

Citywide Transmission of Multidrug-resistant Tuberculosis Under China's Rapid Urbanization: A Retrospective Population-based Genomic Spatial Epidemiological Study

Qi Jiang,^{1,2} Qingyun Liu,^{1,2} Lecai Ji,¹ Jinli Li,¹ Yaling Zeng,¹ Lianguang Meng,¹ Geyang Luo,² Chongguang Yang,³ Howard E. Takiff,^{4,5} Zheng Yang,¹ Weiguo Tan,¹ Weiye Yu,^{1,a} and Qian Gao^{1,2,a}

¹Shenzhen Center for Chronic Disease Control, Shenzhen, China; ²Key Laboratory of Medical Molecular Virology (Ministry of Education, National Health Commission, Chinese Academy of Medical Sciences), School of Basic Medical Sciences, Shanghai Medical College and Shanghai Public Health Clinical Center, Fudan University, Shenzhen, China; ³School of Public Health, Yale University, New Haven, Connecticut, USA; ⁴Integrated Mycobacterial Pathogenomics Unit, Institut Pasteur, Paris, France; and ⁵Nanshan Center for Chronic Disease Control, Shenzhen, China

Background. Population movement could extend multidrug-resistant tuberculosis (MDR-TB) transmission and complicate its global prevalence. We sought to identify the high-risk populations and geographic sites of MDR-TB transmission in Shenzhen, the most common destination for internal migrants in China.

Methods. We performed a population-based, retrospective study in patients diagnosed with MDR-TB in Shenzhen during 2013–2017. By defining genomic clusters with a threshold of 12–single-nucleotide polymorphism distance based on whole-genome sequencing of their clinical strains, the clustering rate was calculated to evaluate the level of recent transmission. Risk factors were identified by multivariable logistic regression. To further delineate the epidemiological links, we invited the genomic-clustered patients to an in-depth social network investigation.

Results. In total, 105 (25.2%) of the 417 enrolled patients with MDR-TB were grouped into 40 genome clusters, suggesting recent transmission of MDR strains. The adjusted risk for student to have a clustered strain was 4.05 (95% confidence interval, 1.06–17.0) times greater than other patients. The majority (70%, 28/40) of the genomic clusters involved patients who lived in different districts, with residences separated by an average of 8.76 kilometers. Other than household members, confirmed epidemiological links were also identified among classmates and workplace colleagues.

Conclusions. These findings demonstrate that local transmission of MDR-TB is a serious problem in Shenzhen. While most transmission occurred between people who lived distant from each other, there was clear evidence that transmission occurred in schools and workplaces, which should be included as targeted sites for active case finding.

Keywords. transmission; multidrug-resistant tuberculosis; whole-genome sequencing; internal migrants; spatial analysis.

While the acquisition of drug resistance due to inadequate therapy of previously drug-sensitive *Mycobacterium tuberculosis* (*Mtb*) strains was the early origin of multidrug-resistant tuberculosis (MDR-TB) or rifampin-resistant (RR) tuberculosis (TB), it is currently recognized that recent transmission is primarily driving the global endemic of MDR-TB [1, 2], and outbreaks of MDR-TB in urban centers have highlighted the urgent need to interrupt MDR-TB transmission. China has the second largest number of MDR-TB cases worldwide, estimated at 73 000 MDR/RR-TB cases, which is 13% of the global burden [3].

Population movement can extend transmission of MDR-TB transmission beyond the boundaries of an individual city, and spread even across a country's borders, thereby make the global prevalence more complicated [4–6]. Internal migration in China, principally young people going from rural areas to work in the large cities, has increased rapidly since 1979 when the government implemented sweeping economic reforms. China has about 287 million migrant workers who accounted for 35% of the nation's workforce until 2017 [7]. Shenzhen was established as China's first special economic zone, and over the past 40 years it has grown from a small village of 30 000 inhabitants to become the core of the “world's factory,” with a population of about 15 million [8, 9]. Unlike the cross-border immigrants in developed countries, this “floating population” is more transient and more difficult to access by the city's healthcare system, which complicates efforts to control transmissible diseases such as TB [10]. Our previous work analyzed TB transmission between internal migrants and local residents of Shanghai, China, and estimated that more than two-thirds of the migrant cases with TB in big genomic clusters were infected after they

Received 25 June 2019; editorial decision 7 August 2019; accepted 26 August 2019; published online August 28, 2019.

^aW. Y. and Q. G. contributed equally to this work.

Correspondence: Q. Gao, Key Laboratory of Medical Molecular Virology, School of Basic Medical Sciences, Fudan University, Room 201, Fuxing Building, Dongan Road No. 131, Shanghai, China 200032 (qiangao@fudan.edu.cn).

Clinical Infectious Diseases® 2020;71(1):142–51

© The Author(s) 2019. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com.
DOI: 10.1093/cid/ciz790

migrated to Shanghai [11]. Where or how the transmission occurs among the migrants, however, is not clear. To further investigate the transmission of MDR-TB among migrants, we utilized whole-genome sequencing (WGS) of clinical MDR-TB strains collected during 2013–2017 in Shenzhen to detect clustered strains. This was then combined with epidemiological data and patient interviews to detect risk factors for transmission and possible geographic groupings or social links that could explain the transmission between genomic-clustered cases.

METHODS

Study Setting and Sample Enrollment

Shenzhen lies at the southern tip of Guangdong province, just across the bay from Hong Kong, and is divided into 10 districts and 74 subdistricts. Tuberculosis is mainly diagnosed and treated at the district-level Centers for Chronic Disease Control (CCDC) and 1 comprehensive infectious diseases hospital. The Chinese national guidelines [12] stipulate that 3 sputum samples should be collected from suspected TB cases and sent for Ziehl-Neelsen staining and smear microscopy and sputum cultures. For cases who would receive their treatment in Shenzhen, the culture-positive strains would be sent to the Shenzhen CCDC for phenotypic drug sensitivity testing with rifampin and isoniazid to identify patients with MDR-TB, who are then treated and managed at this central CCDC. All the patients with clinical MDR-TB strains stored at the Shenzhen CCDC during 2013–2017 were enrolled in this study.

