Short Communication

The Prevalence and Distribution of *Chlamydia trachomatis* Genotypes among Sexually Transmitted Disease Clinic Patients in Guangzhou, China, 2005–2008

Bin Yang¹,²*, He-Ping Zheng¹, Zhan-Qin Feng¹, Yao-Hua Xue¹, Xing-Zhong Wu¹, Jin-Mei Huang¹, Xiu-Juan Xue¹, and Han-Ning Jiang¹

¹Guangdong Provincial Center for Skin Diseases and STI Control and Prevention, Guangzhou; and ²Xiangya School of Medicine, Central South University, Changsha, China


SUMMARY: This study was designed to determine the prevalence and distribution of *Chlamydia trachomatis* genotypes from clinical specimens in Guangzhou, China, obtained in the period 2005–2008. One hundred and ninety-four urogenital *C. trachomatis* samples were collected from sexually transmitted disease clinic patients, and the VS1-VS2 of *OmpA* gene was amplified by nested PCR and sequenced using an ABI-prism 3730 sequencer. Clinical *C. trachomatis* strains were genotyped and analyzed for a mutation with respect to the reference VS1-VS2 sequence. VS1-VS2 fragments with 453 bp were amplified from 194 clinical samples. Upon alignment with the sequences of the reference strains, 189 strains with discernible sequences were typed into 9 genotypes, while 5 with ambiguous sequences were considered to be mixed-serovar samples. The most prevalent genotypes were E (50, 26%), F (46, 24%), J (35, 19%), and D (24, 13%). There was no significant difference in the distribution of any of the genotypes detected during the study period, except for genotype K (*P* < 0.01). A total of 16 (8%, 16/189) genetic variants of the *OmpA* VS1-VS2 of the reference strains were identified. Mutations occurred frequently for genotypes D (2/24, 8%), E (6/50, 12%), F (2/46, 4%), G (1/8, 13%), H (1/12, 8%), and K (4/11, 36%), with most of these being sense mutations that may result in amino acid substitution. Sequencing the *OmpA* VS1-VS2 enabled the genotype and sequence variations within each genotype to be analyzed. Genotypes E, F, J, and D continued to dominate among urogenital *C. trachomatis*, whereas genotype K increased significantly in Guangzhou between 2005 and 2008.

Urogenital *Chlamydia trachomatis* infection, the most prevalent sexually transmitted disease (STD), has become a major public health problem worldwide (1). Thus, although most such infections remain asymptomatic (2), undetected and untreated infections result in persistent transmission and can cause serious sequelae, including pelvic inflammatory disease, ectopic pregnancy, tubal infertility, and epididymitis. Moreover, *C. trachomatis* infection may be a cofactor in the transmission of human immunodeficiency virus (HIV) (3).

The typing of *C. trachomatis* infection is important for epidemiological studies. Fifteen human serovars have been detected to date based on immunoepitope analyses of the major outer membrane protein (MOMP) with monoclonal antibodies (4). However, serological typing is limited in that a culture from a clinical isolate is required, which means that newly emerging types may be missed (5,6). Different molecular methods for genotyping *C. trachomatis* that target the *OmpA* gene have been reported recently (7–9). Indeed, automated sequencing of the *OmpA* gene that encodes MOMP allows a more precise typing of *C. trachomatis* (10,11).

