Clustering and recent transmission of *Mycobacterium tuberculosis* in a Chinese population

Guisheng Xu1,*  
Xuhua Mao2,*  
Jianming Wang1  
Hongqiu Pan3

1Department of Epidemiology, Key Laboratory of Infectious Diseases, School of Public Health, Nanjing Medical University, Nanjing, China;  
2Department of Clinical Laboratory, Yixing People's Hospital, Wuxi, China;  
3Department of Tuberculosis, The Third Hospital of Zhenjiang, Zhenjiang, China

*These authors contributed equally to this work

**Purpose:** The objectives of the present study were to characterize the clinical isolates prevailing in the northeast of Jiangsu and to investigate the mode of transmission. The study also aimed to explore the extent to which *Mycobacterium tuberculosis* strains contributed to drug resistance and the possible factors related to the recent transmission.

**Patients and methods:** We consecutively enrolled 912 culture-confirmed pulmonary tuberculosis (TB) cases from 1 January 2013 to 31 December 2014 in Lianyungang City, which is located in the center of China's vast ocean area and the northeast of Jiangsu province. Isolates were genotyped using 15-locus mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR) typing. The Hunter–Gaston discrimination index (HGDI) was used to estimate the discriminatory power and diversity of molecular markers.

**Results:** Among 741 successfully genotyped isolates, 144 (19.43%) strains formed 46 clusters, while 597 (80.57%) isolates had the unique MIRU pattern. The total HGDI for all 15 loci was 0.999. The average cluster size was 3 (2–13) patients. The estimated proportion of recent transmission was 13.34%. Patients with unfavorable treatment outcomes were infected with clustered strains at a higher proportion than were those with favorable treatment outcomes (adjusted OR: 1.78, 95% CI: 1.14–2.85, *P*=0.012).

**Conclusion:** The probability of recent TB transmission was relatively low in the study site, while the cases mainly arose from the activation of previous infection. Spatial analysis showed that strains forming larger clusters had the characteristics of regional aggregation.

**Keywords:** tuberculosis, molecular epidemiology, genotype, transmission, drug resistance

**Introduction**

Tuberculosis (TB) remains a global public health threat, especially in developing countries.1 Although TB control has been effective in some regions of the world, the emergence of multidrug-resistant (MDR) and extensively drug-resistant TB during the past decade threatens to undermine these advances.2,3

Molecular epidemiology is increasingly used as a complementary tool to conventional epidemiology.2 Genotyping of *Mycobacterium tuberculosis* (MtB) has been used to enhance our understanding of the TB epidemic and determine the routes of transmission. Cases with indistinguishable molecular characteristics of strains more likely contribute to recent transmission, whereas cases caused by strains with unique fingerprints are typically due to the reactivation of latent infection.5

Restriction fragment length polymorphism (RFLP) targeting the insertion sequence 6110 (IS6110) transposable element has been widely used and considered as a gold standard tool.
However, this method requires large amounts of high-quality DNA and has poor discriminatory power for isolates with >6 copies of IS6110.6 Other alternative molecular tools include mixed linker PCR (ML-PCR), double-repetitive-element PCR (DRE-PCR), fast ligation-mediated PCR (FlIP), spacer oligonucleotide typing (spoligotyping), and mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR) typing, among others.5,6

TB is highly prevalent in China, and studies on the mode of transmission have been carried out in some regions; however, few efforts in fingerprinting analysis have been made based on the mutation profiles of isolates from the east region.7–10 Jiangsu is located in the east part of China, absorbs a large number of migrations and has the highest population density in China.11 Exploring the pattern of Mtb transmission and drug resistance in this area could substantially facilitate local and countrywide TB control. The objectives of the present study were to characterize and classify the clinical isolates prevailing in the northeast of Jiangsu and to investigate the mode of transmission. The study also aimed to explore the extent to which Mtb strains contributed to drug resistance and the possible factors related to the recent spread of TB strains.

**Patients and methods**

**Study setting and participants**

This study was conducted in Lianyungang City, which is located in the center of China’s vast ocean area, the northeast of Jiangsu province and the extreme north of the Yangtze River Delta. We have consecutively enrolled 912 culture-confirmed TB cases from 1 January 2013 to 31 December 2014. The Institutional Review Board of Nanjing Medical University approved this study. Written informed consent was obtained from all participants included in the study.