Whole-genome Sequencing

The MDR-TB strains were re-grown on Löwenstein-Jensen medium and their genomic DNA were extracted using the cetyl trimethyl ammonium bromide method [13]. For each sample, a pair-end library was constructed and sequenced on an Illumina platform (Illumina) at the Shenzhen Union of Medicine, with an expected coverage of 100×. The raw data were analyzed using in-house Perl scripts (www.perl.org) to call the single-nucleotide polymorphisms (SNPs) and to identify fixed mutations (frequency $\geq 75\%$) [2]. A genomic cluster was defined as strains with a genetic distance of 12 SNPs or less, suggesting they were the result of recent transmission. Strains that differed by more than 12 SNPs as compared with all other strains were defined as unique strains [2]. The drug-resistance profile was predicted for 14 anti-TB drugs based on the mutations reported to be associated with resistance (Supplementary Table 1) [14]. Pre-extensively drug-resistant (pre-XDR) TB was defined as an MDR-TB strain with additional resistance to fluoroquinolones (FQs) or any second-line injected drug, including kanamycin, amikacin, and capreomycin, while the XDR-TB strain was defined as resistance to both antibiotic classes. Phylogenetic trees were constructed using the maximum-likelihood method on an online platform (<http://samtb.szmbzx.com>). Visualization of the bacteriological information was performed at Interactive

Tree Of Life (<https://itol.embl.de/>). The sequencing data have been deposited with links to BioProject accession number PRJNA559678 in the National Center for Biotechnology Information BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>).

Statistical Methods and Spatial Analysis

Patients' demographic and clinical characteristics were collected from the National TB registration system and the Shenzhen TB clinical system. Statistical analyses were performed in R (version 3.5.1). The chi-square test was used to assess possible associations of genomic clustering. Univariable and multivariate logistic regression was used to estimate the odds ratios (ORs) and 95% confidence intervals (CIs). Factors with a *P* value less than .05 in the final multivariate model were considered to be independently associated with genomic clustering. Spatial analysis and visualization were performed in ArcGIS (version 10.2; Esri). We used point data for mapping of the MDR-TB cases by kernel density estimation methods using Gaussian smoothing to analyze the aggregation of patients with TB [15].

Interview and Social Network Analysis

Patients whose strains belonged to genomic clusters with at least 3 cases were invited to participate in a structured in-depth interview about their social network. After providing the informed consent, the enrolled patients were given a questionnaire to collect data on close contacts, social contacts, and places they frequented in the 5 years before their diagnosis of MDR-TB, including the detailed addresses of their residences, workplaces, and public community centers and facilities. Social networks between the clustered cases were constructed using Cytoscape software (version 3.7.0; <http://www.cytoscape.org/>) [16].

RESULTS

Basic Description of Cases of Multidrug-resistant Tuberculosis in Shenzhen

The Shenzhen CCDC collected and stored 8866 clinical *Mtb* strains sent from 10 districts during 2013–2017, and 450 (5.08%) of them were MDR-TB (Figure 1A). The median age of the patients was 31 years (interquartile range, 27–42 years), and 62.0% (269/434) were male, excluding 16 patients for whom there was no basic information. For 392 cases whose home provinces could be inferred from their ID number, only 12 cases (3.1%, 12/392) were born in Shenzhen, while one-quarter of the cases (26.0%, 102/392) were from neighboring areas in Guangdong province, and the remaining patients (70.9%, 278/392) were originally from the other 22 provinces of China (Figure 1B).

Whole-genome sequencing was successfully performed on strains isolated from 417 (92.7%, 417/450) cases of MDR-TB that were then subject to further genomic and epidemiological analysis (Figure 1A). There were 23 strains that failed

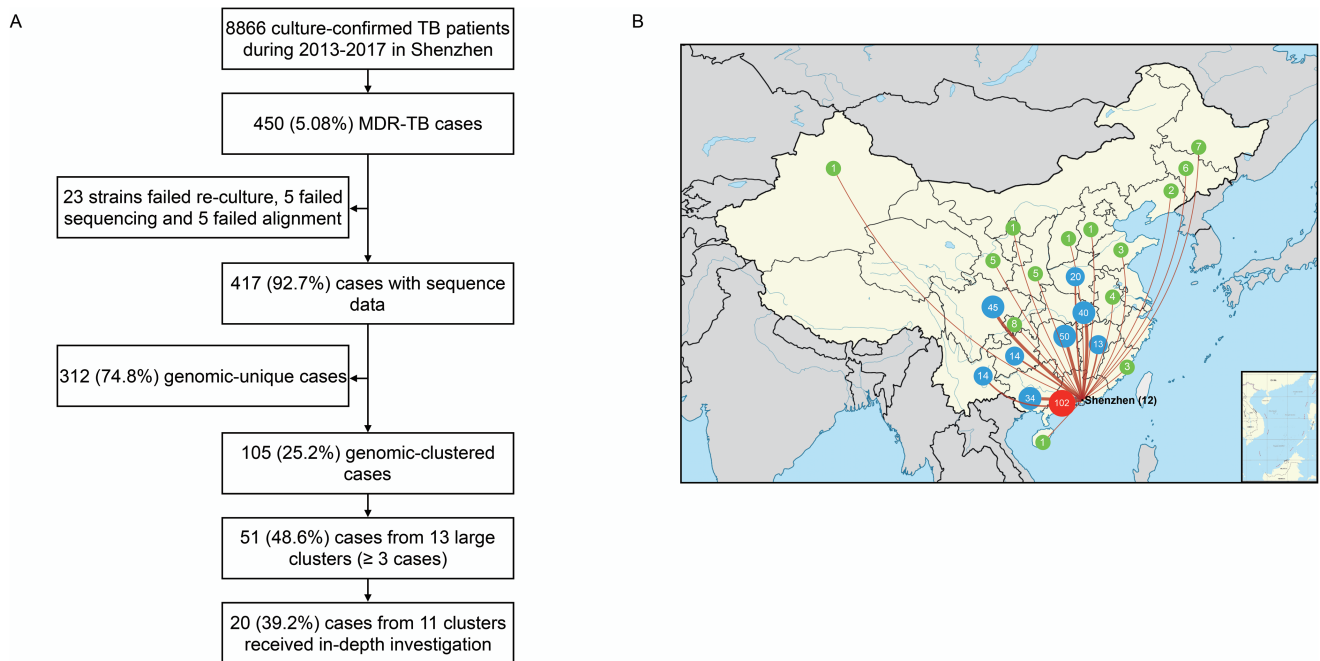


Figure 1. A, Sample enrollment and study flowchart. B, Provinces of origin of patients with multidrug-resistant tuberculosis before migrating to Shenzhen. The numbers of enrolled patients who were born in the different provinces of China are indicated within circles. Green circles indicate provinces where less than 10 patients were born, blue circles indicated provinces where 10–50 patients were born, and the red circle indicates cases from Guangdong province with the exception of Shenzhen. Abbreviations: MDR-TB, multidrug-resistant tuberculosis; TB, tuberculosis.

