

# Whole-Genome Sequencing and Epidemiological Investigation of Tuberculosis Outbreaks in High Schools in Hunan, China

Zuhui Xu<sup>1,2,\*</sup>, Haican Liu<sup>3,\*</sup>, Yanping Liu<sup>4,\*</sup>, Yi Tang<sup>2</sup>, Yunhong Tan<sup>2</sup>, Peilei Hu<sup>2</sup>, Chuanfang Zhang<sup>2</sup>, Chongguang Yang<sup>4</sup>, Kanglin Wan<sup>3</sup>, Qiaozhi Wang<sup>2</sup>

<sup>1</sup>Xiangya School of Public Health, Central South University, Changsha, 410078, People's Republic of China; <sup>2</sup>Tuberculosis Control Institute of Hunan Province, Changsha, 410013, People's Republic of China; <sup>3</sup>State Key Laboratory of Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, 102206, People's Republic of China; <sup>4</sup>School of Public Health (Shenzhen), Shenzhen Campus of Sun Yat-sen University, Shenzhen, 518000, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Qiaozhi Wang, Department of Institute office, Tuberculosis Control Institute of Hunan Province, No. 519 of Xianjiahu Road, Changsha, 410013, People's Republic of China, Tel/fax +86073188809748, Email wangqz664@163.com; Kanglin Wan, State Key Laboratory of Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, No. 155 of Changbai Road, Changping District, Beijing, 102206, People's Republic of China, Tel +86 13910065264, Email wankanglin@icdc.cn

**Background:** Tuberculosis (TB) seriously threatens individual and public health. Recently, TB outbreaks in schools have been reported more frequently in China and have attracted widespread attention. We reported three TB outbreaks in high schools in Hunan Province, China.

**Methods:** When a tuberculosis patient was reported in a school, we carried out field epidemiological investigations, including tuberculin skin testing (TST), chest X-ray (CXR) and laboratory test for all close contacts, and whole-genome sequencing (WGS) analyses to understand the transmission patterns, the causes and the risk factors for the outbreaks, thereby providing a foundation for the control of TB epidemics in schools.

**Results:** A total of 49 students with TB patients were identified in the three schools where TB outbreaks occurred, including nine patients in School A, 14 patients in School B, and 26 patients in School C. In Schools A, B and C, the putative attack rates in the classes of the index case were 13.8% (8/58), 7.6% (5/66), and 40.4% (21/52), while the putative attack rates of expanding screening in the school were 0.3% (1/361), 0.2% (9/3955), and 0.2% (5/2080), respectively. Thirteen patients had patient delay, with a median delay interval of 69 days (IQR 30.5–113 days). Twelve patients had a healthcare diagnostic delay with a median delay interval of 32 days (IQR 24–82 days). Phylogenetic analysis of culture-positive patients revealed that most of them shared a small genetic distance ( $\leq 12$  SNPs), with three separate genetic clusters (including one MDR-TB genomic cluster), indicating the recent transmission of *Mycobacterium tuberculosis* strains.

**Conclusion:** This combination of field investigation and WGS analysis revealed the transmission of three TB outbreaks in schools. Reinforced implementation is needed to improve timely case finding and reduce diagnosis delay in routine TB control in the school population.

**Keywords:** tuberculosis outbreak, high school, epidemiological investigation, whole genome sequencing

## Background

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* (*M. tuberculosis*) is still one of the major global public health problems.<sup>1,2</sup> There were approximately 842,000 new TB patients in China in 2020, and the estimated incidence rate was 59/100,000.<sup>3</sup> Children and adolescent populations are often overlooked in terms of TB incidence and their access to TB care. In 2019, 396,000 children and adolescents aged 10–19 years were reported with TB, accounting for 10% of total

notifications in 95 countries globally.<sup>4</sup> TB outbreaks in schools have been reported in other countries, such as Japan, Mongolia, Korea, Serbia, and Swaziland.<sup>5–8</sup> Despite differences in TB burden and outbreak preparedness, the affected schools were confronted with similar challenges including delayed diagnosis of index cases, lack of experienced medical personnel, a lack of sustained financial support, and difficulty in responding to media and community attention.<sup>6</sup>

It has been increasingly recognized that adolescents and school-aged people are vulnerable to TB infection.<sup>11</sup> A study showed that there were 39,198 TB patients among students in China, accounting for 4.12% of the total TB burden, and the most common outbreak sites were high schools.<sup>12</sup> TB is easily spread within schools due to the large population density, close contact, and poor ventilation in classrooms and dormitories.<sup>13</sup> The outbreak of TB in schools affects the students' health and can even cause major public health events.<sup>13</sup> However, despite the shared common epidemiological links for TB patients in school outbreaks, it is unclear whether these TB patients were direct transmission events in school or acquired from the community outside the school.

Recently, whole-genome sequencing (WGS) has played a vital role in transmission inference<sup>9,14</sup> of *M. tuberculosis*. Since the traditional epidemiological investigation heavily relies on the disease history of close contact, it is difficult to clarify the transmission chain of TB due to the natural history of TB disease. WGS is considered an ultimate genotyping tool for TB outbreak investigations, which provides a high resolution in determining clusters for transmission analysis compared to conventional methods such as mycobacterial interspersed repetitive unit-variable number tandem repeat typing (MIRU-VNTR).<sup>15,16</sup> Recent studies have shown that WGS was applied in TB outbreaks to find the source and determine the transmission route in communities in China.<sup>17,18</sup>

Hunan Province, located in south-central China, is one of the provinces with a high TB burden in China,<sup>9</sup> with an estimated notification of a TB incidence rate of 74 patients per 100,000 population. From 2012 to 2017, 7940 students with TB were notified in Hunan Province, with a registered incidence rate of 13.2 per 100,000 population, suggesting the relative high school tuberculosis epidemic of Hunan Province.<sup>10</sup> Here, we reported the epidemiology study of three TB outbreaks in schools of Hunan Province in 2017–2019, using WGS analysis combined with the field epidemiological investigation to better understand the transmission characteristics and main influence factors of these outbreaks, thereby providing evidence for the prevention and control of TB epidemics in schools.

