Molecular Typing of *Treponema pallidum*: a 5-Year Surveillance in Shanghai, China

Ting Dai,a** Kang Li,b Haikong Lu,b Xin Gu,b Qianqiu Wang,a and Pingyu Zhoua

Institute of Dermatology, Chinese Academy of Medical Sciences and Peking Union Medical College, Nanjing, China; and Department of STD Institute, Shanghai Skin Disease Hospital, Shanghai, China

Previously, a small study showed that 14f was the predominant subtype of *Treponema pallidum* in Shanghai, China. The result was quite different from the genotype distribution in other areas of China. This study aimed to identify the strain types of *Treponema pallidum* in samples collected over a 5-year period in Shanghai. From 2007 to 2011, genital swabs were collected from patients with syphilis from the Shanghai Skin Disease Hospital. Positive specimens were typed by the enhanced typing method by adding a *tp0548* gene to the existing *arp* and *tpr* genotype system. In total, 304 of the 372 enrolled patients yielded fully typeable DNA. Ten *arp* types (4, 6, 8, 9, 11, 12, 13, 14, 15, and 19), 3 *tpr* types (a, d, and o), and 5 *tp0548* types (a, c, f, g, and i) were identified. In total, 12 subtypes were identified with a combination of the *arp* and *tpr* genes. Subtype 14d was found in 270 samples (88.8%). When the combination included the *tp0548* gene, the 12 CDC subtypes identified were divided into 14 strain types. The predominant type was 14d/f (88.8%), followed by 15d/f (3.6%), 13d/f (1.3%), and 19d/c (1.3%). Two of the 44 14d/f-infected patients and both of the 19d/c-infected patients who underwent a lumbar puncture were diagnosed with neurosyphilis. This study showed that the predominant type in Shanghai was 14d/f. While this is in keeping with data from other areas in China, it is different from an earlier report showing that 14f is the most common genotype in Shanghai. Further studies are needed to better understand the association between strain types and neurosyphilis.

China has witnessed a resurgence of syphilis in recent years. The number of reported syphilis cases increased more than 50-fold within 12 years, from 1993 to 2005 (2). The national incidence rates in 2011 and 2010 were 32.04 and 28.86 cases per 100,000 population, respectively, with 429,677 and 358,534 cases reported, causing 92 and 69 deaths, respectively (from China national surveillance data) (3). Shanghai is the largest city and economic center of China, where the migrating population (nonresident) constitutes over one-third of the total population. In 2010, Shanghai had an incidence rate of 76.42 cases per 100,000 population for all stages of syphilis, with 429,677 and 358,534 cases reported, causing 92 and 69 deaths, respectively (from China national surveillance data) (3). Shanghai is the largest city and economic center of China, where the migrating population (nonresident) constitutes over one-third of the total population. In 2010, Shanghai had an incidence rate of 76.42 cases per 100,000 population for all stages of syphilis, with the total reported cases in Shanghai being comparable to those for the entire European Union.

The ability to differentiate between genotypes of *Treponema pallidum* could allow monitoring of changes in the prevalence and geographical distribution of strains over time. In 1998, Pillay et al. published a molecular subtyping scheme for *T. pallidum* based on characterization of the acidic repeat protein gene (*arp*) and restriction fragment length pattern (RFLP) analysis of *tpr* gene subfamily II (*tprE, -G, and -J*) (CDC subtype) (14). In 2010, Marra et al. developed an enhanced strain typing method, adding the *tp0548* gene sequence types to the existing CDC genotype system (9).

A previous small study showed that 14f was the predominant subtype in Shanghai (10). The result was quite different from the limited data on genotype distribution in other areas of China (13, 20, 21, 22), where 14d is the predominant type. This study aimed to identify strain types of *T. pallidum* in a larger number of samples collected over a 5-year period in Shanghai. Associations between strain types and neurosyphilis were also explored.