re-culture and 10 that failed sequencing or alignment, all of which were excluded from the analysis. The distribution of pairwise and minimum SNP distance between each possible pair of the strains is shown in [Supplementary Figure 1](#) to delineate the genetic structure of this population. A maximum-likelihood phylogeny tree was constructed for the 417 MDR-TB strains ([Figure 2](#)): 351 (84.2%) belonged to lineage 2, 64 (15.3%) to lineage 4, and 2 (0.5%) to lineage 1. The tree also displays the resistance profile for 11 anti-TB drugs based on the presence of validated resistance-conferring mutations. Among the MDR cases, 25.7% (107/417) harbored mutations in *gyrA/B* and were classified as FQ-resistant strains, and 17.8% (19/107) of these were XDR-TB strains with additional drug-resistant mutations in the *rrs* or the *eis* promoter. The percentages of strains resistant to the 11 anti-TB drugs are shown in [Supplementary Table 1](#). No mutation was detected in genes associated with resistance to linezolid, clofazimine, or bedaquiline.

Identification of Genomic-clustered Cases and Their Risk Factors

A total of 105 strains (25.2%, 105/417) were grouped into 40 genomic clusters containing 2 to 6 strains. The percentage of clustered strains varied between lineages or sublineages ($\chi^2 = 17.3$, $P = .0002$). The highest percentage was found in the modern Beijing sublineage (32.9%, 71/216), followed by the ancient Beijing sublineage (20.7%, 28/135) and lineage 4 (9.38%, 6/64).

Fluoroquinolone-resistant strains were more frequent in clusters than in nonclusters (38.3% vs 21.3%, $\chi^2 = 13.19$, $P = .0003$). In addition, FQ resistance appeared to develop frequently in transmitted strains ([Supplementary Figure 2](#)), because FQ-resistance mutations were acquired in more than half (51.6%, 16/31) of the clusters involving a strain without FQ resistance: 11 became pre-XDR strains and 1 became an XDR-TB strain.

The majority of patients with MDR-TB (64.5%, 269/417) were new cases without a history of TB treatment, and so had been infected by MDR-TB strains from direct transmission ([Table 1](#)). The proportion of new cases was slightly higher among clustered cases than unique cases, but the difference was not statistically significant (71.4% vs 62.2%, $\chi^2 = 3.03$, $P = .082$). In addition, 67 (16.1%) were retreated patients who could have acquired drug resistance during their previous episode or were reinfected with a resistant strain. Another 59 (14.1%) patients received intermittent first-line anti-TB treatment before seeking care in a district CCDC, where they were correctly diagnosed with MDR-TB.

In the multivariate logistic analysis, adjusted by Beijing genotype and FQ resistance, epidemiologic factors that were independently associated with having a clustered strain included having been born in Shenzhen (OR, 3.66; 95% CI, 1.07–12.8) and living in the Longgang or suburban districts (OR, 2.08; 95% CI, 1.07–3.97) ([Table 2](#)). Surprisingly, students had 4.05 (OR, 4.05; 95% CI, 1.06–17.0) times the risk to be in a genomic cluster than other patients.