## Methods

### Study Design and Settings

According to the diagnostic criteria of pulmonary tuberculosis issued by the National Health and Family Planning Commission of China,<sup>19</sup> the diagnosis of pulmonary tuberculosis is mainly based on etiological (including bacteriology and molecular biology) examination, combined with epidemiological history, clinical symptoms, chest radiologic evidence, and related auxiliary examinations. In this study, the three TB outbreaks occurred in three high schools in the south, central and north of Hunan Province from 2017 to 2019. When we received a report of a student TB patient (index case), we immediately carried out a field epidemiological investigation in the school. We investigated the index case to identify the close contacts, and then conducted TB screening for all contacts. Once an additional new TB patient was identified, the field investigation would expand to the whole floor or school. We further conducted WGS to elucidate the transmission patterns of these outbreaks.

### Epidemiological Investigation

An epidemiological investigation was conducted among the index case and their close contacts. The index case was defined as the first identified TB patient in each school.<sup>17</sup> Close contacts refer to those who have direct contact with index cases, and the first-round screening of close contacts mainly included teachers and students in the same class and dormitory. Systematic TB screening was performed by clinical evaluation, tuberculin skin testing (TST), chest radiography (CXR), and laboratory test for all close contacts. The laboratory test included sputum smear and culture, drug susceptibility test was further conducted if culture positive.<sup>20</sup> If one or more patients of TB were newly discovered in the contact screening, the second round of screening of the close contacts would be expanded to the floor and/or the building of the class and/or dormitory through the school.<sup>20</sup> In addition, the family member contacts would be referred to the local

hospital for screening. Each TB patient was interviewed using a standardized questionnaire by the local Centers for Disease Control and Prevention (CDC) staff ([Additional file 1](#)). The questionnaire included individual information, medical history, exposure history, symptoms of the first onset of illness (cough, expectoration, hemoptysis, low fever, night sweats, fatigue, etc.), the reasons for patient delay and healthcare diagnostic delay, and clinical treatment history. The patient was judged to have a patient delay if the interval between the date of the first symptom and the first visit (doctor) exceeds two weeks.<sup>21</sup> Similarly, healthcare diagnostic delay is defined as more than two weeks between the date of the first health facility visit and the confirmation of a TB diagnosis.<sup>22</sup>

## Sputum Specimens and Laboratory Testing

Sputum specimens were collected from all the patients for the investigation, including routine sputum smears, bacterial isolation, and culture. The *M. tuberculosis* isolates used for WGS were obtained from 15 culture-positive TB patients of the three schools, including one strain of a patient's family member. The standard strain H37Rv was provided by the Tuberculosis Laboratory of the National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention. The number of culture-positive isolates from each school was as follows: School A, 4; School B, 6; School C, 5.

## Bacterial Isolation, Drug Susceptibility Testing, and Resistance Prediction

The sputa were processed by the standard N-acetyl L-cysteine (NALC)/NaOH decontamination method. The decontaminated sediments were resuspended in 2.0 mL of phosphate-buffered saline and 0.5 mL was inoculated into Bactec MGIT 960 tubes. Tubes flagged positively by the MGIT 960 instrument were subjected to further *M. tuberculosis* complex identification through the commercial MPT64 immunochromatographic test (GENESIS, Hangzhou, China).

Drug susceptibility tests (DSTs) for Rifampin (RIF) and Isoniazid (INH) resistance were performed according to the manufacturer's instructions (Becton, Dickinson and Company, Sparks, MD) with the following drug concentrations: RIF 1.0 µg/mL and INH 0.1 µg/mL. The in-silico drug resistance was predicted using TB-Profler v2.8.14 and based on genomic sequencing data, including RIF, INH, Amikacin (AMK), Capreomycin (CPM), Ciprofloxacin (CIP), Ethambutol (EMB), Kanamycin (KM), Moxifloxacin (MXF), Ofloxacin (OFLX), Pyrazinamide (PZA), and Streptomycin (SM), Bedaquilin (BDQ) and Delamanid.

## Whole-Genome Sequencing Analysis

Genomic DNA of *M. tuberculosis* isolates was extracted using the CTAB method for sequencing as previously described.<sup>23,24</sup> DNA libraries were constructed with genomic DNA using kits provided by Illumina according to the manufacturer's instructions. The average sequencing depth of the genome is 124.8, and the coverage is 99.3%. DNA libraries were then selected to perform cluster growth and 150 bp paired-end sequencing on an Illumina HiSeq 2500 according to standard protocols. Paired-end reads were mapped to the reference genome H37Rv (GenBank accession number, NC\_000962.3) with Bowtie2. The SAMtools (version 1.6) and VarScan (version 2.3.9) suite were used to define SNPs, with low-quality SNPs (Phred score  $Q < 20$  and read depth  $< 5$ ) and sites with missing calls in  $> 10\%$  of isolates were removed. Heterogenous sites were called the consensus allele if present in  $\geq 80\%$  of mapped reads. SNPs in repetitive regions, PE/PPE genes and in resistance-conferring genes were excluded from further phylogenetic analysis.