---

*Corresponding author: Mailing address: Guangdong Provincial Center for Skin Diseases and STI Control and Prevention, 10 Xian Lie Dong Heng Rd, Guangzhou, 510500, China. Tel & Fax: +86-2087631390, E-mail: bin_yanggz@yahoo.com.cn*
CT2 (5′-GCR TTR CAR AGA ACR TTY AAY TC-3′). The inner primers were CT3 (5′-ACT TGG TTT TCG ACC GTG TTT TG-3′) and CT4 (5′-GAT TGA GCG TAT TGG AAA GAA GC-3′) (Invitrogen Bio, Shanghai, China). The amplification products of the OmpA VS1-VS2 were directly purified and sequenced by Invitrogen Bio using an ABI prism 3730 sequencer (Applied Biosystems Division, Perkin Elmer, Foster City, Calif., USA) according to the manufacturer’s instructions. Clinical samples were sequenced once, and samples differing from the relevant reference OmpA sequence were reanalyzed using a second PCR preparation to rule out the possibility of Taq errors. The sequences of clinical strains were determined by comparison with the reference strains. Genotypes and mutations of the C. trachomatis strains were determined by comparison with the reference strains, whereas 16 (8%) differed from the respective reference sequence by up to three nucleotides found to be E (50, 26%), F (46, 24%), J (35, 19%), and D (24, 13%) (Table 1). A comparison of the genotype distribution between 2005 and 2008 showed no significant difference for any genotypes detected, except for type K (P < 0.01). The five ambiguous sequences represented a double peak in the sequence chromatogram, which might indicate heterogeneity within a PCR product and could be considered to be a result of mixed infections with different genotypes.

A detailed analysis of the VS1-VS2 nucleotide sequences of the clinical strains was carried out to determine the mutations in VS1-VS2 of C. trachomatis. The results of this analysis indicated only limited sequence differences between the reference sequence and within genotypes (Table 2). Thus, 173 of the 189 clinical specimens which produced discernible sequences showed identical nucleotide sequences to the VS1-VS2 of the reference strains and the most prevalent genotypes found to be E (50, 26%), F (46, 24%), J (35, 19%), and D (24, 13%) (Table 1). A comparison of the genotype distribution between 2005 and 2008 showed no significant difference for any genotypes detected, except for type K (P < 0.01). The five ambiguous sequences represented a double peak in the sequence chromatogram, which might indicate heterogeneity within a PCR product and could be considered to be a result of mixed infections with different genotypes.

A detailed analysis of the VS1-VS2 nucleotide sequences of the clinical strains was carried out to determine the mutations in VS1-VS2 of C. trachomatis. The results of this analysis indicated only limited sequence differences between the reference sequence and within genotypes (Table 2). Thus, 173 of the 189 clinical specimens which produced discernible sequences showed identical nucleotide sequences to the VS1-VS2 of the reference strains and the most prevalent genotypes found to be E (50, 26%), F (46, 24%), J (35, 19%), and D (24, 13%) (Table 1). A comparison of the genotype distribution between 2005 and 2008 showed no significant difference for any genotypes detected, except for type K (P < 0.01). The five ambiguous sequences represented a double peak in the sequence chromatogram, which might indicate heterogeneity within a PCR product and could be considered to be a result of mixed infections with different genotypes.

A detailed analysis of the VS1-VS2 nucleotide sequences of the clinical strains was carried out to determine the mutations in VS1-VS2 of C. trachomatis. The results of this analysis indicated only limited sequence differences between the reference sequence and within genotypes (Table 2). Thus, 173 of the 189 clinical specimens which produced discernible sequences showed identical nucleotide sequences to the VS1-VS2 of the reference strains and the most prevalent genotypes found to be E (50, 26%), F (46, 24%), J (35, 19%), and D (24, 13%) (Table 1). A comparison of the genotype distribution between 2005 and 2008 showed no significant difference for any genotypes detected, except for type K (P < 0.01). The five ambiguous sequences represented a double peak in the sequence chromatogram, which might indicate heterogeneity within a PCR product and could be considered to be a result of mixed infections with different genotypes.

Table 1. Genotype distribution of 189 urogenital C. trachomatis strains in Guangzhou

<table>
<thead>
<tr>
<th>Year</th>
<th>B</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>2 (2)</td>
<td>12 (13)</td>
<td>22 (23)</td>
<td>29 (30)</td>
<td>2 (2)</td>
<td>8 (8)</td>
<td>1 (1)</td>
<td>19 (20)</td>
<td>1 (1)</td>
<td>96</td>
</tr>
<tr>
<td>2008</td>
<td>0</td>
<td>12 (13)</td>
<td>28 (30)</td>
<td>17 (18)</td>
<td>6 (6)</td>
<td>4 (4)</td>
<td>0</td>
<td>16 (17)</td>
<td>10 (11)</td>
<td>93</td>
</tr>
<tr>
<td>Total</td>
<td>2 (1)</td>
<td>24 (13)</td>
<td>50 (26)</td>
<td>46 (24)</td>
<td>8 (4)</td>
<td>12 (6)</td>
<td>1 (1)</td>
<td>35 (19)</td>
<td>11 (6)</td>
<td>189</td>
</tr>
</tbody>
</table>