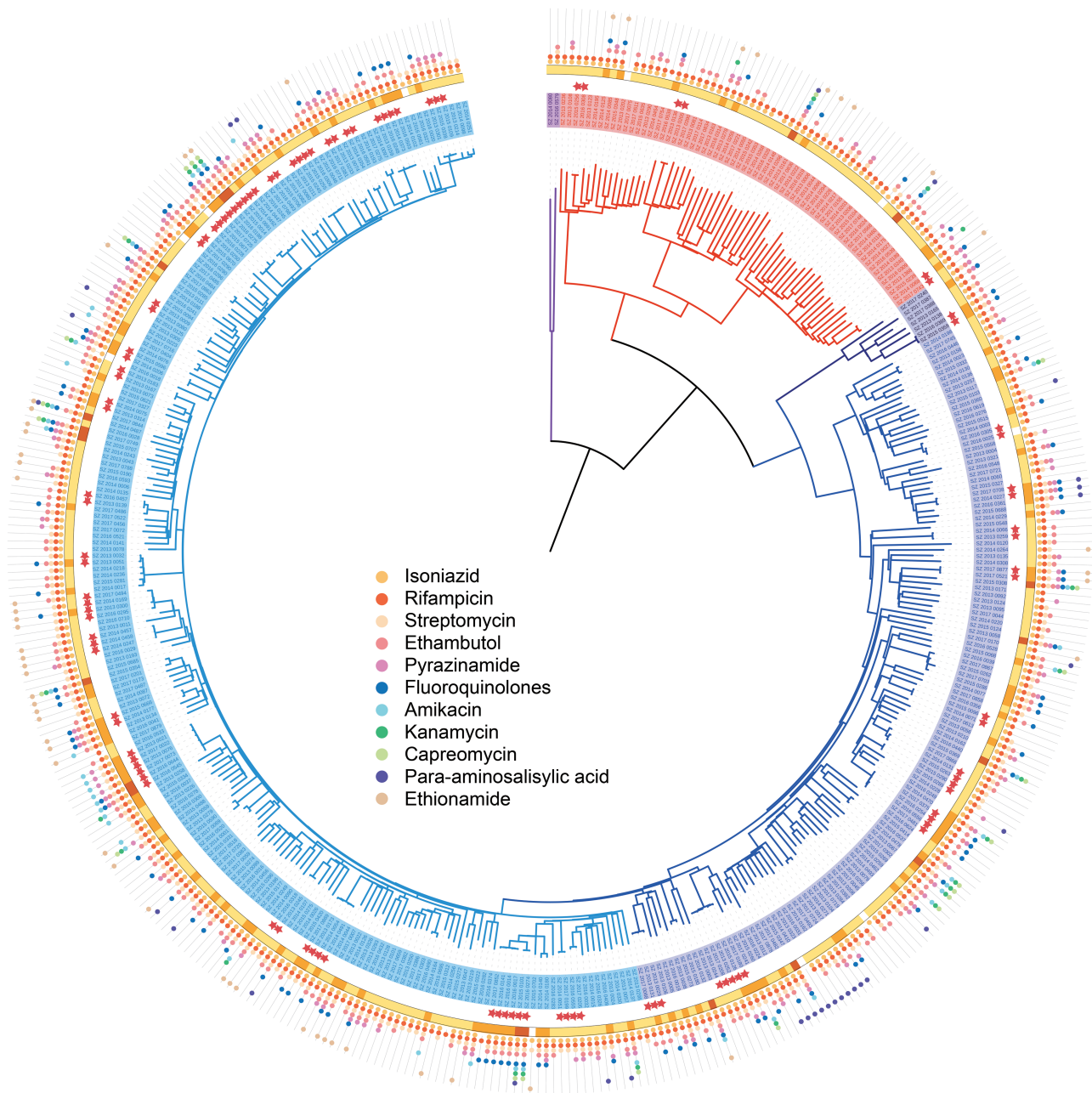


Figure 2. Phylogeny, clustering, and resistance profile of 417 multidrug-resistant tuberculosis (MDR-TB) strains. Purple, red, navy, blue, and cyan branches indicated lineage 1, lineage 4, proto-Beijing, ancient Beijing, and modern Beijing strains, respectively. Solid stars outside the phylogeny indicate genomic-clustered strains differing by ≤ 12 single-nucleotide polymorphisms. The outer yellow-orange-red circle shows the development of MDR to MDR plus fluoroquinolone resistance, and finally to extensive drug resistance. The outermost colored dots indicate the resistance to 11 anti-TB drugs. Abbreviations: MDR, multidrug resistance; MDR-TB, multidrug-resistant tuberculosis.

Spatial Hot Spots of Multidrug-resistant Tuberculosis Cases

To further characterize the spatial distribution of the cases, kernel density maps were drawn using the residential addresses of the patients with clustered and unique MDR-TB (Figure 3A, B). Three areas that were hot spots in the maps of both genomic-clustered and unique cases were in areas of high population density in the centers of Baoan and Luohu districts. Another hot spot for genomic-clustered cases was located in Longgang subdistrict in Longgang district (sub-LG) (arrow in Figure 3B) without a high

population density. Although this initially suggested that sub-LG might be a high-risk site for MDR-TB transmission, it was found that the 6 patients with clustered strains in sub-LG belonged to 6 different clusters and the other cases in each cluster lived in different districts with a residential address 4 to 40 kilometers away.

By comparing the residence districts of the cases within each genomic cluster, we found that in only 12 of the 40 clusters (30%, 12/40) the clustered patients lived in the same district, while the patients within the majority of the clusters (70%,

Table 1. Bacteriological, Demographic, Clinical, and Geographic Characteristics of Genomic-Clustered and Unique Cases of Multidrug-resistant Tuberculosis in Shenzhen, China

	No. (%)			χ^2	P value
	Clustered (n = 105)	Unique (n = 312)	Total (N = 417)		
Bacteriological factors					
Genotype				17.26	.0002
Modern Beijing	71 (67.6)	145 (46.5)	216 (51.8)		
Ancient Beijing	28 (26.7)	107 (34.3)	135 (32.3)		
Non-Beijing	6 (5.7)	60 (19.2)	66 (15.8)		
Drug resistance					
PZA	35 (33.3)	77 (24.7)	112 (26.9)	2.995	.0835
FQ	41 (39.0)	66 (21.2)	107 (25.7)	13.19	.0003
Injectable SLDs	11 (10.5)	36 (11.5)	47 (11.3)	0.0886	.7659
Demographic factors					
Gender				0.0022	.9622
Male	64 (61)	185 (59.3)	249 (59.7)		
Female	39 (37.1)	114 (36.5)	153 (36.7)		
Age, years				2.925	.5704
15–24	21 (20)	41 (13.1)	62 (14.9)		
25–34	42 (40)	135 (43.3)	177 (42.4)		
35–44	17 (16.2)	57 (18.3)	74 (17.7)		
45–54	15 (14.3)	40 (12.8)	55 (13.2)		
≥55	8 (7.6)	25 (8)	33 (7.9)		
Ethnic group				4.152	.0416
Han Chinese	100 (95.2)	272 (87.2)	372 (89.2)		
Non-Han minority	3 (2.9)	27 (8.7)	30 (7.2)		
Occupation				12.29	.0153
Student	7 (6.7)	4 (1.3)	11 (2.6)		
Public service	15 (14.3)	30 (9.6)	45 (10.8)		
Office worker	7 (6.7)	31 (9.9)	38 (9.1)		
Factory worker	23 (21.9)	88 (28.2)	111 (26.6)		
Unemployed/unknown	49 (46.7)	146 (46.8)	195 (46.8)		
Clinical factors					
TB history				4.349	.1137
New case	75 (71.4)	184 (62.2)	269 (64.5)		
Retreated case	11 (10.5)	56 (17.9)	67 (16.1)		
Inadequate treatment	15 (14.3)	44 (14.1)	59 (14.1)		
Total delay				1.772	.6211
≤2 weeks	24 (22.9)	78 (25)	102 (24.5)		
2–4 weeks	32 (30.5)	79 (25.3)	111 (26.6)		
4–8 weeks	14 (13.3)	33 (10.6)	47 (11.3)		
>8 weeks	11 (10.5)	41 (13.1)	52 (12.5)		
Geographic factors					
Birthplace				5.057	.0798
Shenzhen	6 (5.7)	6 (1.9)	12 (2.9)		
Guangdong	26 (24.8)	70 (22.4)	96 (23)		
Other	58 (55.2)	199 (63.8)	257 (61.6)		
Living district				17.80	.0068
Baoan	36 (34.3)	124 (39.7)	160 (38.4)		
Longhua	8 (7.6)	32 (10.3)	40 (9.6)		
Nanshan	5 (4.8)	46 (14.7)	51 (12.2)		
Futian	14 (13.3)	29 (9.3)	43 (10.3)		
Luohu	12 (11.4)	22 (7.1)	34 (8.2)		
Longgang	21 (20)	37 (11.9)	58 (13.9)		
Suburbs	7 (6.7)	8 (2.6)	15 (3.6)		

Abbreviations: FQ, fluoroquinolone; PZA, pyrazinamide; SLD, second-line drug; TB, tuberculosis.

