Ong RT-H. SpoTyping: fast and accurate in silico Mycobacterium spoligotyping from sequence reads. Genome Med 2016;8:19.
[22] Couvin D, David A, Zozio T, Rastogi N. Macro-geographical specificities of the prevailing tuberculosis epidemic as seen through SITVIT2, an updated version of the Mycobacterium tuberculosis genotyping database. Infect Genet Evol 2019;72:31–43.
[23] Palomino J-C, Martin A, Camacho M, Guerra H, Swings J, Portaels F. Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in Mycobacterium tuberculosis. Antimicrob Agents Chemother 2002;46:2720–2.
[24] Martin A, Palomino JC. Resazurin microtiter assay (REMA): procedure manual. Belgium: Institute of Tropical Medicine, Mycobacteriology Unit Antwerp; 2009.
[25] Miotto P, Cabibbe AM, Borroni E, Degano M, Cirilloa DM. Role of disputed mutations in the rpoB gene in interpretation of automated liquid MGIT culture results for rifampin susceptibility testing of mycobacterium tuberculosis. J Clin Microbiol 2018;56 e01599-17.
[26] Programa Estadual de Controle da Tuberculose Centro Estadual de Vigilância em Saúde SE da S do RG do S. Tuberculose: Informe Epidemiológico; 2019.
[27] Salvato RS, Schiefelbein S, Barcellos RB, Praetzel BM, Anusca IS, Esteves LS, et al. Molecular characterisation of multidrug-resistant Mycobacterium tuberculosis isolates from a high-burden tuberculosis state in Brazil. Epidemiol Infect 2019;147:e216.
[28] Esteves LS, Dalla Costa ER, Vasconcellos SEG, Vargas A, Ferreira SLM Junior, Halon ML, et al. Genetic diversity of Mycobacterium tuberculosis isoniazid monoresistant and multidrug-resistant in Rio Grande do Sul, a tuberculosis high-burden state in Brazil. Tuberculosis 2018;110:36–43.
[29] Allix-Béguec C, Arandjelovic I, Bi L, Beckert P, Bonnet M, Bradley P, et al. Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. N Engl J Med 2018;379:1403–15.
[30] Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in Mycobacterium tuberculosis. Eur Respir J 2017;50:1701354.
[31] Mokrousov I, Otten T, Vyshnevskiy B, Narvskaya O. Detection of embB306 mutations in ethambutol-susceptible clinical isolates of Mycobacterium tuberculosis from Northwestern Russia: implications for genotypic resistance testing. J Clin Microbiol 2002;40:3810–13.
[32] Schleusener V, Köser CU, Beckert P, Niemann S, Feuerriegel S. Mycobacterium tuberculosis resistance prediction and lineage classification from genome sequencing: comparison of automated analysis tools. Sci Rep 2017;7:46327.
[33] Mahé P, El Azami M, Barlas P, Tournoud M. A large scale evaluation of tbprofiler and mykrobe for antibiotic resistance prediction in Mycobacterium tuberculosis. PeerJ 2019;7:e6857.
[34] Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis 2015;15:1193–202.
[35] Verza M, Scheffer MC, Salvato RS, Schorner MA, Barazzetti FH, Machado H de M, et al. Genomic epidemiology of Mycobacterium tuberculosis in Santa Catarina. Southern Brazil. Sci Rep 2020;10:12891.
[36] Mokrousov I. Ubiquitous and multifaceted: SIT53 spoligotype does not correlate with any particular family of Mycobacterium tuberculosis. Tuberculosis 2021;126:102024.
[37] Miyahara R, Smittipat N, Juthayothin T, Yanai H, Disratthakit A, Imsanguan W, et al. Risk factors associated with large clusters of tuberculosis patients determined by whole-genome sequencing in a high-tuberculosis-burden country. Tuberculosis 2020;125:101991.
[38] Chee CBE, Lim LKY, Ong RTH, Sng LH, Hsu LY, Lee VJM, et al. Whole genome sequencing analysis of multidrug-resistant tuberculosis in Singapore, 2006–2018. Eur J Clin Microbiol Infect Dis 2021;40:1079–83.
[39] Sacchi FPC, Praca RM, Tatara MB, Simonsen V, Ferrazoli L, Croda MG, et al. Prisons as reservoir for community transmission of tuberculosis. Brazil. Emerg Infect Dis 2015;21:452–5.
[40] Umpeleva T, Belousova K, Golubeva L, Boteva T, Morozova I, Vyazovaya A, et al. Molecular characteristics of Mycobacterium tuberculosis in the “closed” Russian town with limited population migration. Infect Genet Evol 2020;79:104174.