Introduction to Molecular Epidemiology of Tuberculosis

Qian Gao
Fudan University, Shanghai
2018.10.17, Shenzhen
Outline

• Introduction of MolEpi of tuberculosis
• Genotyping methods of *M. tuberculosis*
• Key points should be paid attention to in Molepi Study
• Several examples of Molepi study
The History of MolEpi of Tuberculosis

Genotyping methods + Epidemiological methods = Molepi

Molecular Epidemiology Assumption

- How to differentiate different strain?
  - Colony morphology and Phage typing: Impossible
  - Molecular genotyping: Possible

- Based on genotyping method
  - Identical genotype = same strain: outbreak, recent transmission
  - Unique genotype = different stains, reactivation, reinfection
The Significance of MolEpi

- Insights into the transmission of tuberculosis
  - Dogma: more then 90% of TB patient were caused by reaction
  - Molepi studies showed 30-70% patient caused by recent transmission

- Genotyping for tuberculosis control programs
  - Identification of risk factor for transmission
  - Improving investigations of contacts
  - Evaluation of tuberculosis programs (recent transmission rate)

- Genotyping for clinical management
  - Confirm the cross-contamination in lab
  - Identify the relapse or reinfection
  - Identify the acquired drug resistant of reinfection
Recent transmission or Reactivation?

Dogma: 90% of patients were censed by reactivation.

<table>
<thead>
<tr>
<th>Settings</th>
<th>Duration of Study</th>
<th>Genotyping Methods</th>
<th>Recent Transmission Rate</th>
<th>Risk Factor for Recent Transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malawi</td>
<td>1995-2003</td>
<td>IS6110-RFLP</td>
<td>72%</td>
<td>HIV(+), younger age, et al.</td>
</tr>
</tbody>
</table>

Risk Factor for Recent Transmission

- Homeless people
- Prison
- School and Kindergarten
- Nursing home
Genotyping Methods
Traditional Genotyping Methods

- IS611-RFLP (Restriction – fragment-length polymorphism)
- Spoligotyping
- MIRU-VNTR (Mycobacterial interspersed repeat units- Variable Number Tandem Repeat)

**IS6110-RFLP**

- Differentiate the Strains based on
  - IS6110 copies
  - Positions of IS6110 in the genome
- Characteristic of IS6110-RFLP genotyping
  - High discriminatory power
  - But not for the strains with no or low copy number of IS6110
  - Needs more DNA sample and complicated operation
  - Difficult to compare the results from different lab

Spoligotyping

- Differentiate the Strains based on
  - 43 of DR (direct-repeat) present or absent
- Characteristic of spoligotyping
  - Easier and faster
  - Digitalized results and easy for inter-laboratory comparison
  - low discriminatory power

Differentiate the Strains based on
- Copies of the tandem repeated

Characteristic of MIRU-VNTR
- High discriminatory power based on locus used (12, 16, 24 loci)
- Easier and faster
- Digitalized results and easy for inter-laboratory comparison
- VNTR-24 is recommended by USA CDC
Whole Genome Sequencing (WGS)

- WGS is an increasingly accessible and affordable for *M. tuberculosis* typing
  - The cost is getting cheaper
  - Differentiate the Strains based on the SNP (single nucleotide polymorphism) on entire genome, the mutation rate about 0.3-0.5 SNP per year per genome
- Characteristic of WGS
  - The highest of discriminatory power
  - Great increased the precision of genotyping and contact tracing
  - Elucidated the Mutation rate, drug resistance and phylogeny and evolution of *M. tuberculosis*
WGS’s Two Characteristics

- Traditional genotyping methods: inaccurate when tracing transmission routes
- WGS: tracing transmission routes by delineating the order of nucleotide changes
  - The reverse mutation of M. tuberculosis rarely happens
  - It is not common that different strains of *M. tuberculosis* have same mutations

Takiff H. & Feo O., Lancet Infect Dis 2015
WGS Vs Traditional Genotyping

- Traditional genotyping methods: including less than 1% of the genome
- WGS: including about 90-95% of the genome

WGS Vs Traditional Genotyping

All traditional DNA fingerprints for both isolates were isogenic, with the exception of the MIRU-VNTR locus 1955

- K-1 and K-2 are two clinical isolates, belong to Beijing K-family
- Both isolates were part of a large cluster of closely related organisms

WGS shows substantial genomic diversity

An outbreak of TB occurred over 3 years in Canada. MIRU-VNTR genotyping suggested the outbreak was clonal. WGS data revealed two genetically distinct lineages and suggesting two concomitant outbreaks.