Table 2. Univariable and Multivariable Logistic Regression on the Risk Factors of Multidrug Resistance Clustering

	Univariable Regression		Multivariable Regression	
	OR (95% CI)	P value	OR (95% CI)	P value
Age group, years				
15–24	1.605 (.885–2.847)	.1110	...	
≥25	Ref		...	
Ethnic group				
Han Chinese	3.309 (1.137–14.07)	.0535	...	
Non-Han minority	Ref		...	
Occupation				
Student	5.492 (1.622–21.35)	.0076	4.049 (1.055–17.01)	.0427
Other	Ref		...	
Genotype				
Beijing strains	3.929 (1.774–10.43)	.0021	3.297 (1.352–9.933)	.0168
Non-Beijing strains	Ref		...	
FQ resistance				
FQ-resistant	2.533 (1.570–4.079)	.0001	2.299 (1.330–3.958)	.0027
FQ-susceptible	Ref		...	
Birthplace				
Shenzhen	3.202 (.9779–10.49)	.0488	3.661 (1.072–12.57)	.0347
Other	Ref		...	
Living area				
Longgang, suburbs	2.099 (1.218–3.580)	.0069	2.082 (1.073–3.971)	.0273
Other	Ref		...	

Abbreviations: CI, confidence interval; FQ, fluoroquinolone; OR, odds ratio; Ref, reference.

28/40) lived in different districts. The average geographic distance between the residences of all the paired clustered cases was 8.76 (interquartile range, 2.71–21.54) kilometers (Figure 3C). The distance did not change significantly when adjusting the threshold for defining genomic clusters from 12 SNPs to 5 SNPs or 1 SNP ($P = .4805$ and $P = .4493$, by Wilcoxon rank-sum test; Figure 3D, E). If living within the same community is defined as living within 2 kilometers of each other, only 22 out of 96 pairs (22.9%) of patients with clustered strains could be said to live in the same community, where transmission could plausibly have occurred, while the majority (77.1%, 74/96) of the paired clustered patients lived distantly from one another across the city.

Epidemiological Links of Genomic-clustered Cases

To further delineate the transmission links between the genomic-clustered cases, an in-depth interview of social networks was administered to at least 1 patient belonging to 11 of the 13 clusters with at least 3 cases. Likely due to the high mobility of the migrant population, 26 (51.0%, 26/51) patients could not be contacted, and an additional 5 patients who were contacted (9.8%, 5/51) did not wish to participate. The remaining 20 (39.2%, 20/51) clustered patients from 11 genomic clusters agreed to be interviewed either by phone (16 patients) or face to face (4 patients). From the patients' residential areas and workplaces during the previous 5 years, it was possible to establish epidemiological links for 12 patients in 5 clusters: 2

clusters included married couples, 1 cluster included 2 colleagues from the same workplace, and 2 clusters included patients who had attended the same schools. Patients in another 4 clusters were probably linked because they lived in the same community, and 3 cases of cluster 44 came from the same county in Sichuan province where the transmission probably occurred (Figure 4, Supplementary Table 2). The confirmed epidemiological links, which were missed in the routine clinical history, were only revealed by WGS-informed social network investigation. For example, initial information showed that 2 of the patients in cluster 46 were students, but the schools that were reported in the routine record differed: 1 case reported only her junior high school while the other reported only the senior high school. The third case in this cluster was listed as unemployed and no school was noted. It was discovered in the further interview that they had all attended the same junior high school during 2013–2015, where the transmission of their MDR-TB likely occurred. Therefore, WGS can provide a precise method of guiding the epidemiological investigation to identify confirmed transmission links.

DISCUSSION

This retrospective, citywide, genomic epidemiological study provides a comprehensive picture of the transmission of MDR-TB during a 5-year interval in the newly developed metropolis of Shenzhen, China. We found that 25% of the patients with MDR-TB had genomic-clustered strains, consistent with