**Materials and Methods**

**Study population.** Between August 2007 and December 2011, eligible patients attending the Shanghai Skin Disease Hospital Sexually Transmitted Disease (STD) Clinic were asked to participate in this study. An initial diagnostic evaluation was performed, which included dark-field microscopy or *Treponema pallidum* particle agglutination (TPPA) and the rapid plasma reagin (RPR) tests. Patients who had moist lesions (chancres, condyloma lata, or mucous patches) that were confirmed as early syphilis were eligible for enrollment.

Information collected included the patient’s demographic data, clinical characteristics, and laboratory findings. Syphilis was staged according to clinical characteristics, dark-field microscopy, and serotesting. The diagnosis of neurosyphilis was confirmed if the patient had positive serum RPR and TPPA test results and (in the absence of substantial contamination of cerebrospinal fluid [CSF] with blood) positive CSF TPPA and reactive CSF venereal disease research laboratory (VDRL) test results. In addition, neurosyphilis can be highly suspected if the patient has a nonreactive CSF VDRL test result but has positive serum RPR, serum TPPA, and CSF TPPA test results along with either or both of the following: (i) elevated CSF proteins or elevated CSF white blood cell (WBC) count in the absence of other known causes of the abnormalities or (ii) clinical neurological or psychiatric manifestations without other known causes of these clinical abnormalities (19). HIV testing was also performed for patients who agreed. This study was approved by the Ethics Committee of the Shanghai Skin Disease Hospital. Informed consent was obtained from all participants.
DNA extraction and molecular typing of *T. pallidum*. Specimens were obtained by rubbing the base of the lesion with a sterile swab. DNA was extracted using a QIAamp DNA minikit (Qiagen) according to the manufacturer's instructions in a biosafety cabinet in a separated room to minimize potential contamination with exogenous DNA. DNA was extracted using a QIAamp DNA minikit (Qiagen) according to the manufacturer's instructions in a biosafety cabinet in a separated room to minimize potential contamination with exogenous DNA. To extract DNA from the lesions, DNA was obtained by rubbing the base of the lesion with a sterile swab. DNA extraction and molecular typing of *T. pallidum*. Specimens were obtained by rubbing the base of the lesion with a sterile swab. DNA was extracted using a QIAamp DNA minikit (Qiagen) according to the manufacturer's instructions in a biosafety cabinet in a separated room to minimize potential contamination with exogenous DNA. To extract DNA from the lesions, DNA was obtained by rubbing the base of the lesion with a sterile swab. DNA extraction and molecular typing of *T. pallidum*. Specimens were obtained by rubbing the base of the lesion with a sterile swab. DNA was extracted using a QIAamp DNA minikit (Qiagen) according to the manufacturer's instructions in a biosafety cabinet in a separated room to minimize potential contamination with exogenous DNA. To extract DNA from the lesions, DNA was obtained by rubbing the base of the lesion with a sterile swab. DNA extraction and molecular typing of *T. pallidum*. Specimens were obtained by rubbing the base of the lesion with a sterile swab. DNA was extracted using a QIAamp DNA minikit (Qiagen) according to the manufacturer's instructions in a biosafety cabinet in a separated room to minimize potential contamination with exogenous DNA. To extract DNA from the lesions, DNA was obtained by rubbing the base of the lesion with a sterile swab. DNA extraction and molecular typing of *T. pallidum*. Specimens were obtained by rubbing the base of the lesion with a sterile swab. DNA was extracted using a QIAamp DNA minikit (Qiagen) according to the manufacturer's instructions in a biosafety cabinet in a separated room to minimize potential contamination with exogenous DNA. To extract DNA from the lesions, DNA was obtained by rubbing the base of the lesion with a sterile swab.

### Results

In total, 372 patients were enrolled during the 5-year study period. By *tpp47* gene screening, 316 (86%) of the 372 specimens were positive, indicating the presence of *T. pallidum* DNA; among them, 304 had fully typeable *T. pallidum* DNA. Twelve of the 372 specimens were positive but could not be fully typed; among these 12, two were found to be negative for both *tpr* and *arp* gene PCR analysis, while eight were typeable by the *tpr* gene only.