## Estimate the Mutation Rate

<table>
<thead>
<tr>
<th>MTBC markers</th>
<th>Mutation rate estimates</th>
<th>Homoplasy index</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoligotype</td>
<td>$2.0 \times 10^{-2} - 9.0 \times 10^{-2}$ per year&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Yes/relatively high</td>
<td>Preliminary screen of genetic diversity, excluding possible laboratory contaminations</td>
</tr>
<tr>
<td>Regions of difference/targeted interrogation of phylogenetic SNPs</td>
<td>Not determined</td>
<td>No</td>
<td>(Sub-)lineage classification</td>
</tr>
<tr>
<td>IS6110 sequence</td>
<td>0.0135 changes per copy per year&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Yes/low to moderate (genotypes with low numbers of IS6110 copies)</td>
<td>(Local) molecular epidemiological investigations, differentiation between relapse/re-infection</td>
</tr>
<tr>
<td></td>
<td>0.0161 changes per copy per year&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIRU-VNTR loci</td>
<td>$7.0 \times 10^{-4} - 1.5 \times 10^{-2}$ per locus per year&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Yes/moderate</td>
<td>(Global) molecular epidemiological investigations, differentiation between relapse/re-infection, screening for potential TB transmission clusters, screening for lineage identification</td>
</tr>
<tr>
<td></td>
<td>$3.3 \times 10^{-4} - 9.8 \times 10^{-3}$ per locus per year&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$10^{-4}$ per locus per year&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$2.5 \times 10^{-3} - 2.6 \times 10^{-2}$ per locus per year&lt;sup&gt;f&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$1.2 \times 10^{-3} - 2.6 \times 10^{-3}$ per locus per year&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Estimate the Mutation Rate

<table>
<thead>
<tr>
<th>WGS/genome wide SNP analysis</th>
<th>0.24–0.34 SNPs per genome per year&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No/very low</th>
<th>Molecular epidemiological investigations, differentiation between relapse/re-infection, high resolution outbreak investigation, drug resistance prediction, robust phylogenetic analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.26–0.66 SNPs per genome per year&lt;sup&gt;f&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.3–0.5 SNPs per genome per year&lt;sup&gt;k&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.3–0.7 SNPs per genome per year&lt;sup&gt;i&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.93–1.56 SNPs per genome per year&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.13–0.27 SNPs per genome per year&lt;sup&gt;m&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0073–0.013 SNPs per genome per year&lt;sup&gt;h&lt;/sup&gt; (long-term rate)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References are as follows: <sup>a</sup>Reyes and Tanaka (2010), <sup>b</sup>Tanaka and Rosenberg (2001), <sup>c</sup>Rosenberg et al. (2003), <sup>d</sup>Ragheb et al. (2013), <sup>e</sup>Wirth et al. (2008), <sup>f</sup>Aandahl et al. (2012), <sup>g</sup>Eldholm et al. (2016), <sup>h</sup>Eldholm et al. (2015), <sup>i</sup>Roetzer et al. (2013), <sup>j</sup>Ford et al. (2011), <sup>k</sup>Ford et al. (2013), <sup>l</sup>Walker et al. (2013a), <sup>m</sup>Bos et al. (2014), and <sup>n</sup>Comas et al. (2013)
Key points should be paid attention
Two Caveats to MolEpi

- Require the population based study
  - To get accurate clustered rate requires the evaluation of a large percentage of TB cases in the population and over a long period.

- Require the epidemiologic information
  - Careful to interpreter the genotyping data
  - Same genotyping may not reflect recent transmission
  - Is WGS data better?