**Data collection**

Sputum samples from each TB suspect were collected for acid-fast bacilli (AFB) smear microscopy, which was performed at local hospitals. All AFB smear-positive samples were routinely transported to the Fourth People’s Hospital of Lianyungang for culture and drug susceptibility testing (DST). Drug resistance to rifampicin (RIF), isoniazid (INH), ofloxacin (OFLX), and kanamycin (KA) was assessed based on the proportion method as recommended by WHO/IUATLD (International Union against Tuberculosis and Lung Disease). Genomic DNA was obtained by suspending mycobacterial colonies in 100–200 μl distilled H2O and incubating at 85°C for 30 min. All isolates were genotyped using 15-locus MIRU-VNTR typing (Mtub04, ETRC, ETRE, MIRU40, MIRU10, MIRU16, Mtub21, Qub11b, ETRA, Mtub30, MIRU26, ETRE, Mtub39, Qub26, and Qub4156). The primers used to amplify specific loci are described in Table 1. PCR was performed on an amplifier (MJ Research, Inc., Watertown, MA, USA), and the amplification products were separated by agarose gel electrophoresis. The product sizes were estimated by comparison with the 50 bp and 100 bp ladder markers. The copy number per locus was calculated using the method described by Supply et al.12 We applied the “pheatmap” package of R software to perform the clustering analysis.13

**Data analysis**

Data were entered and analyzed using SPSS 18.0 software (SPSS Inc., Chicago, IL, USA). Characteristics of cases with different strains were compared using the χ² or Fisher’s exact test. A multivariate logistic regression model was applied to estimate the effect of factors related to drug resistance, clustering, and cluster size. The test level was set at 0.05.

**Table 1** Primers and HGDI for the 15 MIRU-VNTR loci

| Loci   | Sequence (5¢→3¢)                                      | HGDI  |
|--------|-------------------------------------------------------|-------|
| Mtub04 | CTGGCCCGCCATCAGGCAGCATTATT GGCAGCAGGCGGGGATTTCTTC   | 0.511 |
| ETRC   | CGAGGATGGACATGGCCGTTATCT AATGACTTTGAAACGGATTGTA     | 0.036 |
| ETRE   | GGGCACAGCCGCAACTGC GGGCACAGGACGACTGC                 | 0.344 |
| MIRU40 | GGGTGTTGAGTGACAACTGTGT    GGGTGTACGCGGCGAAACATGAGTA | 0.509 |
| MIRU10 | GTCCTAGAACAACTGAGCGCTCG  GCCACCTTGATGACTGGTACCTACT | 0.624 |
| MIRU16 | TCGGTGACTGCGGTACGTCAAGTA CCGTCTGACGCGCCTGAGTA       | 0.231 |
| Mtub21 | AGATCCAGRTGTTGCTGTCG    CAACATTCCGCGCTTCTGTA        | 0.747 |
| Qub11b | CGTAAAGGGGGATGCGGGGAATAGG CGAAGTAAAGTGTGGCAT        | 0.762 |
| ETRA   | AAAATCGGTTCATCACCTTTCTAT CGAAGCTTGGGCGGCCGATTT      | 0.397 |
| Mtub30 | CTTGAAGCGCCCGGCTCTCATCTGT ACTTGACCGCACCAGGCTAAGTA  | 0.382 |
| MIRU26 | TAGGCTTACGTCGAAACTGTTGAC CATAGGCGACCGCGGAAATG      | 0.727 |
| ETRE   | ACTGTATTGCTCATACGCGTTTA GTCGCACTGTGCTTGTAG         | 0.597 |
| Mtub39 | CGGTTGGAGGCGATGACCTCTTC TTAGAGGCAGCGGGAAAGCTTAG     | 0.428 |
| Qub26  | CGCGCGTCGGGCCGGAGGTCTCTCCGCG GACGTCAGCGCTGTGAGAT   | 0.610 |
| Qub4156| TGGTGGCTAGGCACGCTCTGCGG  TACACCGCGGCGGCTTAC         | 0.242 |
| Total  | –                                                      | 0.999 |