The alignment of concatenated SNPs of the 15 strains was used to construct a maximum-likelihood (M-L) phylogenetic tree with Mega X,<sup>25,26</sup> using the "GTR" nucleotide substitution model with 500 bootstrap samples. We compared the pairwise genomic distance and defined a genomic cluster as the genetic distance of strains that were no more than 12 SNPs.<sup>27,28</sup>

## Ethics Approval and Consent to Participate

Ethical approval was obtained from the Research Ethics Committee of the Hunan Provincial Chest Hospital (protocol number: KLS2019092501). The informed consent form was obtained from all study participants aged  $\geq 18$  years or their parents or legal guardians if  $< 18$  years before enrollment, and ethical principles conformed to the Declaration of Helsinki.

## Statistical Analysis

Statistical analyses were performed using the SPSS statistical package (version 21.0). Continuous variables are reported as the mean  $\pm$  standard deviation, and nonnormally distributed variables are reported as the median (quartile range). Categorical variables are shown as the frequency (%). Categorical variables were compared by the  $\chi^2$  test or Fisher's exact test as appropriate.  $P < 0.05$  was regarded as statistically significant.

## Results

### Index Case Investigation

In December 2017, a 16-year-old male student in School A went to the county's central hospital because of cough and expectoration for two weeks. He was diagnosed as a rifampin-resistant patient, and received standard treatment at Hunan Provincial Chest Hospital. The sputum smear was negative, but the sputum culture was positive. The time interval from the first symptom to definitive diagnosis was 46 days, and the reason he self-reported for the delay in seeking medical treatment was the stigma of TB and to conceal the disease (he suspected he had TB before visiting doctors). We also found his mother was an MDR patient, indicating a potential household transmission.

In February 2019, a 16-year-old male student in the 11<sup>th</sup> grade of School B was reported to have TB. He had symptoms of cough, expectoration, low fever, and fatigue for weeks before the TB diagnosis. At that time, he made a common neglect for consideration of a cold and received no treatment because of the lacking knowledge of TB. On February 25, because of the unrelieved symptoms, he went to the hospital to seek further care and was diagnosed with TB followed by standardized anti-TB treatment. The sputum smear was negative but the sputum culture was positive, and small cavities were found in his lung. The time interval from the first symptom to definitive diagnosis was 37 days.

In February 2018, a 17-year-old male student in School C was diagnosed with TB. Because of the lack of the knowledge of TB symptom, he did not pay enough attention to the repeated expectoration and cough since July 2017. Until February 6, 2018, he went to the emergency department of the county hospital for hemoptysis caused by a collision with the chest and was diagnosed with TB. Then he underwent standardized anti-TB treatment in the hospital. Small cavities were found in his lung with sputum smear-negative and sputum culture-negative. The time interval from the first symptom to definitive diagnosis was 247 days.

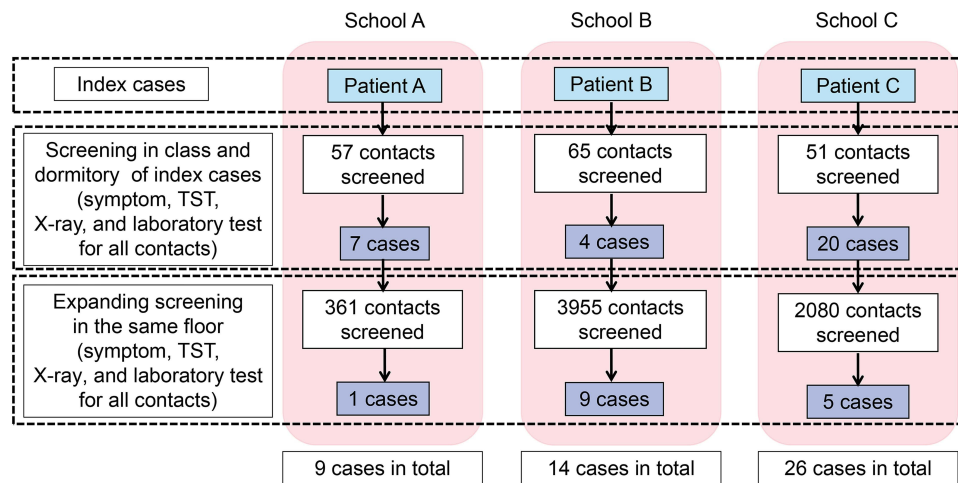
### Close Contact Investigation

A total of 6569 close contacts of the index cases were identified, including 418 from School A, 4020 from School B, and 2131 from School C. In the first-round screening for close contacts, 31 TB patients were identified. In addition to the index case, there were 8 (13.8%, 8/58), 5 (7.6%, 5/66), and 21 (40.4%, 21/52) patients in the class of index cases in School A, School B, and School C. In the second round of screening (except for the first round screened), a total of 15 TB patients were identified, including 1 (0.3%, 1/361) in School A, 9 (0.2%, 9/3955) in School B, and 5 (0.2%, 5/2080) in School C, respectively. In total, 49 new TB patients were identified during the two rounds of TB screening among the school population, including 20 pathogen-positive patients (with microbiological findings) and 29 pathogen-negative patients. The survey procedure is shown in [Figure 1](#). The putative attack rates of the classes and schools of index cases are shown for each outbreak in [Table 1](#). There was a significant difference in the putative attack rates between the class and the school ([Tables 1 and 2](#)).

### Epidemiological Characteristics

This investigation identified a total of 49 TB patients in the three schools, with which three school TB outbreaks were identified. There were 36 (73.5%) male and 13 (26.5%) female patients. Except for one patient who was a 31-year-old female teacher, the others were all students with a median age of 17 years (IQR 17–18). Most TB patients (34, 69.4%) were in the 12<sup>th</sup> grade. Thirty-four patients (69.4%) were in the same class as the index cases, and 20 (40.8%) patients were bacterially positive. The epidemiological characteristics of the 49 patients diagnosed with TB are shown in [Table 2](#).