Cluster Rate = \( \frac{n}{N} \) or \( \frac{n-1}{N} \)

- \( n \) = No. of clustered isolates
- \( N \) = No. of total of isolates
- \( l \) = No. of cluster
Estimates of Recent Transmission Rate

Simulation model estimates of the influence of sampling proportion, “n” method, 100 samples

Estimates of Recent Transmission Rate

- The longer time leads to increased cluster rate
How to Define the Identical Genotype?

- Depends on genotyping methods, WGS might be the “gold standard”
  - An artificial concept and not absolute, 5 SNP Or more SNP?
  - SNPs accumulation was not linearly correlated with time in short time interval
- The change of molecular markers significantly affect the threshold
  - IS6110-RFLP half life 3.2 years; VNTR mutation rate: $10^{-2.06}$ per locus per year
  - SNP mutation rate: 0.3-0.5 SNP per year per genome
How to Define the Identical Genotype

- VNTR genotype define: 1 locus different is same or not?
  - Isolates from the same patient: might be the same
  - Isolates from different patient: might be different

<table>
<thead>
<tr>
<th>SNP distance</th>
<th>Within patient</th>
<th>United Kingdom*</th>
<th>Songjiang, Shanghai</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNP distance &lt;= 12</td>
<td>20(100%)</td>
<td>46(11.3%)</td>
<td>13(15.1%)</td>
</tr>
<tr>
<td>SNP distance &gt; 12</td>
<td>0(0%)</td>
<td>360(88.7%)</td>
<td>73(84.9%)</td>
</tr>
</tbody>
</table>

Unpublished data
Why Need to Develop an Optimal VNTR Set for Local?

- The population structure of *Mycobacterium tuberculosis* varies in different regions.
- Beijing strains are genetically highly similar, which leads to limited discriminatory power of VNTR-15/24.
- Reducing the number of loci tested is good for application.

<table>
<thead>
<tr>
<th>Method</th>
<th>Clusters</th>
<th>HGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>VNTR-15</td>
<td>6</td>
<td>0.99</td>
</tr>
<tr>
<td>VNTR-24</td>
<td>6</td>
<td>0.992</td>
</tr>
<tr>
<td>IS6110-RFLP</td>
<td>3</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Sebastian et al. Proc Natl Acad Sci USA. 2006
A Dong Shen et al. J Clin Microbiol. 2008
How to Develop an Optimal VNTR Set

Population based sample collection reflect the true HGI value
- Hospital based or random selected isolates will missed clustered isolates, which result in overestimate of discriminatory power

$$HGI = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} n_j(n_j - 1)$$
Several Examples of MolEpi
Development of VNTR Set in China

- VNTR Genotyping
  - MIRU12
  - VNTR15
  - VNTR24, VNTR 24+4

- VNTR in China
  - not standardized, many different methods were used, MIRU 12, VNTR15, VNTR24 et al.

- Objective: to develop a VNTR typing method that can achieve high resolution with a small number of loci
Population-based Collections of the Isolates

<table>
<thead>
<tr>
<th>Study fields</th>
<th>Total</th>
<th>Beijing genotype</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sichun</td>
<td>216</td>
<td>115</td>
<td>53</td>
</tr>
<tr>
<td>Guangxi</td>
<td>176</td>
<td>109</td>
<td>62</td>
</tr>
<tr>
<td>Shanghai</td>
<td>396</td>
<td>314</td>
<td>79</td>
</tr>
<tr>
<td>Shandong</td>
<td>206</td>
<td>160</td>
<td>78</td>
</tr>
<tr>
<td>Henan</td>
<td>197</td>
<td>177</td>
<td>90</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>184</td>
<td>159</td>
<td>86</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1375</strong></td>
<td><strong>1034</strong></td>
<td><strong>75</strong></td>
</tr>
</tbody>
</table>
Discriminatory Powers of 25 VNTR Loci
## Optimal VNTR Combinations