**Abbreviations:** HGDI: Hunter–Gaston discrimination index; MIRU, mycobacterial interspersed repetitive unit; VNTR, variable number tandem repeat.
A cluster was defined as two or more isolates from different patients with identical MIRU patterns, whereas nonclustered patterns were referred to as unique. If the cluster contained five or more cases, it was defined as a large cluster. The index case in each large cluster was temporally defined as the first case diagnosed within the cluster. The rate of recent transmission was calculated by the formula: \[ \frac{Tc-Nc}{Ta} \times 100\% \], where \( Tc \) was the total number of clustered isolates, \( Nc \) was the number of clusters, and \( Ta \) was the total number of isolates. The “pheatmap” package of R software was used to plot the MIRU–VNTR results. The dendrograms based on 15 VNTR locus data were conducted using the unweighted pair group method with arithmetic mean protocol. According to the first-class clustering results, the samples were divided into MIRU-type I and II. The Hunter–Gaston discrimination index (HGDI) for the loci in MIRU-15 sets was used to estimate the discriminatory power of the genotyping method and the diversity of molecular markers.\(^{14,15}\)

\[
HGDI = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} n_j(n_j - 1),
\]

where \( N \) is the total number of strains, \( n_j \) is the number of strains of the j genotype, and \( s \) is the MIRU-VNTR locus genotype number.

Results
We genotyped 912 clinical samples, and 741 (81.25%) isolates provided complete data for the entire MIRU-15 set. Among the 741 TB patients, 135 (18.22%) were previously treated, 274 (36.98%) were aged over 60 years, and 587 (79.22%) were male. DST showed that the drug resistance to RIF, INH, OFLX, and KA was 8.37%, 11.74%, 9.04% and 0.68%, respectively. The proportion of MDR-TB was 5.94%, which was higher among patients aged <60 years (Table 2).

According to the genotypes of the MIRU-15 loci, strains were categorized into different groups. The total HGDI for all loci was 0.999 (Table 1). The estimated proportion of recent transmission was 13.34%. As shown in Figure 1, two main groups were categorized, and we defined them as MIRU-type I (n=214, 28.88%) and MIRU-type II (n=527, 71.12%). The ETRE, MIRU10, MTUB04, MIRU40, Mtub21, Mtub30, Mtub39, and MIRU26 loci had more repeated sequences in the MIRU-type II group than in the MIRU-type I group. The MIRU-type I and II classification was significantly associated with gender (\( \chi^2 = 6.21, P = 0.013 \)), age (\( \chi^2 = 4.67, P = 0.031 \)), and clustering (\( \chi^2 = 53.15, P < 0.001 \)) (Table 3).

One hundred forty-four (19.43%) strains formed 46 clusters, while 597 (80.57%) isolates had a unique MIRU pattern (Figure 1). The average cluster size was 3 (2–13) cases. Over 38.19% of clusters contained two cases, and the largest cluster included 13 cases. We have plotted patients in a geographic information system map. There were 9 patients in cluster 1, which was mainly distributed in the north part of Lianyungang. Cluster 2 included 5 patients, all living in the south part of the study area. There were 5 patients in cluster 3, and they were distributed throughout the whole city. Cluster 4 included 12 patients, mainly distributed in the...
east or south area (Figure 2). We have further categorized 144 clustered strains into two groups: small clusters (including 2–4 cases) and large clusters (including 5 or more cases). No significant factors were observed to be related with cluster size (Table 4).

As shown in Table 5, strains in the MIRU-type II group were more likely to form a cluster (adjusted OR: 13.25, 95% CI: 5.73–30.62, P<0.001). Patients with unfavorable treatment outcomes were infected with clustered strains at a higher proportion than were those with favorable
Infection and Drug Resistance 2018:11

Discussion

We have consecutively enrolled TB cases from one region of China and genotyped them using the 15-locus MIRU-VNTR. The proportion of clustering in the study area was 19.43%, and the recent transmission rate was 13.34%, which was lower than that reported in some developed countries. For example, a study conducted in London reported that 46% were clustered and the estimated proportion attributable to recent transmission was 34%.16 Another study performed in the UK reported that the proportion of cases attributable to recent transmission was 15%.17 In the Netherlands, 46% of strains were found in clusters with identical RFLPs, and 35% were attributed to active transmission.18 In the United States, recent transmission was estimated to be 48% from 1996 to 2000 but decreased to 23% in 2005–2009.19,20 The proportion of recent transmission was relatively low in our study, indicating the higher possibility of recrudescence in the study area. This may partly contribute to the higher prevalence of latent infection in the Chinese population.21