The symptom distributions of TB patients are shown in [Figure 2](#). Among 49 patients, 28 (57.1%) had TB-like symptoms and the others showed no related symptoms. Among these 28 patients, the most common symptoms were



**Figure 1** Diagram of the close contact survey process. Laboratory test: sputum smear and culture for suspects, DST also conducted if culture positive.

cough and expectoration, of which the incidence was 85.7% and 53.6%, respectively, and the following symptoms, such as low fever, fatigue and night sweats, were 10.7%, 10.7% and 7.1%, respectively.

A total of 13 patients had patient delay, with a median delay interval of 69 days (IQR 30.5–113 days). Among them, eight patients had a patient delay because they were unaware of the connection between the symptoms of cough or expectoration with TB. A total of 12 patients had a healthcare diagnostic delay, with a median delay interval of 32 days (IQR 24–82 days). The main reasons for these delays were that the patients were misdiagnosed with bronchitis and other non-tuberculosis diseases during healthcare seeking.

## Drug Resistance Profiles

Four isolates from School A had phenotypic rifampin-resistance, and two of these isolates were also isoniazid-resistant (ie, multidrug-resistant tuberculosis, MDR-TB). The patients in School B and School C were all phenotypically susceptible to rifampin and isoniazid. Based on WGS-predicted genetic drug resistance profiles, the isolates from School B and School C were pan-susceptible. MDR-TB and rifampin-resistant tuberculosis (RR-TB) were identified among three isolates from patients in School A, involving two students and one family member. We further identified that at least three patients in school A had strains harboring the second-line injectable anti-TB drugs (*rrs*, A1401X) (Figure 3 and Additional file 2), indicating the direct transmission of drug resistant TB both in households and schools. The comparisons between the DST and in-silico predicted drug resistance was shown in Additional file 3.

## Transmission Inference

We conducted genomic and phylogenetic analyses based on the *M. tuberculosis* strains collected from 15 culture-positive patients from the three schools. All these *M. tuberculosis* strains were identified as the sublineage L2.2.1 (subgroup of the Beijing family strain), except one strain belonging to lineage 4.1 (Figure 3).

Aligning reads against the reference strain revealed 1803 SNPs that were used to reconstruct an M-L phylogeny tree. The M-L clearly showed that each school formed separate clusters, while all the pairwise genetic distances within each cluster were less than ten SNPs, with a median distance of four SNPs, indicating the recent transmission and spread of these outbreaks in a short period. However, not all the new patients notified during expanded screening belonged to the outbreak chain, as one TB patient had an isolate with more than 100 SNPs from the other strains in the same school (School A), which was unlikely to belong to the same transmission chain and suggested a different route of infection. The pairwise SNP distances between 15 isolates and H37Rv are shown in Table 3. The cluster of School A involved a mother and her son, indicating both household and school transmission events of MDR-TB strains. The strains from six students in School B and five in School C formed two separated clusters (Figure 3 and Additional File 4). Based on the SNP distance matrix and reconstructed phylogeny trees, we confirmed that these three school TB outbreaks were independent

**Table 1** The Status of Contact Survey of Tuberculosis Outbreaks in Three Schools

	Index Case						No. of First Round Screening in Class and Dormitory of the Index Case	Putative Attack Rate in the Class* n (%)	No. of Expanding Screening in the Same Floor <sup>#</sup>	Putative Attack Rate of Expanding Screening n (%)	P value (Putative Attack Rate in the Class VS in the School)
	Sputum Smear	Sputum Culture	Cavity	First Symptom	Time Interval from First Symptom to Definitive Diagnosis (d)	Reason for Finding Delay					
School A	Negative	Positive	No	Cough, Expectoration	46	Conceal the disease	57	8 (13.8)	361	1 (0.3)	P< 0.01
School B	Negative	Positive	Yes	Cough, Expectoration, Hemoptysis	37	Delay in seeking medical treatment	65	5 (7.6)	3955	9 (0.2)	P< 0.01
School C	Negative	Negative	Yes	Cough, Expectoration, Hemoptysis	247	Delay in seeking medical treatment	51	21 (40.4)	2080	5 (0.2)	P< 0.01

**Notes:** \*Putative attack rate in the class = No. of cases in class and dormitory of the index cases (including the index case) /No. of students in the class and dormitory of the index case. <sup>#</sup>The expanding screening on the same floor was conducted for the persons except in the first round.