<table>
<thead>
<tr>
<th>No. of loci (no. of possible combinations)</th>
<th>No. of combinations with HGI higher than VNTR-15</th>
<th>Optimal combinations with highest the HGI*</th>
<th>HGI (mean ± STDEV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All strains</td>
<td>Non-Beijing strains</td>
</tr>
<tr>
<td>2 (190)</td>
<td>0</td>
<td>1-5</td>
<td>0.900±0.041</td>
</tr>
<tr>
<td>3 (1140)</td>
<td>0</td>
<td>1-2-5</td>
<td>0.948±0.027</td>
</tr>
<tr>
<td>4 (4845)</td>
<td>0</td>
<td>1-2-4-5</td>
<td>0.966±0.025</td>
</tr>
<tr>
<td>5 (15504)</td>
<td>0</td>
<td>1-2-3-4-5</td>
<td>0.974±0.022</td>
</tr>
<tr>
<td>6 (38760)</td>
<td>0</td>
<td>1-2-3-4-5-8</td>
<td>0.980±0.020</td>
</tr>
<tr>
<td>7 (77520)</td>
<td>0</td>
<td>1-2-3-4-5-6-8</td>
<td>0.984±0.017</td>
</tr>
<tr>
<td>8 (125970)</td>
<td>0</td>
<td>1-2-3-4-5-6-8-12</td>
<td>0.987±0.013</td>
</tr>
<tr>
<td>9 (167960)</td>
<td>8</td>
<td>1-2-3-4-5-6-7-8-10-12</td>
<td>0.989±0.011</td>
</tr>
<tr>
<td>10 (184756)</td>
<td>219</td>
<td>1-2-3-4-5-6-7-8-10-12</td>
<td>0.991±0.008</td>
</tr>
<tr>
<td>11 (167960)</td>
<td>1506</td>
<td>1-2-3-4-5-6-7-8-10-12-17</td>
<td>0.992±0.008</td>
</tr>
<tr>
<td>12 (125970)</td>
<td>4864</td>
<td>1-2-3-4-5-6-7-8-10-12-14-17</td>
<td>0.993±0.006</td>
</tr>
<tr>
<td>13 (77520)</td>
<td>8836</td>
<td>1-2-3-4-5-6-7-8-10-12-14-15-17</td>
<td>0.994±0.006</td>
</tr>
<tr>
<td>14 (38760)</td>
<td>9513</td>
<td>1-2-3-4-5-6-7-8-9-10-12-14-15-17</td>
<td>0.994±0.006</td>
</tr>
<tr>
<td>15 (15504)</td>
<td>6599</td>
<td>1-2-3-4-5-6-7-8-9-10-12-14-15-17-18</td>
<td>0.994±0.006</td>
</tr>
<tr>
<td>16 (4845)</td>
<td>3081</td>
<td>1-2-3-4-5-6-7-8-9-10-12-14-15-16-17-18</td>
<td>0.994±0.006</td>
</tr>
<tr>
<td>17 (1140)</td>
<td>941</td>
<td>1-2-3-4-5-6-7-8-9-10-12-14-15-16-17-18-20</td>
<td>0.995±0.006</td>
</tr>
<tr>
<td>18 (190)</td>
<td>181</td>
<td>1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-20</td>
<td>0.995±0.006</td>
</tr>
<tr>
<td>19 (20)</td>
<td>20</td>
<td>1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-20</td>
<td>0.995±0.006</td>
</tr>
<tr>
<td>20 (1)</td>
<td>1</td>
<td>1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20</td>
<td>0.995±0.006</td>
</tr>
</tbody>
</table>
VNTR (9+3) Genotyping for China

- Optimized 9-locus (VNTR-9) plus 3 hypervariable loci (HV-3) as standard for nationwide genotyping of MTB in China
  - VNTR-9 can be used as the first-line method for large-scale genotyping
  - HV-3 can be used to subtype the VNTR-9 clustered strains to identify the transmission in local