In 2013, a population-based study conducted in rural China reported that the age-standardized and sex-standardized rates of skin-test positivity (≥10 mm) ranged from 15% to 42%, and the QuantiFERON positivity rates ranged from 13% to 20%.22 Another cross-sectional study that was undertaken in eastern of China reported that the positive rate of the interferon-gamma release assay was 19.98%.22

According to the genotypes of MIRU-15 loci, two main groups were categorized (MIRU-type I and II). The classification was significantly associated with gender and age. This finding might be explained as different recent transmission rates among specific populations. Previous studies have reported that recent TB infection is associated with a variety of factors, including youth, male sex, belonging to an ethnic minority, homelessness, drug use, excessive alcohol consumption, homelessness, and previous incarceration.23 Factors related to clustering varied greatly between areas. In a UK study conducted by Hamblion et al, the clustered patients

---

**Table 3** Characteristics of strains genotyped in Lianyungang

| Variables          | All, N (%) | MIRU-type I, N (%) | MIRU-type II, N (%) | \( \chi^2 \) | \( P \) |
|--------------------|------------|--------------------|---------------------|-------------|--------|
| Gender             |            |                    |                     |             |        |
| Male               | 587 (79.22)| 182 (85.05)        | 405 (76.85)         | 6.21        | 0.013  |
| Female             | 154 (20.78)| 32 (14.95)         | 122 (23.15)         |             |        |
| Age (years)        |            |                    |                     |             |        |
| <60                | 467 (63.02)| 122 (57.00)        | 345 (65.50)         | 4.67        | 0.031  |
| ≥60                | 274 (36.98)| 92 (43.00)         | 182 (34.50)         |             |        |
| TB history         |            |                    |                     |             |        |
| New cases          | 606 (81.78)| 180 (84.11)        | 426 (80.83)         | 1.10        | 0.295  |
| Previously treated | 135 (18.22)| 34 (15.89)         | 101 (19.17)         |             |        |
| Location           |            |                    |                     |             |        |
| Urban              | 207 (27.79)| 50 (23.36)         | 157 (29.79)         | 3.12        | 0.077  |
| Rural              | 534 (72.21)| 164 (76.64)        | 370 (70.21)         |             |        |
| Clustering         |            |                    |                     |             |        |
| Unique             | 597 (80.57)| 208 (97.20)        | 389 (73.81)         | 53.15       | <0.001 |
| Clustered          | 144 (19.43)| 6 (2.80)           | 138 (26.19)         |             |        |
| Treatment outcomea |            |                    |                     |             |        |
| Successful         | 507 (84.78)| 158 (87.78)        | 349 (83.49)         | 1.79        | 0.181  |
| Unsuccessful       | 91 (15.22) | 22 (12.22)         | 69 (16.51)          |             |        |

Note: *With missing values.

Abbreviations: MIRU, mycobacterial interspersed repetitive unit; TB, tuberculosis.
### Table 4 Factors related to cluster size