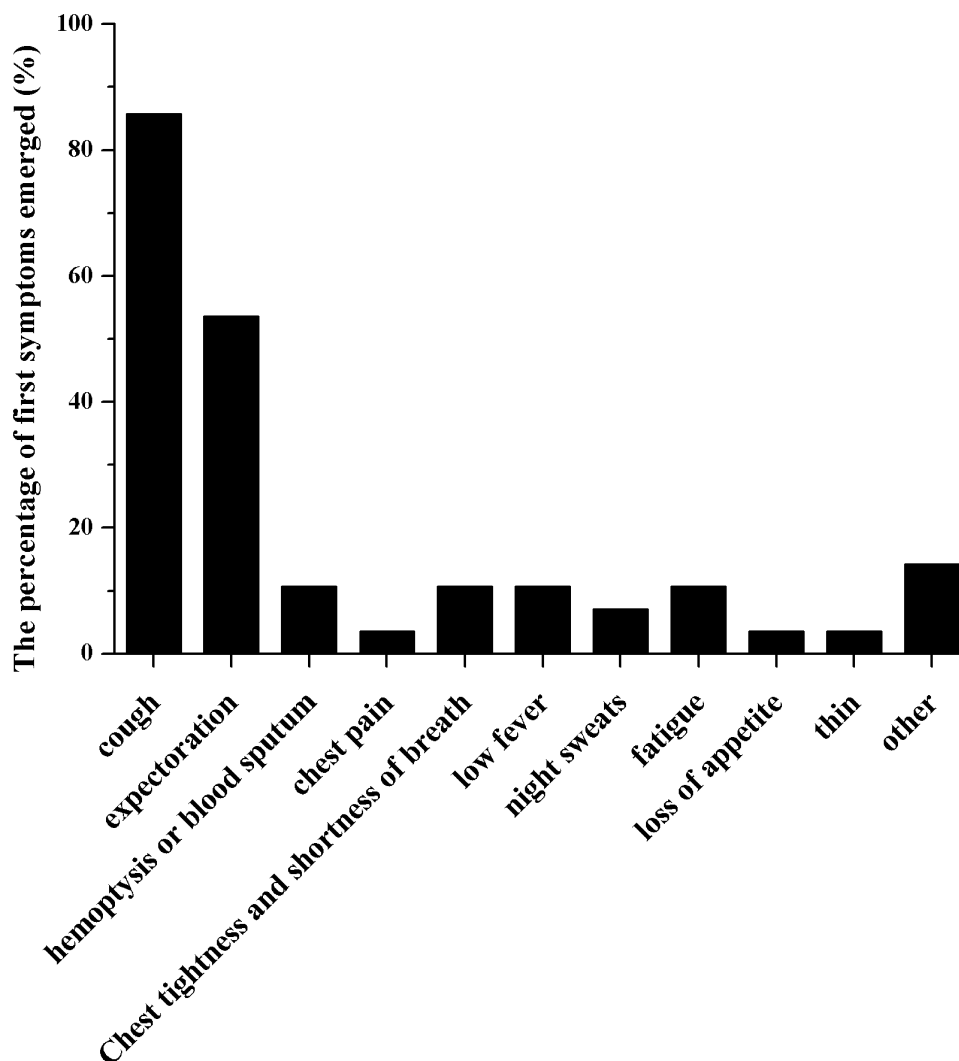
**Table 2** Demographics and Clinical Characteristics of the TB Cases (N = 49)

Characteristics	School A n (%)	School B n (%)	School C n (%)	Total n (%)	P value
Gender					
Male	9(100.0)	8(57.1)	19(73.1)	36(73.5)	0.07
Female	0(0.0)	6(42.9)	7(26.9)	13(26.5)	
Age					
16	0(0.0)	7(50.0)	1(3.8)	8(16.3)	< 0.01
17	4(44.4)	7(50.0)	12(46.2)	23(46.9)	
18	5(55.6)	0(0.0)	12(46.2)	17(34.7)	
31	0(0.0)	0(0.0)	1(3.8)	1(2.0)	
Occupation					
Student	9(100.0)	14(100.0)	25(96.2)	48(98.0)	1.00
Teacher	0(0.0)	0(0.0)	1(3.8)	1(2.0)	
Grade					
10th	0(0.0)	0(0.0)	1(3.8)	1(2.0)	< 0.01
11th	0(0.0)	13(92.9)	1(3.8)	14(28.6)	
12th	9(100.0)	1(7.1)	24(92.3)	34(69.4)	
The same class or dormitory with the index cases					
Yes	8(88.9)	5(35.7)	21(80.8)	34(69.4)	0.009
No	1(11.1)	9(64.3)	5(19.2)	15(30.6)	
Living in dormitory					
Yes	1(11.1)	3(21.4)	3(11.5)	7(14.3)	0.857
No	8(88.9)	11(78.6)	23(88.5)	42(85.7)	
Contact history					
Yes	1(11.1)	0(0.0)	1(3.8)	2(4.1)	0.414
No	8(88.9)	14(100.0)	25(96.2)	47(95.9)	
TST					
≥ 10mm	4(44.4)	3(21.4)	17(65.4)	24(49.0)	0.031
< 10mm	5(55.6)	11(78.6)	9(34.6)	25(51.0)	
Cavity					
Yes	3(33.3)	5(35.7)	2(7.7)	10(20.4)	0.05
No	6(66.7)	9(64.3)	24(92.3)	39(79.6)	
Pathogen positive for AFB logy					
Yes	7(77.8)	7(50.0)	6(23.1)	20(40.8)	0.012
No	2(22.2)	7(50.0)	20(76.9)	29(59.2)	
RIF-resistant					
Yes	5(55.6)	0(0.0)	0(0.0)	5(10.2)	< 0.01
No	3(33.3)	7(50.0)	5(19.2)	15(30.6)	
Not detected	1(11.1)	7(50.0)	21(80.8)	29(59.2)	

events (i.e., SNP distance >100 between any cases from two schools, [Table 3](#) and [Additional File 4](#)). Overall, the WGS analysis confirmed the TB outbreak among these three schools and differentiated transmission events from non-school sources, which were more likely from the household.

## Discussion

The current study reported three outbreaks of TB in high schools in Hunan Province, China, which were confirmed by epidemiological investigation and WGS analysis. A total of 6569 students and staff in schools were screened by TST and CXR examinations, and 49 TB patients were identified by laboratory confirmation and/or clinical diagnosis, including one teacher and 48 students. Whole-genome sequencing analysis confirmed the recent transmission of TB among students and teachers, but also revealed that not all of the screened new TB patients were involved in the school