$\chi^2$ 0.47 0.01 1.26 3.65 1.28 1.29 0.97 0.21 8.13

$P$ >0.05 >0.05 >0.05 >0.05 >0.05 >0.05 >0.05 <0.01

Table 2. Mutations found in 189 clinical strains compared to the reference strains

<table>
<thead>
<tr>
<th>Genotype</th>
<th>No. of strains with mutation (%)</th>
<th>No. of base mutation</th>
<th>Nucleotide mutation</th>
<th>Mutation position</th>
<th>Variable sequence</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>1/24 (4)</td>
<td>1</td>
<td>A → G</td>
<td>289</td>
<td>VS1</td>
<td>Silent</td>
</tr>
<tr>
<td>D</td>
<td>1/24 (4)</td>
<td>2</td>
<td>C → T</td>
<td>264</td>
<td>VS1</td>
<td>P → L</td>
</tr>
<tr>
<td>E</td>
<td>2/50 (4)</td>
<td>1</td>
<td>C → T</td>
<td>264</td>
<td>VS1</td>
<td>P → L</td>
</tr>
<tr>
<td>E</td>
<td>2/50 (4)</td>
<td>1</td>
<td>A inserted</td>
<td>515</td>
<td>VS2</td>
<td>S → C</td>
</tr>
<tr>
<td>E</td>
<td>2/50 (4)</td>
<td>3</td>
<td>A → C</td>
<td>257</td>
<td>VS1</td>
<td>Silent</td>
</tr>
<tr>
<td>F</td>
<td>1/46 (2)</td>
<td>1</td>
<td>T inserted</td>
<td>265</td>
<td>VS1</td>
<td>A → S</td>
</tr>
<tr>
<td>F</td>
<td>1/46 (2)</td>
<td>2</td>
<td>G → C</td>
<td>266</td>
<td>VS1</td>
<td>A → P</td>
</tr>
<tr>
<td>G</td>
<td>1/8 (13)</td>
<td>1</td>
<td>G → A</td>
<td>484</td>
<td>VS2</td>
<td>G → S</td>
</tr>
<tr>
<td>H</td>
<td>1/12 (8)</td>
<td>1</td>
<td>T inserted</td>
<td>266</td>
<td>VS1</td>
<td>T → Y</td>
</tr>
<tr>
<td>K</td>
<td>4/11 (36)</td>
<td>1</td>
<td>A → G</td>
<td>293</td>
<td>VS1</td>
<td>N → S</td>
</tr>
</tbody>
</table>
the function and antigenicity of the MOMP.

The MOMP of *C. trachomatis* encoded by the *OmpA* gene is highly immunogenic and is thought to play a role in protective immunity (14,15). Nucleotide substitution almost invariably results in a different amino acid sequence and these changes therefore account for the antigenic differences between *C. trachomatis* serovars. Thus, *OmpA* is a suitable target for serotyping *C. trachomatis*. Yuan et al. have analyzed the sequences of four variable sequence regions (VS1–VS4) of the *OmpA* gene for all 15 *C. trachomatis* serovars and have demonstrated that the nucleotide sequences for VS1 and VS2 are sufficiently different among all serovars (16). *OmpA* is a single copy gene. The previous technique for genotyping *C. trachomatis* targeted the whole *OmpA* gene but had low sensitivity (49%) and a rather large (1.1 kb) gene fragment was required (17). In this study, two pairs of PCR primers were designed for *OmpA* constant regions flanking variable segments VS1-VS2, and the method nested PCR amplification used showed a high sensitivity, similar to that for chlamydial plasmid-PCR (13), and proved able to genotype *C. trachomatis* from urogenital samples. Sequence analysis of the *OmpA* VS1-VS2 of *C. trachomatis* from clinical specimens showed that 189 (97%) of the 194 *C. trachomatis* samples showed a discernible sequence and could be successfully analyzed into nine genotypes, whereas the remaining 5 (3%) were ambiguous sequences that were considered to be mixed infections with different genotypes. This technology provides an efficient and sensitive means for molecular epidemiological analysis of *C. trachomatis* infections among high-risk groups and sexual partners.