The median age of these typeable patients was 39.6 years. A total of 65.5% of participants were men, and 69.0% resided in or near the investigating city. Of the 304 patients, 62.5% (190) had primary syphilis, 26.0% (79) had secondary syphilis, and 11.5% (35) had chancres together with lesions suggestive of secondary syphilis (skin rashes). A total of 117 patients provided sexual orientation information, and 5.1% of them (6/117) were homosexual/bisexual. Of the 304 patients, 247 accepted the offer of an HIV test, and of these, 3.6% (9/247) were HIV positive. An RPR test was performed in all participants, but data for 8 patients are not available, since the test was performed in other hospitals. Of the remaining 296 participants, 33.1% (98/296) had RPR titers equal to or less than 1:8 (the RPR test was negative for 38 of the 98 patients), 33.4% (99/296) had titers ranging from 1:16 to 1:64, and 33.4% (99/296) had RPR titers greater than 1:64.

A total of 10 *arp* types (4, 6, 8, 9, 11, 12, 13, 14, 15, and 19) were identified in these specimens, with type 14 being the most common type detected (90.1%, 274/304) (Fig. 1A). Three *tpr* RFLP patterns (a, d, and o) were found, the majority of which belonged to the type d (98.0%, 297/304) (Fig. 1B). By combining the *arp* and

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>tpp47</em> sense</td>
<td>5′-CGTGTGGTATCAACTAGTG-3′</td>
</tr>
<tr>
<td><em>tpp47</em> antisense</td>
<td>5′-CTAACCCTTACTCTTGACGC-3′</td>
</tr>
<tr>
<td>ARP-1</td>
<td>5′-CAAGTCAGGAGCGGACTGTCG-3′</td>
</tr>
<tr>
<td>ARP-2</td>
<td>5′-GCTATCACTGGGGATGTCG-3′</td>
</tr>
<tr>
<td>A-1</td>
<td>5′-ACTGGCCTCTGCCACACTTGA-3′</td>
</tr>
<tr>
<td>B-2</td>
<td>5′-CTACAGGAGAGGGTGAAGCT-3′</td>
</tr>
<tr>
<td>IP-6</td>
<td>5′-CAGGTTTGGCCGTAAAGCC-3′</td>
</tr>
<tr>
<td>IP-7</td>
<td>5′-AATCAGGGGAGATACCGTCG-3′</td>
</tr>
<tr>
<td><em>tp0548</em> sense</td>
<td>5′-GTCCTCTATGATATCCTGTCA-3′</td>
</tr>
<tr>
<td><em>tp0548</em> antisense</td>
<td>5′-GTCATGGACTCCTGGAGTG-3′</td>
</tr>
</tbody>
</table>

**TABLE 1. Primers used in *T. pallidum* strain typing**

**RESULTS**

In total, 372 patients were enrolled during the 5-year study period. By *tpp47* gene screening, 316 (86%) of the 372 specimens were positive, indicating the presence of *T. pallidum* DNA; among them, 304 had fully typeable *T. pallidum* DNA. Twelve of the 372 specimens were positive but could not be fully typed; among these 12, two were found to be negative for both *tpr* and *arp* gene PCR analysis, while eight were typeable by the *tpr* gene only.

The median age of these typeable patients was 39.6 years. A total of 65.5% of participants were men, and 69.0% resided in or near the investigating city. Of the 304 patients, 62.5% (190) had primary syphilis, 26.0% (79) had secondary syphilis, and 11.5% (35) had chancres together with lesions suggestive of secondary syphilis (skin rashes). A total of 117 patients provided sexual orientation information, and 5.1% of them (6/117) were homosexual/bisexual. Of the 304 patients, 247 accepted the offer of an HIV test, and of these, 3.6% (9/247) were HIV positive. An RPR test was performed in all participants, but data for 8 patients are not available, since the test was performed in other hospitals. Of the remaining 296 participants, 33.1% (98/296) had RPR titers equal to or less than 1:8 (the RPR test was negative for 38 of the 98 patients), 33.4% (99/296) had titers ranging from 1:16 to 1:64, and 33.4% (99/296) had RPR titers greater than 1:64.