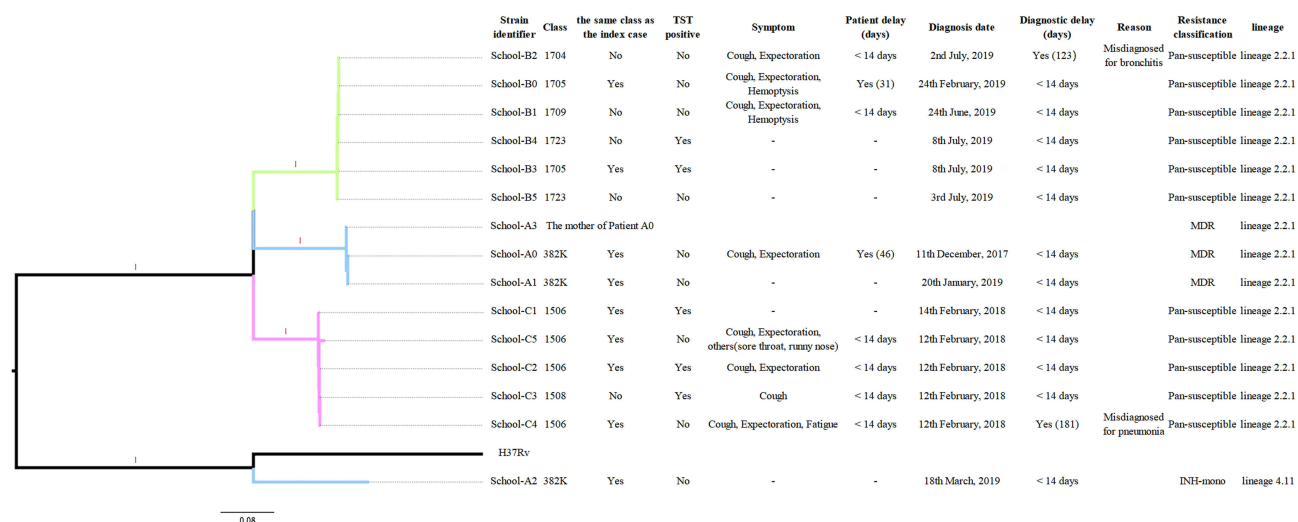


**Figure 2** The first symptom proportion of the patients.

transmission chain. Instead, the reason might be singletons and probably resulted from the infection events outside the school.

Students in high school are one of the vulnerable and high-risk population groups to TB in which it is easy to develop and spread infectious diseases and even cause outbreaks. Reports of TB outbreaks in schools have increased recently in China, including a previous report of a TB outbreak of 90 students in a high school in the same province.<sup>7,29–31</sup> In 2017, the National Health and Family Planning Commission and the Ministry of Education of China released a guideline to strengthen the early diagnosis and treatment of pulmonary TB in all levels of schools.<sup>20</sup> High school education is the most important stage as students need to take college entrance examinations.<sup>32</sup> They usually face lots of pressure to study and reduce time in exercise, which could lead to weakened immunity.<sup>33</sup> Furthermore, they often attach little importance to upper respiratory tract infection symptoms, such as cough and expectoration, causing a patient delay. All the index cases experienced a long diagnosis delay, and the time intervals of patient delay were also relatively long in the three schools (69 days, IQR 30.5–113 days). In general, the student TB patients (both index cases and secondary patients) were more likely to overlook TB-like symptoms such as cough and expectoration. Another explanation for diagnosis delay could be the misdiagnosis of non-TB diseases such as bronchitis. For example, the TB outbreak in School C was caused for this reason. In summary, reinforcing routine education regarding TB control knowledge is a necessary step for high schools in China.





**Figure 3** Maximum-likelihood tree of 15 strains annotated with epidemiological characteristics related to drug resistance. Branches are colored by different schools. MDR, rifampicin and isoniazid resistant. INH-mono, rifampicin susceptible and isoniazid resistant. Pan-susceptible, rifampicin and isoniazid susceptible. Patients A0 and B0 were index cases.

We conducted a drug susceptibility assay on all collected isolates in the three schools and found that only isolates from School A showed resistance to rifampin, isoniazid, and second-line injectable anti-TB drugs. We found the index patient of School A with a previous history of TB, and his mother was also an MDR-TB patient. Therefore, the index patient might have been infected with *M. tuberculosis* by his mother and caused the school transmission, which was verified by WGS. The hide of disease of the student contributed to the spread of drug-resistant TB in School A.

WGS has shown a high resolution in determining TB transmission dynamics in many settings worldwide.<sup>27,34,35</sup> WGS analysis has advanced power in measuring genetic relatedness between strains compared to other PCR-based genotyping tools, which have been approved to study TB transmission in strains circulating in China.<sup>27,36</sup> The ability to determine the transmission relationship is important in TB outbreaks in schools since the outbreak occurrences of infectious disease in the student population are always in high tension. Furthermore, the utility of the WGS tool can differentiate unrelated transmission from school outbreaks despite common epidemiological links (e.g., same class or dormitory). A study in Guangxi Province, China, also showed that previous exposure to TB in the household led a student to catch TB at school, and then spread it to his classmates.<sup>37</sup> It emphasized that the source of school TB outbreaks might not be in the school but be infected outside the school. Again, this finding highlighted the importance of improving the identification of TB among adolescents. The application of WGS has gained insight into the investigation of TB outbreaks in schools and confirmed transmission inference with traditional epidemiological investigation.

There were several limitations in our study. First, not all the TB patients had culture-positive isolates. The small number of *M. tuberculosis* isolates from TB patients limited the role of WGS in transmission inference, including the transmission chain analysis. We also cannot fully rule out the possibility of recent transmission of the singleton strain in our study. Second, the study's relatively short duration limited the identification of more TB patients that might involve in the transmission chain but had not developed TB based on the nature of the disease. Despite these limitations, our study highlighted the critical role of epidemiological investigation and WGS in the transmission of school TB.

## Conclusions

In summary, this study combined the field investigation and the WGS analysis and revealed at least three independent TB outbreaks in the schools. Patients' diagnosis delay mainly contributed to the recent transmission of TB in this vulnerable population. Furthermore, the identification of the MDR-TB cluster involving family and school links reinforced the importance of TB control in both general and school populations. The TB outbreaks among senior high school students also raise more concern on the awareness of TB-like symptoms during routine healthcare activities, which can benefit TB control in the early stage of the epidemic and prevent the further spread in school populations.