The identification and genotyping of *C. trachomatis* from clinical samples are vital for molecular epidemiology and vaccination studies. A difference in the distribution of *C. trachomatis* serovars was noted between Thailand and Japan, with serovars F (35%) and E (18) being the most prevalent in Thailand and serovars D (32%), F (17%), and E (17%) in Japan, in the mid-1990s (18). Furthermore, a longitudinal study in Japan showed that the most frequently identified serovars in 1999–2001 (E, 28%; D, 19%; G, 14%; and F, 12%) were different to those in 2003–2005 (D, 25%; E, 24%; G, 19%; and I, 12%) (19). Serovars B, Ba, and I showed significant differences during the three observation periods. However, there was no significant difference in the geographic distribution of *C. trachomatis* serovars among high-risk women in China, with serotype E being predominant in the South (32%) and East (27%), and serotype F in the Southwest (28%) (20). Another study also showed that a similar distribution existed in four cities in southern China (21). In our study, *C. trachomatis* from the clinical samples in Guangzhou more commonly included genotypes E (50, 26%), F (46, 24%), J (35, 19%), and D (24, 13%); the other genotypes (B, G, H, I, and K) were encountered infrequently. A comparison of the genotype distribution between 2005 and 2008 showed no significant difference in any of the genotypes detected, except genotype K (*P* < 0.01). The most frequently encountered genotypes (E and F) in our clinical specimens were different from those reported previously from studies in the Netherlands (E, 23%; F, 31%) (7), Sweden (E, 47%; F, 17%) (22), and Korea (E, 45%; F, 20%) (23). Genotype E (22%) was the most prevalent in Taiwan (24), followed by D (19%), F (16%), J (15%), K (11%), G (11%), H (6%), and Ba (2%), with a geographical difference in the prevalence of genotype H (*P* < 0.018) being found between northern and southern Taiwan. The detection of two B strains in this study is in accordance with a previous report of the presence of Ba strains in the genital tract (25).

Previous studies have reported the genetic variability of *OmpA* gene sequences (26,27). In Guangzhou, we found 173 (92%) identical sequences and 16 (8%) genetic variants of *OmpA* VS1-VS2 for the reference strains. The base mutation or insertion was located in the VS1 or VS2 region, although no variants were identified for genotypes B, I, and J. Furthermore, genotype J, which was one of the most constant genotypes in Guangzhou, was found to be quite different from the Taiwanese genotype J strains found to have 100% mutation (22). This finding suggests that mutations occur frequently among genotypes D (2/24, 8%), E (6/50, 12%), F (2/46, 4%), G (1/8, 13%), H (1/12, 8%), and K (4/11, 36%), with most of them being sense mutations that might result in amino acid substitution (Table 2). These changes may also account for the immunogenic differences between the MOMPs of *C. trachomatis* genotypes.

In conclusion, nine genotypes were found to be commonly associated with urogenital infections in our samples. Genotypes E, F, J, and D continued to dominate urogenital *C. trachomatis* infections, although genotype K increased significantly in Guangzhou between 2005 and 2008. Furthermore, sequencing the VS1-VS2 of the *OmpA* gene enables sequence variations within each genotype to be analyzed and the different *C. trachomatis* strains to be differentiated. Genotyping of the VS1-VS2 of *OmpA* therefore enables a practical approach to epidemiologic investigations in *C. trachomatis* infection and transmission.

**Acknowledgments** The authors thank Dr. Joseph D. Tucker for proofreading this manuscript.

This work was supported by the Guangdong Province Science Funds (no. 2005B3420111) and the Medical Science Funds of Guangdong Provincial Bureau of Public Health (no. A2005150).

**REFERENCES**