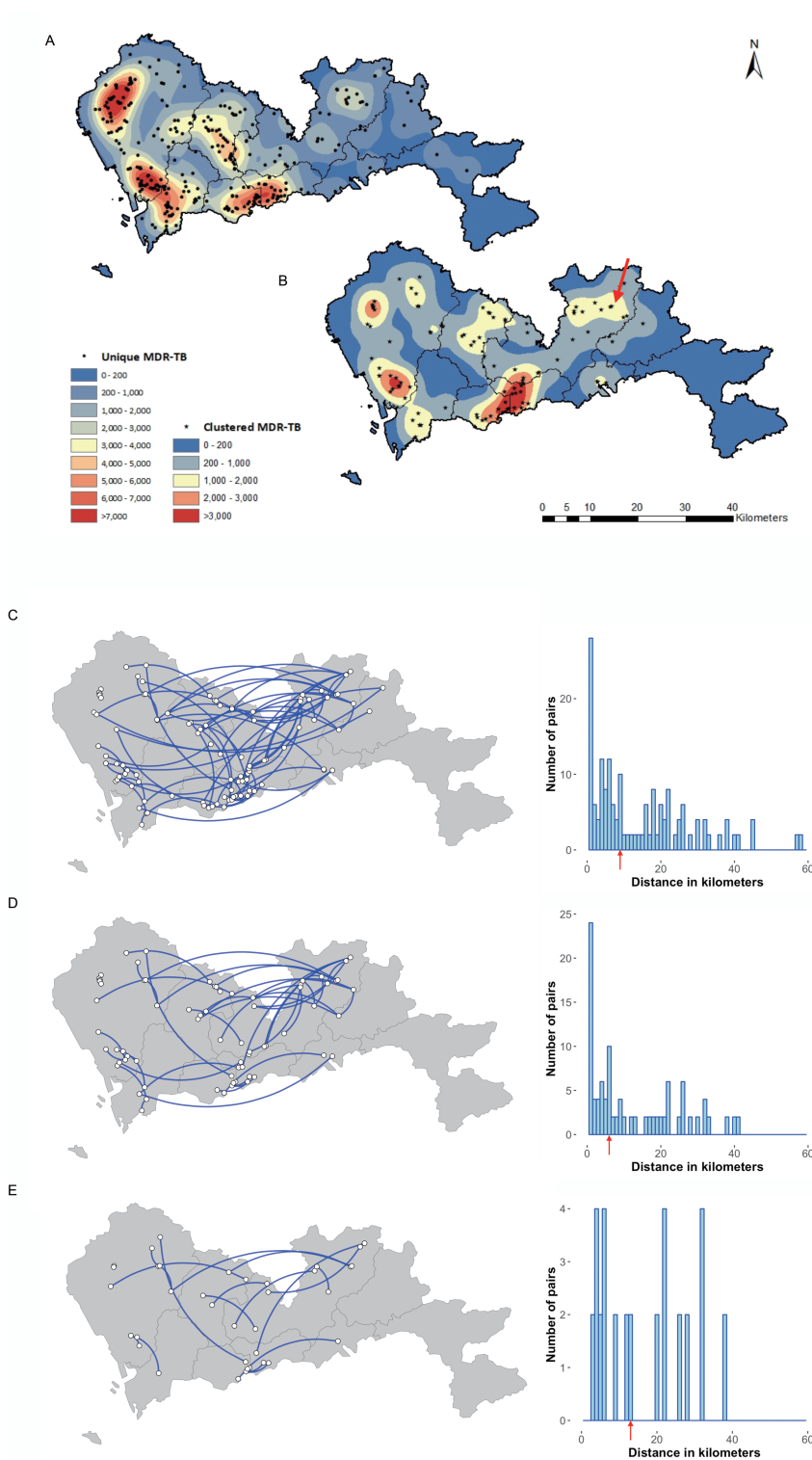


Figure 3. Kernel density maps of unique (A) and clustered (B) multidrug-resistant tuberculosis cases and genomic links and distances between the residences of genomic-linked cases, whose isolates differed by ≤ 12 single-nucleotide polymorphisms (SNPs) (C), ≤ 5 SNPs (D), and ≤ 1 SNP (E), respectively. The maps in panels A and B were colored according to varied kernel density estimation, with the red center representing areas with the most intense concentration of genomic unique (dots) and clustered (stars) patients, respectively. The arrow in panel B indicates the Longgang subdistrict of Longgang district (sub-LG) where patients had the highest risk to be involved in a genomic cluster. Arrows in the histogram indicate the median geographic distance between the paired clustered cases. Abbreviations: MDR-TB, multidrug-resistant tuberculosis; SNP, single-nucleotide polymorphism.

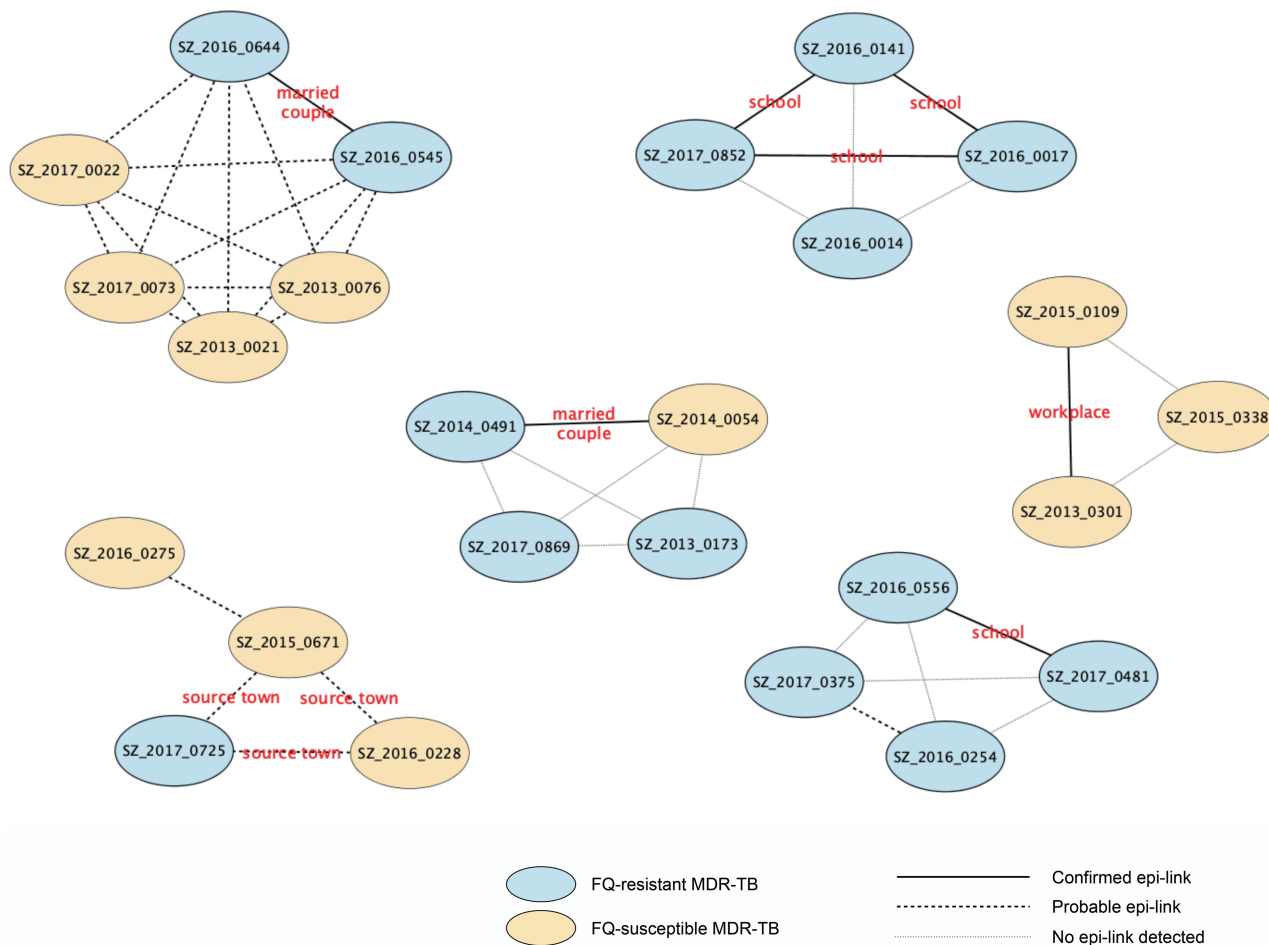


Figure 4. Social network for 6 large clusters with confirmed or probable epidemiological links. Abbreviations: epi-link, epidemiological link; FQ, fluoroquinolone; MDR-TB, multidrug-resistant tuberculosis.

recent transmission. Possible epidemiological links were identified in only 22 of the 96 clustered pairs, where the clustered patients lived within the same community, while confirmed epidemiological links were identified among classmates in schools and colleagues in workplaces as well as with transmission between married couples.