### VNTR (9+3) Vs. VNTR (24+4)

<table>
<thead>
<tr>
<th></th>
<th>黑龙江</th>
<th>广西</th>
<th>上海</th>
<th>四川</th>
<th>河南</th>
<th>总数</th>
</tr>
</thead>
<tbody>
<tr>
<td>菌株总数</td>
<td>163</td>
<td>137</td>
<td>202</td>
<td>188</td>
<td>161</td>
<td>851</td>
</tr>
<tr>
<td>成簇菌株数*</td>
<td>18</td>
<td>10</td>
<td>47</td>
<td>8</td>
<td>21</td>
<td>104</td>
</tr>
<tr>
<td>单基因型菌株数*</td>
<td>128</td>
<td>125</td>
<td>151</td>
<td>176</td>
<td>138</td>
<td>718</td>
</tr>
<tr>
<td>一致菌株数</td>
<td>146</td>
<td>135</td>
<td>198</td>
<td>184</td>
<td>159</td>
<td>822</td>
</tr>
<tr>
<td>一致率(%)</td>
<td>89.6</td>
<td>98.5</td>
<td>98.0</td>
<td>97.8</td>
<td>98.8</td>
<td>96.6</td>
</tr>
</tbody>
</table>

Recurrent Tuberculosis
-reinfection or relapse ?

- 5-20% cases are expected to be recurrent even cured by DOTS
- Dogma:

![Diagram showing the relationship between infection, disease, cured, relapse, and reinfection]

- Recurrence
- Relapse (reactivation)
- Infection
- Disease
- Cured
- Reinfection
- Relapse
Recurrent Tuberculosis in Shanghai: -reinfection or relapse?

- Retrospective, population-based analysis of recurrent tuberculosis from 2000 to 2012 in Shanghai city, China
- HIV Prevalence in Shanghai is low
- Compared the DNA genotypes between isolates of initial episode with those of subsequent episode.
- 42% patients with paired isolates had unmatched genotype patterns (re-infection)

Shen X., et al., Tuberculosis, 2017
Transmitted or Acquired DR among Treated Patients?

- **Dogma:** Treated patients have acquired drug resistance
- **Real acquired resistance:** Resistance mutations in bacterial genome result in acquired resistance
- **Resistant patients with TB history may come from:**
  - Real acquired resistance
  - Exogenous reinfection
  - Mixed infection
Are Resistant Patients with TB History Really Acquired Resistance?

- TB cases during 2009~2015 in Shanghai
  - Sample duration of paired sputum from individual case ≥ 90 ds: 390 cases (780 isolates)

  Molecular DST
  - decreased resistance N= 39 (10%)
  - increased resistance N= 81 (20.8%)
  - identical resistance N= 270 (69.2%)

- VNTR(9+3) genotyping, WGS
  - Identical genotypes – acquired resistance
  - Different genotypes – reinfection/mixed infections

- Median interval time of isolates collected: 342 (90~2200) days
  - 1~4 years: 173 cases
  - >4 years: 19 cases

- Fluorescence
  - Wide Type (H37Rv)
  - rpoB531 (TCG→TTG), rpoB526 (CAC→TAC), rpoB5826 (CAC→CTC), rpoB533 (CTG→CCG)
60% Treated Resistant Patients were Transmitted Resistance

- Increasing resistance among treated mostly (~60%) caused by transmission
- 84% (27/32) resistance was transmitted resistance

81 cases with increasing resistance

VNTR (9+3)

Different genotypes
N = 48 (59.3%)

Identical genotypes
N = 33 (40.7%)

Acquired resistance

To be published

50% Treated Patients were Reinfection

- Among patients whose resistance didn’t change during treatment, 50% were reinfected with another strains, indicating serious transmission.
Recent Transmission of TB

- Recent transmission: develop disease shortly (1-2ys) after infection
- Reactivation: develop disease far from infection

\[\text{Recent transmission: development shortly (1-2ys) after infection} \]
\[\text{Reactivation: development far from infection} \]

\[\text{LTBI - 95\%} \]
\[\text{Active TB - 5\%} \]

\[\text{90\%} \]
How to Differentiate Recent Transmission?