| Variables     | Small (2–4 cases) | Large (≥ 5 cases) | OR   | 95% CI       | P    | aOR* | 95% CI* | P  |
|---------------|------------------|-------------------|------|--------------|------|------|---------|----|
|               | Size of cluster, N (%) |                  |      |              |      |      |         |    |
| Gender        |                  |                   |      |              |      |      |         |    |
| Male          | 80 (84.21)        | 42 (85.71)        | I    |              |      |      |         |    |
| Female        | 15 (15.79)        | 7 (14.29)         | 0.89 | 0.34–2.35    | 0.812| 0.91 | 0.34–2.41| 0.845|
| Age (years)   |                  |                   |      |              |      |      |         |    |
| <60           | 60 (63.16)        | 28 (57.14)        | I    |              |      |      |         |    |
| ≥60           | 35 (36.84)        | 21 (42.86)        | 1.29 | 0.64–2.60    | 0.483| 1.28 | 0.63–2.59| 0.492|
| Location      |                  |                   |      |              |      |      |         |    |
| Urban         | 30 (31.58)        | 10 (20.41)        | I    |              |      |      |         |    |
| Rural         | 65 (68.42)        | 39 (79.59)        | 0.56 | 0.25–1.26    | 0.159| 0.57 | 0.25–1.31| 0.188|
| TB history    |                  |                   |      |              |      |      |         |    |
| New cases     | 80 (84.21)        | 34 (69.39)        | I    |              |      |      |         |    |
| Previously treated | 15 (15.79)    | 15 (30.61)        | 2.35 | 1.04–5.35    | 0.041| 2.30 | 1.00–5.30| 0.050|
| DR-TB         |                  |                   |      |              |      |      |         |    |
| No            | 80 (84.21)        | 42 (85.71)        | I    |              |      |      |         |    |
| Yes           | 15 (15.79)        | 7 (14.29)         | 0.90 | 0.34–2.35    | 0.812| 0.90 | 0.34–2.39| 0.833|
| MDR-TB        |                  |                   |      |              |      |      |         |    |
| No            | 91 (95.79)        | 46 (93.88)        | I    |              |      |      |         |    |
| Yes           | 4 (4.21)          | 3 (6.12)          | 1.50 | 0.32–6.91    | 0.615| 1.53 | 0.34–7.15| 0.591|
| Treatment outcomeb |            |                   |      |              |      |      |         |    |
| Successful    | 71 (87.65)        | 31 (77.50)        | I    |              |      |      |         |    |
| Unsuccessful  | 10 (12.35)        | 9 (22.50)         | 2.06 | 0.76–5.57    | 0.154| 2.04 | 0.75–5.53| 0.160|

**Notes:** *a* adjusted for age and gender; *b* with missing values; DR-TB: drug resistance to RIF, INH, OFLX or KA; MDR-TB: drug resistance to RIF and INH.

**Abbreviations:** TB, tuberculosis; DR, drug-resistance; MDR, multidrug-resistant; MIRU, mycobacterial interspersed repetitive units; RIF, rifampicin; INH, isoniazid; OFLX, ofloxacin; KA, kanamycin.

### Table 5 Factors related to the clustering of strains

| Variables       | Clustered, N (%) | OR   | 95% CI       | P    | aOR* | 95% CI* | P  |
|-----------------|-----------------|------|--------------|------|------|---------|----|
|                 | No              | Yes  |              |      |      |         |    |
| Gender          |                 |      |              |      |      |         |    |
| Male            | 465 (77.89)     | 122 (84.72) | I    |              |      |      |         |    |
| Female          | 132 (22.11)     | 22 (15.28)  | 0.64 | 0.39–1.04    | 0.071| 0.64 | 0.39–1.05| 0.079|
| Age (years)     |                 |      |              |      |      |         |    |
| <60             | 379 (63.48)     | 88 (61.11)  | I    |              |      |      |         |    |
| ≥60             | 218 (36.52)     | 56 (38.89)  | 1.11 | 0.76–1.61    | 0.597| 0.94 | 0.65–1.37| 0.750|
| TB history      |                 |      |              |      |      |         |    |
| New cases       | 492 (82.41)     | 114 (79.17) | I    |              |      |      |         |    |
| Previously treated | 105 (17.59)    | 30 (20.83)  | 1.23 | 0.78–1.94    | 0.366| 1.19 | 0.75–1.88| 0.463|
| Location        |                 |      |              |      |      |         |    |
| Urban areas     | 167 (28.00)     | 40 (27.88)  | I    |              |      |      |         |    |
| Rural areas     | 430 (72.00)     | 104 (72.22) | 0.99 | 0.66–1.52    | 0.963| 1.03 | 0.65–1.53| 0.906|
| DR-TB           |                 |      |              |      |      |         |    |
| No              | 476 (79.73)     | 122 (84.72) | I    |              |      |      |         |    |
| Yes             | 121 (20.27)     | 22 (15.28)  | 0.71 | 0.43–1.17    | 0.18 | 0.72 | 0.44–1.18| 0.186|
| MDR-TB          |                 |      |              |      |      |         |    |
| No              | 560 (93.80)     | 137 (95.14) | I    |              |      |      |         |    |
| Yes             | 37 (6.20)       | 7 (4.86)    | 0.77 | 0.34–1.77    | 0.543| 0.80 | 0.35–1.83| 0.589|
| MIRU type       |                 |      |              |      |      |         |    |
| I               | 208 (34.84)     | 6 (4.17)    | I    |              |      |      |         |    |
| II              | 389 (65.16)     | 138 (95.83) | 12.30| 5.34–28.33   | <0.001| 13.25 | 5.73–30.62| <0.001|
| Treatment outcomeb |            |      |              |      |      |         |    |
| Successful      | 374 (77.92)     | 82 (67.77)  | I    |              |      |      |         |    |
| Unsuccessful    | 106 (22.08)     | 39 (32.23)  | 1.68 | 1.08–2.56    | 0.021| 1.78 | 1.14–2.85| 0.012|

