Genotyping of *Chlamydia trachomatis* from the endocervical specimens of high-risk women in Hungary

Fruzsina Petrovay,1 Eszter Balla,1 István Németh2 and Éva Gönczől3

1Department of Bacteriology, National Centre for Epidemiology, Budapest, Hungary
2Developmental Drug Metabolism and Pharmacokinetics, Gedeon Richter Plc, Budapest, Hungary
3Department of Virology, National Centre for Epidemiology, Budapest, Hungary

The distribution of different *Chlamydia trachomatis* serovars in Hungary has not been reported previously. The objective of this study was to determine the distribution and prevalence of *C. trachomatis* serovars in a high-risk population by genotyping. The endocervical specimens of 484 female sex workers (FSWs) were screened for *C. trachomatis* by plasmid PCR. Genotyping was performed in all *C. trachomatis*-positive samples by PCR-based RFLP analysis of the *omp1* gene. A total of 32 specimens (6.6 %) were positive for *C. trachomatis*. Age was an important risk factor for *C. trachomatis* infection in FSWs. The highest prevalence was detected in women under the age of 20 (18.8 %). All positive specimens were successfully genotyped and seven serovars were identified. The most prevalent was serovar D (34.4 %), followed by E (21.9 %), F (18.8 %), G (9.4 %), J (9.4 %), H (3.1 %) and I (3.1 %). A heterogeneous distribution of *C. trachomatis* serovars was observed in the study group, where the most common serovars were D, E and F comprising 75 % of the positive samples. This PCR-based RFLP method could be used in epidemiological studies on the prevalence of *C. trachomatis* infection to provide more information and to compare the serovar distribution among different cohorts.

INTRODUCTION

*Chlamydia trachomatis* infection is one of the most prevalent sexually transmitted infections (STIs) worldwide. The estimated global incidence of new cases of *C. trachomatis* infection is about 92 million, in western Europe it is 5 million, and the asymptomatic occurrence is 50 and 70 % in men and women, respectively (WHO, 2001).

The immunodominant surface antigen of *C. trachomatis* is the major outer-membrane protein (MOMP) (Caldwell et al., 1981). The *omp1* gene encoding the MOMP consists of four variable domains (VD1–4), which are flanked and interspaced by five constant domains. The nucleotide sequence polymorphisms in the variable domains account for the different antigenic properties of *C. trachomatis* serovars (Yuan et al., 1989). *C. trachomatis* strains have been classified into at least 15 different serovars based on immuno-epitope analysis of the MOMP with polyclonal and monoclonal antibodies (Caldwell et al., 1981; Wang et al., 1985; Yuan et al., 1989). Additionally, on the basis of amino acid sequence homology, the 15 serovars of *C. trachomatis* have been classified into 3 groups: group B (B, D, E, L1 and L2), group C (A, C, H, I, J, K, L1 and L3) and an intermediate group (F and G) (Yuan et al., 1989).

Serovars A–C are associated with trachoma, which is hyperendemic in Africa, Asia, and Central and South America, and the second major cause of blindness in the world (Thylefors et al., 1995). Serovars D–K are commonly associated with urogenital infections and perinatal conjunctivitis or pneumonia worldwide (Numazaki, 2004; Paavonen & Eggert-Kruse, 1999). In women, a broad spectrum of clinical manifestations can be present due to urogenital *C. trachomatis* infections including urethritis and cervicitis, and, if left untreated, the ascending infection may cause secondary complications in the upper genital tract, including pelvic inflammatory disease and infertility (Paavonen & Eggert-Kruse, 1999; Paavonen & Lehtinen, 1996). Serovars L1–L3 are associated with lymphogranuloma venereum, a systemic disease most prevalent in tropical and subtropical areas. Nevertheless, several outbreaks among homosexual male populations in western countries have been reported recently, indicating the emerging role of lymphogranuloma venereum (Van de Laar, 2006).