Our finding that 25% of MDR-TB strains were clustered is probably an underestimate because our sample did not include all incident MDR-TB cases. Prior to 2018, culture was only performed on samples that were positive by microscopy, but this test only detects 50–60% of active TB cases [17]. In addition, about one-third of patients diagnosed with TB in Shenzhen requested to return to their home towns for treatment, so their *Mtb* clinical strains were not stored in the Shenzhen CCDC. Therefore, we estimate that our sample included less than 40% of incident MDR-TB strains in Shenzhen over the study period 2013–2017, implying that local MDR-TB transmission may be an even larger problem than our results indicate [18]. The city has made a concerted effort to improve TB control by implementing cultures for all suspected patients and promoting

the use of molecular technologies to detect drug resistance. The finding of a high frequency of FQ resistance in clustered strains highlights the importance of testing MDR-TB strains for FQ resistance so that an effective drug regimen can be implemented promptly to reduce the transmission of pre-XDR or XDR-TB. Our results were also consistent with the widely reported advantage that Beijing strains were more likely to be transmitted [19], especially modern Beijing strains (L2.3), which have dominated the *Mtb* population in China in the past 100 years [20]. This phenomenon might result from the intrinsic enhanced virulence by blocking secretion of ESX-5 substrates through mutations in *ppe38* [21], as well as the historic growth and extensive migration of Beijing people that facilitated the transmission and expansion of this sublineage [20].

We did not identify any specific populations in Shenzhen with an increased risk of transmitting MDR-TB, with the exception of students. There were very few cases of MDR-TB in homeless individuals, drug abusers, or prisoners, which are high-risk populations for TB transmission in developed countries [22]. Although students constituted only 2.6% of all the

patients with MDR-TB, they were more likely to have genomic-clustered isolates, implying a higher risk of recent transmission of TB than the rest of the population. This tendency probably resulted from diagnostic delays, as reported in an MDR-TB outbreak within a school setting in northern China [23]. Students are more likely to suffer from the stigma of a TB diagnosis [24] and, according to treatment protocols in China, would be forbidden from attending school during the whole duration of treatment. After a TB outbreak in a middle school involving 29 students over 6 months in 2017 [25], the China Center for Disease Control and Prevention instituted a series of guidelines for screening students to actively detect cases in schools [26]. Our findings emphasize that students are a target population for increased vigilance, and rapid molecular testing for drug resistance should be included in school screening.

Shenzhen has many factories, crowded offices, and construction sites, and these workplaces are sites with repeated interpersonal contacts and were also identified in our study as presumed transmission links between genomic-clustered patients. Worksites have received less attention as possible transmission hot spots than patients' residences [27], but our study suggests that workplaces in Shenzhen might be another important target for case finding. In large cities such as Shenzhen, the residents, especially migrant workers, often travel extensively within the city for work, medical care, or other activities, which could explain the citywide pattern of TB transmission.

Transmission of the MDR-TB in Shenzhen appeared to be citywide, with large distances between the residential addresses of the genomic-clustered patients, regardless of the SNP cutoff to define a genomic cluster. Consensus on a fixed number of SNPs that represents a threshold has not been reached, and its usage should be based on varied population settings, such as taking the time of samples into account, which was suggested by Stimson et al [28]. A similar scenario was reported in South Africa, where interdistrict transmission of XDR-TB is thought to occur when patients acquire the infection in the district capital city and then return to their rural homes [29]. Although we obtained limited in-depth data due to the retrospective design and the high population mobility, we identified epidemiological links outside the household in 3 clusters, including 2 in schools and another in their workplace. Tuberculosis transmission in Shenzhen could also occur through casual contacts in marketplaces, transport vehicles, and other sites of social interactions [30], but this would be difficult to document and difficult to devise intervention strategies. However, our results suggest that targeted case finding in schools and workplaces would be more feasible and effective at reducing TB transmission, which, in the absence of an effective vaccine, is the only way to reduce TB incidence [31].

In summary, this study of MDR-TB in a new Chinese city with a large population of migrant workers found that transmission is predominantly citywide between patients who live distant from each other. Geographic areas at high risk of TB

transmission, including schools and workplaces, deserve increased vigilance in TB control.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Authors' contributions. Q. J., W. Y., and Q. G. designed and managed the study. Q. J., L. J., and L. M. performed the epidemiological investigation; J. L., Y. Z., and Z. Y. performed the experiments and collected the laboratory data; L. M. collected the clinical data; Q. J. cleaned the data and performed statistical analysis; Q. J. and C. Y. performed the spatial analysis; Q. L. and G.L performed the sequence analysis and interpretation. Q. J., Q. L., H. E. T., and Q. G. prepared the article. All authors contributed to and gave input to the final article.

Acknowledgments. The authors thank the tuberculosis public health teams in Shenzhen Center for Chronic Disease Control (CCDC). Epidemiological investigation was initiated only after receiving consent from the patients. The study protocol was approved by the Ethical Review Board of Shenzhen CCDC (no. 20180310).

Disclaimer. The views expressed in this paper do not necessarily reflect those of the funding bodies.

Financial support. This work was supported by the Sanming Project on Medicine in Shenzhen (grant number SZSM201611030 to W. T. and Q. G.) and the National Science and Technology Major Project of China (grant numbers 2017ZX10201302-006 and 2018ZX10715012-005 to Q. G.). This work also was supported by the Natural Science Foundation of China (grant numbers 91631301 and 81661128043 to Q. G.), the Robert E. Leet and Clara Guthrie Patterson Trust Award (2019 to C. Y.), and the MIDAS Center for Communicable Disease Dynamics (grant number U54 GM088558 to C. Y.).