**Notes:** *a* adjusted for age and gender; *b* with missing values; DR-TB: drug resistance to RIF, INH, OFLX or KA; MDR-TB: drug resistance to RIF and INH.

**Abbreviations:** TB, tuberculosis; DR, drug-resistance; MDR, multidrug-resistant; MIRU, mycobacterial interspersed repetitive units; RIF, rifampicin; INH, isoniazid; OFLX, ofloxacin; KA, kanamycin.
were more likely to be born in the UK, to have been diagnosed with pulmonary TB, to have a previous TB history, to have issues with substance abuse or alcohol abuse, and to have been imprisoned. Another UK study conducted by Anderson et al found that being born in the UK and using illicit drugs were significantly associated with clustering. Using the 15-locus MIRU method, a study in Jiangxi province, China, reported that patients who failed treatment or patients with MDR isolates were more likely to be in clusters. In our study, we observed that clustered cases were more likely to have an unfavorable treatment outcome. With the extension of sputum bacteriological conversion, the risk of ongoing transmission among the community also increased. Compared with that in other studies, the size of clustered cases was small in the current study, which may result in a relatively small sample size over a short time period.

Strain genotyping data, when combined with epidemiological data, enabled the identification of TB patients involved in the same chain of transmission. We assume that cases sharing the same strain (in the same cluster) were recently infected in Lianyungang. We have plotted cases forming large clusters on the map. Cases in cluster 1 or 2 mainly lived together in a small area. We hypothesized that the strains were probably transmitted through the same source of infection.

In 1997, Supply et al identified a novel minisatellite-like structure in the Mtbe genome and named such structures MIRUs. These MIRUs are located mainly in intergenic regions and are dispersed throughout the Mtbe genome. In 2001, Supply et al proved the usefulness of MIRUs for epidemiologic study by conducting PCR analysis of 12 variable tandem repeat loci with specific primers followed by gel electrophoresis. Supply et al proposed a 15-locus system with specific primers followed by gel electrophoresis. The MIRU-VNTR loci and drug resistance have been explored in previous studies. For example, the ETRB locus was associated with INH resistance, and the MIRU20 locus was associated with EMB resistance.

In 1997, Supply et al identified a novel minisatellite-like structure in the Mtbe genome and named such structures MIRUs. These MIRUs are located mainly in intergenic regions and are dispersed throughout the Mtbe genome. In 2001, Supply et al proved the usefulness of MIRUs for epidemiologic study by conducting PCR analysis of 12 variable tandem repeat loci with specific primers followed by gel electrophoresis. Supply et al proposed a 15-locus system with specific primers followed by gel electrophoresis. The MIRU-VNTR loci and drug resistance have been explored in previous studies. For example, the ETRB locus was associated with INH resistance, and the MIRU20 locus was associated with EMB resistance.

In the current study, the total HGDI for all 15 loci was 0.999. The locus Qub11b had the highest HGDI (0.762), while the HGDI of ETRC was very low (0.036). Identical MIRU-VNTR patterns are considered to be in a cluster. It was reported that the proportion of clustering decreased with a greater number of MIRU-VNTR loci typed, with increasing TB incidence and with increasing maximum cluster size.