A total of 10 *arp* types (4, 6, 8, 9, 11, 12, 13, 14, 15, and 19) were identified in these specimens, with type 14 being the most common type detected (90.1%, 274/304) (Fig. 1A). Three *tpr* RFLP patterns (a, d, and o) were found, the majority of which belonged to the type d (98.0%, 297/304) (Fig. 1B). By combining the *arp* and
tpr gene types, we observed 12 CDC subtypes, with subtype 14d being most frequently detected (88.8%, 270/304). For tp0548, five sequence types (a, c, f, g, and i) were identified (Fig. 2), with type f accounting for 97.0% (295/304) of the samples. By enhancing the CDC subtyping method with sequence analysis of the tp0548 gene, we were able to separate the 12 CDC subtypes into 14 different strain types (Fig. 3). The majority (88.8%, 295/304) of the fully typeable specimens belonged to strain type 14d/f. The second most common strain type was 15d/f, present in 11 (11/304, 3.3%) specimens, followed by 13d/f and 19d/c, both found in four (4/304, 1.3%) specimens. The other subtypes included 14a/f (3/304, 1.0%), 4d/f, 15d/f, and 9o/c (2 [7%] of each); 8d/g, 11d/f, 11o/a, 12d/f, 14a/i, and 19d/f (one of each). Over the 5-year study period, 14d/f remained the predominant circulating type in Shanghai (11/11 samples in 2007, 35/40 in 2008, 127/144 in 2009, 46/51 in 2010, and 51/58 in 2011), with some other strain types appearing and disappearing. (Fig. 3).

Of the 304 patients who had typeable samples, 52 patients agreed to undergo lumbar puncture, and six subtypes were found: 14d/f (44 samples); 4d/f, 15d/f, and 19d/c (two of each); and 11o/c and 14a/f (one of each). Among these 52 patients, four were diagnosed with neurosyphilis, of which two had a 14d/f strain and two had a 19d/c strain. All four neurosyphilis patients were HIV negative. No neurosyphilis was found in patients with the other four subtypes. The association between 14d/f and neurosyphilis is not statistically significant (P > 0.05).

DISCUSSION

China has experienced a resurgence in syphilis in the last 20 years (2). Epidemiological studies of many infectious diseases, including STDs such as HIV infection, human papilloma virus infection, gonorrhea, and chlamydia, have been enhanced by molecular typing of their respective causative agents (6, 8, 18). Molecular biology methods are becoming increasingly accessible to many diagnostic laboratories and have the potential to enhance existing STD surveillance and control efforts in China.

To our knowledge, this is the largest T. pallidum typing study done to date, with more than 300 positive samples over the 5-year study period. Fourteen types were identified from these samples using enhanced nomenclature derived by combining the number of the arp gene tandem repeats, the tpr gene RFLP types and the tp0548 gene sequence type. A previous small study indicated the CDC subtype 14f is the predominant genotype in Shanghai (10). However, in several small studies from Guangdong and Hunan provinces in China, as well as a recent large national study, CDC subtype 14d is the most common subtype (13, 15, 20–22). In our study, 5-year data also showed that 14d is the predominant subtype, and we did not identify subtype 14f from a single sample. This is in keeping with a recent large study across different geographic areas in China which did not identify 14f in any samples. Furthermore, Shanghai is the biggest city and economic center of eastern China. Because of its transportation network, Shanghai has close relationships with other areas of the country. In Shanghai, the migrating population constitutes over one-third of the total population, and therefore, it seems reasonable that subtype 14d is predominant both in Shanghai and adjacent areas. We used positive and negative controls in the study. Additionally, the test was repeated and results confirmed at the Institute Pasteur of Shanghai. Thus, we believe 14d is the predominant subtype in Shanghai. Outside China, CDC subtype 14d is also the most common molecular type in Scotland, Canada, South Africa, and Colombia, while 14f was found to be predominant in studies performed in Arizona, North Carolina, and South Carolina in the United States (1, 4, 5, 7, 11, 16, 17).