- Molecular Epidemiology assumption
  - Identical genotype (Cluster strains) - recent transmission
  - Unique genotype – reactivation
  - Genotyping: IS6110-RFLP, VNTR, Whole Genome Sequence
Study Design

- Population-based prospective study, small scale, full coverage
- 5 sites covering 4 million population
- Represent different location, economic and TB epidemic in China

<table>
<thead>
<tr>
<th>sites</th>
<th>area (km²)</th>
<th>population(000)</th>
<th>prevalence (/100 000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wuchang, HLJ</td>
<td>3,756</td>
<td>520</td>
<td>512</td>
</tr>
<tr>
<td>Weishi, HN</td>
<td>1,307</td>
<td>868</td>
<td>497</td>
</tr>
<tr>
<td>Songjiang, SH</td>
<td>604</td>
<td>1 634</td>
<td>96</td>
</tr>
<tr>
<td>Wusheng, SC</td>
<td>966</td>
<td>838</td>
<td>544</td>
</tr>
<tr>
<td>Pingguo, GX</td>
<td>2,473</td>
<td>457</td>
<td>477</td>
</tr>
</tbody>
</table>
Study Strategy

- Establish the epidemiological fields
- Screen all suspected TB patients
- Culture positive TB patients
- Identify MTB
- Genotyping
- DST
- Clustered
- Unique
- Epidemiology investigation
- Elucidate the recent transmission

Questionnaire

Screening includes

- Culture positive TB patients
- Identify MTB
- Genotyping
- DST
- Epidemiology investigation
- Elucidate the recent transmission
Established of Field Sites

- County-level lab
  - Sputum microscopy
  - Sputum culture
- Provincial lab
  - Epidemiologic investigation
  - DST
  - Internet-based database
  - VNTR genotyping
- Fudan University
  - Data analysis
  - Staff management
  - Quality control
Sample Collection

- From June 2009 to June 2012, 17,905 suspects people were screened for tuberculosis,
- 2274 (12.7%) culture-confirmed patients were diagnosed, most (71.3%) of them were male, with median age of 41 yrs (range 15-93)

<table>
<thead>
<tr>
<th>Fields</th>
<th>No. of Cases</th>
<th>Male (%)</th>
<th>Median Age, yrs</th>
<th>DR(%)</th>
<th>INH(%)</th>
<th>RIF(%)</th>
<th>MDR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guangxi</td>
<td>324</td>
<td>78.1</td>
<td>44</td>
<td>14.3%</td>
<td>11.5%</td>
<td>8.1%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Sichuan</td>
<td>414</td>
<td>77.2</td>
<td>44</td>
<td>17.0%</td>
<td>14.6%</td>
<td>11.7%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Henan</td>
<td>481</td>
<td>76.3</td>
<td>52</td>
<td>11.3%</td>
<td>10.0%</td>
<td>7.3%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Shanghai</td>
<td>797</td>
<td>64.0</td>
<td>32</td>
<td>12.1%</td>
<td>11.5%</td>
<td>6.0%</td>
<td>5.1%</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>258</td>
<td>67.8</td>
<td>48</td>
<td>14.0%</td>
<td>10.6%</td>
<td>7.9%</td>
<td>4.4%</td>
</tr>
<tr>
<td>Total</td>
<td>2274</td>
<td>71.3</td>
<td>41</td>
<td>13.3%</td>
<td>11.6%</td>
<td>7.8%</td>
<td>6.0%</td>
</tr>
</tbody>
</table>
1/3 TB was Caused by Recent Transmission

- During June 2009 to June 2012, 2238 culture (+) patients were enrolled, most (71.3%) were male, median age 41 ys (15~93)

- Cluster rate = 31%, indicating 31% cases were resulted from recent transmission