**Table 3** The Pairwise SNP Distance Matrix of the 15 School TB Isolates and Reference Strain

	H37Rv	School-C1	School-C2	School-C3	School-C4	School-C5	School-A0	School-A1	School-A2	School-A3	School-B0	School-B1	School-B2	School-B3	School-B4	School-B5
H37Rv	0															
School-C1	1135	0														
School-C2	1140	5	0													
School-C3	1136	3	4	0												
School-C4	1136	5	4	2	0											
School-C5	1120	19	20	18	18	0										
School-A0	1205	370	373	373	373	387	0									
School-A1	1206	373	374	376	374	388	3	0								
School-A2	656	1030	1035	1031	1031	1025	1098	1099	0							
School-A3	1204	369	372	372	372	386	7	10	1097	0						
School-B0	1134	351	352	352	350	364	413	414	1025	410	0					
School-B1	1137	350	349	351	351	365	412	413	1028	409	3	0				
School-B2	1136	351	350	352	352	366	413	414	1027	412	4	3	0			
School-B3	1128	357	356	356	356	372	419	420	1023	418	12	9	10	0		
School-B4	1133	354	353	353	351	367	416	415	1026	413	5	4	7	7	0	
School-B5	1134	351	350	350	350	366	413	414	1025	412	6	3	4	6	3	0

**Note:** The red-white-blue color scale represents the pairwise SNP distance from small to large.

## Abbreviations

TB, tuberculosis; TST, tuberculin skin testing; CXR, chest X-ray; WGS, whole-genome sequencing; *M. Tuberculosis*, *Mycobacterium tuberculosis*; MIRU-VNTR, mycobacterial interspersed repetitive unit-variable number tandem repeat typing; CDC, Centers for Disease Control and Prevention; DST, drug susceptibility tests; RIF, rifampin; INH, isoniazid; AMK, amikacin; CPM, capreomycin; CIP, ciprofloxacin; EMB, ethambutol; KM, kanamycin; MFX, moxifloxacin; OFLX, ofloxacin; PZA, pyrazinamide; SM, streptomycin; BDQ, bedaquilin; IQR, interquartile range; M-L, maximum-likelihood; MDR-TB, multidrug-resistant tuberculosis; RR-TB, rifampin-resistant tuberculosis.

## Data Sharing Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Acknowledgment

We thank all the investigators from the study sites for their contribution.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This research was supported by the Natural Science Foundation of Hunan Province, China (grant number: 2018JJ2215).

## Disclosure

The authors declare that there are no conflicts of interest.

## References

1. Chen J, Chen L, Zhou M., et al. Transmission of multidrug-resistant tuberculosis within family households by DTM-PCR and MIRU-VNTR genotyping. *BMC Infect Dis.* 2022;22(1):192. doi:10.1186/s12879-022-07188-7
2. Ohiengbomwan OT, Komolafe IO, Alayande S, Njor BE, Onisile DF, Oguzie J. Cross-sectional community-based assessment of knowledge, attitude and practices on tuberculosis in Osun State, South-west, Nigeria. *Health Soc Care Community.* 2022;30(5). doi:10.1111/hsc.13762
3. World Health Organization. Global tuberculosis report 2021. World Health Organization; 2021.
4. World Health Organization. Global tuberculosis report 2020. World Health Organization; 2020.
5. Itaki M, Endo M, Ikedo K, et al. A multidrug-resistant tuberculosis outbreak in a language school: tokyo, Japan, 2019–2020. *Int J Mycobacteriol.* 2021;10(1):37–42.
6. Rahevar K, Yuen T, Oh KH, et al. Tuberculosis outbreaks in schools: experiences from the Western Pacific Region. *Western Pac Surveill Response J.* 2021;12(1):1–5. doi:10.5365/wpsar.2020.11.3.005
7. Stosic MB, Plavska D, Mavroević N, et al. Tuberculosis outbreak among high school students in Novi Pazar, Serbia 2016: a retrospective-cohort study. *J Infect Dev Ctries.* 2019;13(2):101–110. doi:10.3855/jidc.10952
8. Ustero PA, Kay AW, Ngo K, et al. School and household tuberculosis contact investigations in Swaziland: active TB case finding in a high HIV/TB burden setting. *PLoS One.* 2017;12(6):e0178873. doi:10.1371/journal.pone.0178873
9. He W, Tan Y, Liu C, et al. Drug-resistant characteristics, genetic diversity, and transmission dynamics of rifampicin-resistant *Mycobacterium tuberculosis* in Hunan, China, revealed by whole-genome sequencing. *Microbiol Spectr.* 2022;10(1):e0154321. doi:10.1128/spectrum.01543-21
10. Zhang C, Tang Y, Xu Z, Xiao T, Wang Q. Epidemiological characteristics of pulmonary tuberculosis among students in Hunan Province from 2012 to 2017. *Chinese J Infect Control.* 2018;17(11):1008–1012.
11. Snow KJ, Sismanidis C, Denholm J, Sawyer SM, Graham SM. The incidence of tuberculosis among adolescents and young adults: a global estimate. *Eur Respir J.* 2018;51(2):1702352. doi:10.1183/13993003.02352-2017
12. Chen W, Chen Q, Xia Y, Cheng S. Analysis on the characteristics of the national tuberculosis epidemic among students from 2008 to 2012. *Chin J Antituberculosis.* 2013;35(12):949–954.
13. Bao H, Liu K, Wu Z, et al. Tuberculosis outbreaks among students in mainland China: a systematic review and meta-analysis. *BMC Infect Dis.* 2019;19(1):972. doi:10.1186/s12879-019-4573-3
14. Merker M, Egbe NF, Ngangue YR, et al. Transmission patterns of rifampicin resistant *Mycobacterium tuberculosis* complex strains in Cameroon: a genomic epidemiological study. *BMC Infect Dis.* 2021;21(1):891. doi:10.1186/s12879-021-06593-8