A total of 10 arp types (4, 6, 8, 9, 11, 12, 13, 14, 15, and 19), three tpr RFLP patterns (a, d, and o), and five tp0548 sequence types (a, c, f, g, and i) were identified in this study. Of note, arp types 4, 11, and 19, tpr type o, and tp0548 sequence types a, g, and i had never been previously reported in China. It seems that the diversity of T. pallidum subtypes might be higher with greater syphilis prevalence, as is seen in South Africa, where there appears to be a high level of strain diversity (15). It is interesting that the proportion of the predominant type 14d is much higher in Shanghai (88.8%) than in other areas, where the proportion of the predominant type varied from 30% to 76%. Even with enhanced typing method, most of the tp0548 types were f (295/304, 97.0%) and all the subtype 14d samples had the tp0548 type f. Over the 5-year period, different strain types have appeared and disappeared, with 14d/f being the predominant circulating type (Fig. 3). The much smaller...
numbers of the non-14d/f subtypes may indicate smaller and more localized transmission networks in a relatively small geographic area, or the relatively short period of the syphilis resurgence in Shanghai. Another explanation is that type 14d/f might be more virulent than other types or might generate a weaker host immune response, resulting in a longer duration of infection. There could also be differences in sensitivity of the strains to antibiotics (23, 24).

Although current syphilis treatment guidelines do not recommend a lumbar puncture in asymptomatic, HIV-negative patients with early syphilis, it is useful to obtain CSF to examine associations between disease progression and strain types. The recent study by Marra et al. indicated that strain type 14d/f might be associated with a higher risk of neurosyphilis (9). Some T. pallidum types may be more neuroinvasive or better able to escape host immune responses in the central nervous system (CNS) than others. However, the association between 14d/f and neurosyphilis is not significant in Shanghai. In our study, 14d/f accounted for nearly 90% of the cases in Shanghai. Two of the 44 (4.3%) patients infected with strain type 14d/f were initially diagnosed as having secondary syphilis and confirmed as having neurosyphilis after CSF examination. Another study from Africa showed that 14a was predominant in neurosyphilis patients while 14d was most prevalent in epidemiologic studies (12), which indicated that type 14a might be associated with a higher risk of neurosyphilis. Again, all four of the 14a strains in our study (three 14a/f strains and one 14a/i strain) were not isolated from neurosyphilis cases. Type 19d/c was not reported in the study by Marra et al. (9). In our study, a total of four patients were infected with type 19d/c. Two of these, initially diagnosed as having secondary syphilis, were confirmed as having neurosyphilis after CSF examination. Both of them were HIV negative and without other immunosuppressive disease.

Because of the current syphilis treatment guidelines and the difficulty of isolating T. pallidum in neurosyphilis patients, very limited information about the strain types was obtained. Thus, it might be too early to draw a conclusion about the higher-risk strain types associated with neurosyphilis. Further studies on a larger sample of patients with neurosyphilis may help to demonstrate the preferential strain type of neurosyphilis.

ACKNOWLEDGMENTS

We appreciate Sheila A. Lukehart (University of Washington School of Medicine) and Haibo Wang (Institute Pasteur of Shanghai) for their technical support. We thank Christina Marra (University of Washington School of Medicine) for providing Nichols strain. We also thank John Saunders (St Bartholomew’s Hospital, London, United Kingdom) for help with reviewing the paper.

This study was supported by The National Natural Science Foundation of China (NSFC: 30972664), Basic Research Project of Shanghai Science and Technology Commission (STC: 11JC1411600), and Shanghai Natural Science Foundation (STC: 09ZR1428300).

We have no conflicts of interest to declare.

REFERENCES