Limitations
Our study has some limitations. First, samples were collected over a relatively short time period. The probability of clustering might be underestimated. A long-term cohort study can solve this problem. Second, China has a high proportion of the Beijing lineage of Mtbe. In this study, we did not analyze the lineage of the Beijing family. Third, epidemiological data are essential for identifying the route of transmission. In this study, we lack necessary information such as the history of contact with the patients or the frequency of visiting areas with high population density. Thus, we could not clearly explain the chain of transmission for patients in the same cluster.

Conclusion
The probability of recent TB transmission was relatively low in the study site, and the cases mainly came from the activation of previous infection. Spatial analysis showed that some strains in the larger clusters had the characteristics of regional aggregation.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (81473027), Qing Lan Project (2014), Six Talent Peaks Project in Jiangsu province (2014-YY-023), Jiangsu province Science and Technology Funds (BL2014067), Key Research and Development Plan of Zhenjiang (SHW2016008, SH2017022), and Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions
GX and JW conceived, initiated and led the study. GX, XM, JW and HP analyzed the data with input from all the authors. GX and JW prepared the manuscript. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work. All authors reviewed and approved the manuscript.

Disclosure
The authors report no conflicts of interest in this work.

References
1. Raviglione M, Sulis G. Tuberculosis 2015: Burden, challenges and strategy for control and elimination. Infect Dis Rep. 2016;8(2):6570.
2. Gandhi NR, Nunn P, Dheda K, et al. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. Lancet. 2010;375(9728):1830–1843.
3. Dheda K, Gumbo T, Maartens G, et al. The epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant, extensively drug-resistant, and incurable tuberculosis. Lancet Respir Med. 2017;5(4):291–300.

4. Bonura C, Gomgnimbou MK, Refregier G, et al. Molecular epidemiology of tuberculosis in Sicily, Italy: what has changed after a decade? BMC Infect Dis. 2014;14:602.

5. Banu S, Uddin MK, Islam MR, et al. Molecular epidemiology of tuberculosis in rural Matlab, Bangladesh. Int J Tuberc Lung Dis. 2012;16(3):319–326.

6. Li PW, Aung WW, Lee JS, Choi GE, Chang CL. Molecular strain typing of mycobacterium tuberculosis: a review of frequently used methods. J Korean Med Sci. 2013;31(11):1673–1683.

7. Chen H, He L, Huang H, et al. Mycobacterium tuberculosis lineage distribution in Xinjiang and Gansu Provinces, China. Sci Rep. 2017;17(7):1068.

8. Liu HC, Deng JP, Dong HY, et al. Molecular typing characteristic and drug susceptibility analysis of mycobacterium tuberculosis isolates from Zigong, China. Biomed Res Int. 2016;2016:6790985.

9. Hu Y, Mathema B, Zhao Q, et al. Acquisition of second-line drug resistance and extensive drug resistance during recent transmission of Mycobacterium tuberculosis in rural China. Clin Microbiol Infect. 2015;21(12):1093 e1099–e1093 e1018.

10. Luo T, Yang C, Peng Y, et al. Whole-genome sequencing to detect recent transmission of Mycobacterium tuberculosis in settings with a high burden of tuberculosis. Tuberculosis (Edinb). 2014;94(4):434–440.

11. Shao Y, Yang D, Xu W, et al. Epidemiology of anti-tuberculosis drug resistance in a Chinese population: current situation and challenges ahead. BMC Public Health. 2011;11:110.

12. Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C. Variable genome. Mol Microbiol. 2000;36(3):762–771.

13. Gingeras TR, Ghandour G, Wang E, et al. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic Mycobacterium DNA arrays. Genome Res. 1995;8(5):435–448.

14. Chen J, Tsolaki AG, Shen X, Jiang X, Mei J, Gao Q. Deletion-targeted multiplex PCR (DTM-PCR) for identification of Beijing/W genotypes of Mycobacterium tuberculosis. Tuberculosis (Edinb). 2007;87(5):446–449.

15. Imperiale BR, Moyano RD, AB DIG, et al. Genetic diversity of Mycobacterium avium complex strains isolated in Argentina by MIRU-VNTR. Epidemiol Infect. 2017;145(7):1382–1391.