15. Roetzer A, Diel R, Kohl TA, et al. Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study. *PLoS Med*. 2013;10(2):e1001387. doi:10.1371/journal.pmed.1001387
16. Walker TM, Ip CL, Harrell RH, et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet Infect Dis*. 2013;13(2):137–146. doi:10.1016/S1473-3099(12)70277-3
17. Li H, Liu C, Liang M, et al. Tuberculosis outbreak in an educational institution in Henan Province, China. *Front Public Health*. 2021;9:737488. doi:10.3389/fpubh.2021.737488
18. Yu H, Zhang Y, Chen X, et al. Whole-genome sequencing and epidemiological analysis of a tuberculosis outbreak in a high school of southern China. *Infect Genet Evol*. 2020;83:104343. doi:10.1016/j.meegid.2020.104343
19. National Health and Family Planning Commission of China. Diagnosis for pulmonary tuberculosis. Health Industry Standards of China (WS288—2017); November 09, 2017.
20. General Office of the National Health and Family Planning Commission, General Office of the Ministry of Education. Regulations on prevention and control of tuberculosis in schools; July, 2017.
21. Wang Y, Long Q, Liu Q, Tolhurst R, Tang S. Treatment seeking for symptoms suggestive of TB: comparison between migrants and permanent urban residents in Chongqing, China. *Trop Med Int Health*. 2008;13(7):927–933. doi:10.1111/j.1365-3156.2008.02093.x
22. Martinez L, Xu L, Chen C, et al. Delays and pathways to final tuberculosis diagnosis in patients from a Referral Hospital in Urban China. *Am J Trop Med Hyg*. 2017;96(5):1060–1065. doi:10.4269/ajtmh.16-0358
23. Anwaerjiang A, Wang Q, Liu H, et al. Prevalence and molecular characteristics based on whole genome sequencing of Mycobacterium tuberculosis Resistant to four anti-tuberculosis drugs from Southern Xinjiang, China. *Infect Drug Resist*. 2021;14:3379–3391. doi:10.2147/IDR.S320024
24. Honore-Bouakline S, Vincensini JP, Giacuzzo V, Lagrange PH, Herrmann JL. Rapid diagnosis of extrapulmonary tuberculosis by PCR: impact of sample preparation and DNA extraction. *J Clin Microbiol*. 2003;41(6):2323–2329. doi:10.1128/JCM.41.6.2323-2329.2003
25. Stecher G, Tamura K, Kumar S. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol Biol Evol*. 2020;37(4):1237–1239. doi:10.1093/molbev/msz312
26. Kumar S, Stecher G, Li M, Nnyaz C, Tamura K:MEGAX. Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547–1549. doi:10.1093/molbev/msy096
27. 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(3):275–284. doi:10.1016/S1473-3099(16)30418-2
28. Meumann EM, Horan K, Ralph AP, et al. Tuberculosis in Australia's tropical north: a population-based genomic epidemiological study. *Lancet Reg Health West Pac*. 2021;15:100229. doi:10.1016/j.lanwpc.2021.100229
29. Chen W, Xia Y, Li X, et al. A tuberculosis outbreak among senior high school students in China in 2011. *J Int Med Res*. 2012;40(5):1830–1839. doi:10.1177/030006051204000521
30. Hou J, Pang Y, Yang X, et al. Outbreak of Mycobacterium tuberculosis Beijing Strain in a High School in Yunnan, China. *Am J Trop Med Hyg*. 2020;102(4):728–730. doi:10.4269/ajtmh.19-0533
31. Wang S, Tang Y, Zhong L, et al. Pulmonary tuberculosis outbreak in a high school, China. *Austin Intern Med*. 2021;5(1):1053.
32. Wang L, Yeerjiang Y, Gao HF, Pei JF, Zhang RX, Xu WH. Self-reported anxiety level and related factors in senior high school students in China during the outbreak of coronavirus disease 2019. *J Affect Disord*. 2022;301:260–267. doi:10.1016/j.jad.2022.01.056
33. You NN, Zhu LM, Li GL, et al. A tuberculosis school outbreak in China, 2018: reaching an often overlooked adolescent population. *Epidemiol Infect*. 2019;147:e303. doi:10.1017/S0950268819001882
34. Bryant JM, Schürch AC, van Deutekom H, et al. Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. *BMC Infect Dis*. 2013;13(1):110. doi:10.1186/1471-2334-13-110
35. Han Z, Li J, Sun G, et al. Transmission of multidrug-resistant tuberculosis in Shimen community in Shanghai, China: a molecular epidemiology study. *BMC Infect Dis*. 2021;21(1):1118. doi:10.1186/s12879-021-06725-0
36. Kikuchi T, Nakamura M, Hachisu Y, Hirai S, Yokoyama E. Molecular epidemiological analysis of Mycobacterium tuberculosis modern Beijing genotype strains isolated in Chiba Prefecture over 10 years. *J Infect Chemother*. 2022;28(4):521–525. doi:10.1016/j.jiac.2021.12.020
37. Pan D, Lin M, Lan R, et al. Tuberculosis transmission in households and classrooms of adolescent cases compared to the community in China. *Int J Environ Res Public Health*. 2018;15(12):12. doi:10.3390/ijerph15122803

## Infection and Drug Resistance

Dovepress

### Publish your work in this journal

Infection and Drug Resistance is an international, peer-reviewed open-access journal that focuses on the optimal treatment of infection (bacterial, fungal and viral) and the development and institution of preventive strategies to minimize the development and spread of resistance. The journal is specifically concerned with the epidemiology of antibiotic resistance and the mechanisms of resistance development and diffusion in both hospitals and the community. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/infection-and-drug-resistance-journal>