<table>
<thead>
<tr>
<th>Sites</th>
<th>total strains</th>
<th>clustered strains</th>
<th>clusters</th>
<th>cluster rate (%)</th>
<th>cluster size</th>
<th>max cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wusheng, SC</td>
<td>414</td>
<td>90</td>
<td>42</td>
<td>21.7</td>
<td>2.1</td>
<td>4</td>
</tr>
<tr>
<td>Pingguo, GX</td>
<td>324</td>
<td>117</td>
<td>47</td>
<td>36.1</td>
<td>2.6</td>
<td>6</td>
</tr>
<tr>
<td>Weishi, HN</td>
<td>481</td>
<td>149</td>
<td>57</td>
<td>30.9</td>
<td>2.6</td>
<td>7</td>
</tr>
<tr>
<td>Songjiang, SH</td>
<td>797</td>
<td>255</td>
<td>107</td>
<td>32.0</td>
<td>2.5</td>
<td>7</td>
</tr>
<tr>
<td>Wuchang, HLJ</td>
<td>258</td>
<td>94</td>
<td>34</td>
<td>36.0</td>
<td>3.0</td>
<td>13</td>
</tr>
<tr>
<td>total</td>
<td>2274</td>
<td>705</td>
<td>287</td>
<td>31.0</td>
<td>2.5</td>
<td>13</td>
</tr>
</tbody>
</table>
Most (78.7%) of the clusters were comprised of two patients.
Cluster rate of MDR-TB is much higher than DS-TB (43.7% vs 31.0%, p=0.005)

MDR strains transmit easier than susceptible strains (aOR=1.86, 95%CI 1.25-2.63)

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Adjusted OR</th>
<th>95%CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDR-TB</td>
<td>1.86</td>
<td>1.25-2.63</td>
<td>0.001</td>
</tr>
<tr>
<td>Beijing Strains</td>
<td>1.56</td>
<td>1.23-2.96</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

WGS to Analysis the Recent Transmission of MDR-TB

- 2009-2012, all culture (+) TB patients from 31 designated hospitals in Shanghai
- DST: L-J proportion method (RIF & INH)
- Genotyping
  - VNTR (9+3) : differentiate recent transmission except for resistant strains
  - WGS of clustered isolates explains recent transmission in detail
Primary Outcomes

- During 2009-2012, 7982 isolates collected
- 367 (4.6%, 95%CI 4.1-5.1) were MDR-TB
- 60% were new cases
- 73% male, median age 39 ys (16-88 ys)
WGS Analysis

- 125 (38.6%) were clustered by VNTR9+3
- WGS of 122 VNTR-clustered isolates, 32% (103/324) were confirmed recent transmission with a cutpoint of 12 SNPs
- 38 clusters with 2-8 cases
- 69% (64/93) clustered cases had epi-links
- 43% (44/103) retreated resistant patients resulted from transmission
## Risk Factors of Recent Transmission

- Diagnosis delay (>2 months), elderly
- No related to gender, TB history, smear(+)
- Public entertainment or consumer places like card rooms, community markets were hotspots for transmission

<table>
<thead>
<tr>
<th>Factors</th>
<th>aOR*</th>
<th>95%CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>diagnosis delay (≥2 ms)</td>
<td>2.29</td>
<td>1.19-4.07</td>
<td>0.005</td>
</tr>
<tr>
<td>45-64 ys</td>
<td>2.15</td>
<td>1.18-3.90</td>
<td>0.009</td>
</tr>
<tr>
<td>≥65 ys</td>
<td>3.18</td>
<td>1.36-7.41</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Tracking the Transmission of MDR
Accumulation of New Mutations during Transmission

- 36% (37/103) clustered isolates obtained non-fixed mutations, being selected in vivo
- 87% (33/38) clusters accumulated new resistance-conferring mutations in transmission: 42% developed to pre-XDR, 11% to XDR-TB

Yang C., et al., Lancet Infect Dis 2017
Transmission is the major Reason for Epidemic MDR-TB

New cases (60%) - Clinical history

Retreated cases (40%)

Laboratory genotyping:
- recent transmission
- remote transmission
- exogenous reinfection
- mixed infections
- acquired resistance

Transmitted > 80%
Acquired < 20%
“More then 80% of incident MDR tuberculosis cases in most present-day epidemic settings result from transmission of MDR tuberculosis rather than selection of de-nove resistance during previous treatment of the index case”.

Kendall E., *et.al.*, Lancet Respir Med. 2015
Summary

- Molepi has revolutionized our understanding of the transmission of tuberculosis
- WGS has greatly increased the precision of genotyping and contact tracing
- Prospective, population-based Molepi still limited, especially in the TB high burden countries.
- Hope more Molepi research to discover the new pattern of TB transmission and promote the TB control program in the TB high burden countries.
Thank you!