Potential conflicts of interest. The authors report no potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. McBryde ES, Meehan MT, Doan TN, et al. The risk of global epidemic replacement with drug-resistant *Mycobacterium tuberculosis* strains. *Int J Infect Dis* 2017; 56:14–20.
2. Yang C, Luo T, Shen X, et al. Transmission of multidrug-resistant *Mycobacterium tuberculosis* in Shanghai, China: a retrospective observational study using whole-genome sequencing and epidemiological investigation. *Lancet Infect Dis* 2017; 17:275–84.
3. World Health Organization. Global tuberculosis report 2018. Geneva: World Health Organization, 2018.
4. Walker TM, Merker M, Knoblauch AM, et al; MDR-TB Cluster Consortium. A cluster of multidrug-resistant *Mycobacterium tuberculosis* among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. *Lancet Infect Dis* 2018; 18:431–40.
5. Lalor MK, Casali N, Walker TM, et al. The use of whole-genome sequencing in cluster investigation of a multidrug-resistant tuberculosis outbreak. *Eur Respir J* 2018; 51:1702313.
6. Bainomugisa A, Pandey S, Donnan E, et al. Cross-border movement of highly drug-resistant *Mycobacterium tuberculosis* from Papua New Guinea to Australia through Torres Strait protected zone, 2010–2015. *Emerg Infect Dis* 2019; 25:406–15.
7. National Bureau of Statistics. Survey on migrant workers 2017. Beijing: National Bureau of Statistics, 2018.
8. Chan KW. China: internal migration. In: Ness I, ed. The encyclopedia of global human migration. 2013; 2:980–95.
9. Shenzhen Statistics Bureau, NBS Survey Office in Shenzhen. Shenzhen statistical yearbook 2017. Beijing: China Statistics Press, 2017.
10. Hu X, Cook S, Salazar MA. Internal migration and health in China. *Lancet* 2008; 372:1717–9.

11. Yang C, Lu L, Warren JL, et al. Internal migration and transmission dynamics of tuberculosis in Shanghai, China: an epidemiological, spatial, genomic analysis. *Lancet Infect Dis* **2018**; 18:788–95.
12. The Ministry of Health, China Center for Disease Control and Prevention. Guidelines for the implementation of China's tuberculosis control program. Beijing: China Peking Union Medical College Press, **2008**.
13. Schiebelhut LM, Abboud SS, Gómez Daglio LE, Swift HF, Dawson MN. A comparison of DNA extraction methods for high-throughput DNA analyses. *Mol Ecol Resour* **2017**; 17:721–9.
14. Papaventsis D, Casali N, Kontsevaya I, Drobniowski F, Cirillo DM, Nikolayevskyy V. Whole genome sequencing of *Mycobacterium tuberculosis* for detection of drug resistance: a systematic review. *Clin Microbiol Infect* **2017**; 23:61–8.
15. Zelner JL, Murray MB, Becerra MC, et al. Identifying hotspots of multidrug-resistant tuberculosis transmission using spatial and molecular genetic data. *J Infect Dis* **2016**; 213:287–94.
16. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003**; 13:2498–504.
17. Steingart KR, Henry M, Ng V, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis* **2006**; 6:570–81.
18. Glynn JR, Vynnycky E, Fine PE. Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. *Am J Epidemiol* **1999**; 149:366–71.
19. Yang C, Luo T, Sun G, et al. *Mycobacterium tuberculosis* Beijing strains favor transmission but not drug resistance in China. *Clin Infect Dis* **2012**; 55:1179–87.
20. Liu Q, Ma A, Wei L, et al. China's tuberculosis epidemic stems from historical expansion of four strains of *Mycobacterium tuberculosis*. *Nat Ecol Evol* **2018**; 2:1982–92.
21. Kay GL, Sergeant MJ, Zhou Z, et al. Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nat Commun* **2015**; 6:6717.
22. Alland D, Kalkut GE, Moss AR, et al. Transmission of tuberculosis in New York City: an analysis by DNA fingerprinting and conventional epidemiologic methods. *N Engl J Med* **1994**; 330:1710–6.
23. Wu X, Pang Y, Song Y, et al. Implications of a school outbreak of multidrug-resistant tuberculosis in northern China. *Epidemiol Infect* **2018**; 146:584–8.
24. Ngamvithayapong-Yanai J, Luangjina S, Thawthong S, Bupachat S, Imsangaun W. Stigma against tuberculosis may hinder non-household contact investigation: a qualitative study in Thailand. *Public Health Action* **2019**; 9:15–23.
25. Wang XD. Tuberculosis outbreak reported at Hunan middle school. 2017-11-17. Available at: http://www.chinadaily.com.cn/china/2017-11/17/content_34664648.htm. Accessed 6 April 2019.
26. National Health Commission, Ministry of Education of the People's Republic of China. Standards for TB prevention and control in schools (2017). Beijing: National Health Commission, **2017**.
27. Cudahy PGT, Andrews JR, Bilinski A, et al. Spatially targeted screening to reduce tuberculosis transmission in high-incidence settings. *Lancet Infect Dis* **2019**; 19:e89–95.
28. Stimson J, Gardy J, Mathema B, Crudu V, Cohen T, Colijn C. Beyond the SNP threshold: identifying outbreak clusters using inferred transmissions. *Mol Biol Evol* **2019**; 36:587–603.
29. Nelson KN, Shah NS, Mathema B, et al. Spatial patterns of extensively drug-resistant tuberculosis transmission in KwaZulu-Natal, South Africa. *J Infect Dis* **2018**; 218:1964–73.
30. Murray EJ, Marais BJ, Mans G, et al. A multidisciplinary method to map potential tuberculosis transmission “hot spots” in high-burden communities. *Int J Tuberc Lung Dis* **2009**; 13:767–74.
31. Vesga JF, Hallett TB, Reid MJA, et al. Assessing tuberculosis control priorities in high-burden settings: a modelling approach. *Lancet Glob Health* **2019**; 7:e585–95.