16. Hamblion EL, Le Menach A, Anderson LF, et al. Recent TB transmission, clustering and predictors of large clusters in London, 2010-2012: results from first 3 years of universal MIRU-VNTR strain typing. Thorax. 2016;71(8):749–756.

17. Anderson LF, Tamne S, Brown T, et al. Transmission of multidrug-resistant tuberculosis in the UK: a cross-sectional molecular and epidemiological study of clustering and contact tracing. Lancet Infect Dis. 2014;14(5):406–415.

18. van Soolingen D, Borgdorff MW, de Haas PE, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. J Infect Dis. 1999;180(3):726–736.

19. Ellis BA, Crawford JT, Braden CR, et al. Molecular epidemiology of tuberculosis in a sentinel surveillance population. Emerg Infect Dis. 2002;8(11):1197–1209.

20. Moonan PK, Ghosh S, Oeltmann JE, Kammerer JS, Cowan LS, Navin TR. Using genotyping and geospatial scanning to estimate recent mycobacterium tuberculosis transmission, United States. Emerg Infect Dis. 2012;18(3):458–465.

21. Liu Y, Huang S, Jiang H, et al. The prevalence of latent tuberculosis infection in rural Jiangsu, China. Public Health. 2017;146:39–45.

22. Gao L, Lu W, Bai L, et al. Latent tuberculosis infection in rural China: baseline results of a population-based, multicentre, prospective cohort study. Lancet Infect Dis. 2015;15(3):310–319.

23. Nava-Aguilera E, Andersson N, Harris E, et al. Risk factors associated with recent transmission of tuberculosis: systematic review and meta-analysis. Int J Tuberc Lung Dis. 2009;13(1):17–26.

24. Chen KS, Liu T, Lin RR, Peng YP, Xiong GC. Tuberculosis transmission and risk factors in a Chinese anthomyint mining community. Int J Tuberc Lung Dis. 2016;20(1):57–62.

25. Surie D, Fane O, Finlay A, et al. Molecular typing, spatial, and field epidemiology suggesting TB transmission in community, not hospital, Gaborone, Botswana. Emerg Infect Dis. 2017;23(3):487–490.

26. Abrahams KA, Cox JA, Spivey VL, et al. Identification of novel imidazo[1,2-a]pyridine inhibitors targeting M. tuberculosis Qbr2. PLoS One. 2012;7(12):e52951.

27. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guillot C. Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. Mol Microbiol. 1999;34(2):257–267.

28. Zhang H, Li D, Zhao L, et al. Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. Nat Genet. 2013;45(10):1255–1260.

29. Supply P, Magdalena J, Himpens S, Locht C. Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. Mol Microbiol. 1997;26(5):991–1003.

30. Ravansalar H, Tadayon K, Ghazvini K. Molecular typing methods used in studies of Mycobacterium tuberculosis in Iran: a systematic review. Iran J Microbiol. 2016;8(5):338–346.

31. Pan XL, Zhang CL, Nakajima C, et al. A quantitative and efficient approach to select MIRU-VNTR loci based on accumulation of the percentage differences of strains for discriminating divergent Mycobacterium tuberculosis sublineages. Emerg Microbes Infect. 2017;6(7):e68.

32. Nabyonga L, Kaseete DP, Katahizi FA, et al. Determination of circulating Mycobacterium tuberculosis strains and transmission patterns among pulmonary TB patients in Kawempe municipality, Uganda, using MIRU-VNTR. BMC Res Notes. 2011;4:280.

33. Yuan X, Zhang T, Kawakami K, et al. Molecular characterization of multidrug- and extensively drug-resistant Mycobacterium tuberculosis strains in Jiangxi, China. J Clin Microbiol. 2012;50(7):2404–2413.

34. Liu J, Li J, Liu J, et al. Genotypic diversity of mycobacterium tuberculosis clinical isolates in the multiethnic area of the Xinjiang Uygur autonomous region in China. Biomed Res Int. 2017;2017:3179535.

35. Cheng XF, Jiang C, Zhang M, et al. Mycobacterial interspersed repetitive unit can predict drug resistance of Mycobacterium tuberculosis in China. Front Microbiol. 2016;7:378.

36. Mears J, Abubakar I, Cohen T, McHugh TD, Sonnenberg P. Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review. BMJ Open. 2015;5(1):e005636.