The genotyping methods used recently in many studies for the identification of *C. trachomatis* serovars have mainly involved PCR-based RFLP and sequencing of the *omp1* gene.
gene (Dean et al., 1995; Gao et al., 2007; Lan et al., 1993, 1995; Lima et al., 2007; Lysén et al., 2004; Morré et al., 2000a; Ngandjio et al., 2004; Sturm-Ramirez et al., 2000; Van Duynhoven et al., 1998). Serovar determination of C. trachomatis by different typing methods is a challenging area of chlamydial research. The geographical distribution and prevalence of different serovars give insights into the molecular epidemiology of C. trachomatis and typing may reveal the circulation of C. trachomatis serovars and variants in a certain community (Lysén et al., 2004; Sturm-Ramirez et al., 2000). The possible relationship between urogenital serovars and clinical manifestations has also been studied, but the results are controversial. Some studies have found associations between clinical symptoms and certain serovars (Dean et al., 1995; Morré et al., 2000a; Van Duynhoven et al., 1998), whilst others have concluded that urogenital infections are not strongly influenced by serovars (Geisler et al., 2003; Lysén et al., 2004). Newer molecular approaches, in particular amplified fragment length polymorphism analysis, a fingerprinting technique (Morré et al., 2000b), and different multi-locus sequence typing methods, have also been developed to study the genomic differences in C. trachomatis serovars and to gain more information on their association with clinical disease (Klint et al., 2007; Pannekoek et al., 2008).

In Hungary, limited information is available on C. trachomatis prevalence in certain populations and there are no data about the circulating serovar distribution (Deák et al., 1997; Ujházy et al., 2007). The aims of the present study were to determine and characterize the prevalence of C. trachomatis serovars in a high-risk population by the genotyping method of omp1 PCR-based RFLP analysis.

METHODS

C. trachomatis strains. Extracted DNA samples from C. trachomatis strains of different serovars were kindly provided by Dr. H. Jalal (Addenbrooke’s Hospital, Cambridge, UK) for use as reference controls as described in a published study (Jalal et al., 2006).

Clinical specimens. Endocervical samples were obtained from 484 street-based female sex workers (FSWs) attending a mobile clinic during an anonymous screening in Budapest and the suburban area from January 2006 to July 2006. The specimens were collected with Dacron-tipped swabs and transported to the laboratory. The swabs were placed into sterile Eppendorf tubes containing 400 μl transport medium (Bartels FlexTrans; Trinity Biotech). After incubation at 4 °C for 24 h, the swabs were removed following vortexing and the samples were stored at −20 °C until further analysis.

DNA extraction. The endocervical samples were centrifuged at 10 000 g for 10 min, the pellets were resuspended in PBS and the DNA was extracted using a High Pure DNA template preparation kit (Roche Diagnostics) according to the manufacturer’s instructions. The elution step with 75 μl elution buffer was repeated once and the extracted DNA was stored at −20 °C.

PCR detection of C. trachomatis. The samples were first screened by a PCR specific for the human β-globin gene to act as a control for the efficacy of DNA extraction and the presence of PCR inhibitors using primers PC04 (GenBank accession no. A26623) and GH20 (GenBank accession no. A26624). For the detection of C. trachomatis DNA in the samples, a diagnostic PCR was used to amplify a 495 bp fragment of the cryptic plasmid gene of C. trachomatis with slightly modified primers of CTP1 and CTP2 described by Ossewaarde et al. (1992) (Table 1). The PCR was performed in an Eppendorf thermocycler, in a volume of 20 μl containing 10 x ReddyMix buffer (ABgene), 200 μM dNTPs, 0.5 μM each primer and 0.5 U ThermoPrime Taq DNA polymerase (ABgene). The PCR conditions consisted of a denaturation step at 95 °C for 4 min, followed by 40 cycles of amplification at 95 °C for 30 s, 64 °C for 30 s and 72 °C for 45 s, with an elongation step at 72 °C for 4 min. The PCR products were visualized on a 1 % agarose gel stained with ethidium bromide.

Genotyping. C. trachomatis-positive samples were selected for omp1 PCR-based RFLP genotyping. The primers used for the omp1 PCR have been described by Lan et al. (1993, 1995) (Table 1). Briefly, a semi-nested PCR was performed with the outer primer set consisting of NLO and NRO, with 3 μl extracted DNA sample added to the reaction mixture in a final volume of 20 μl. Following amplification, 5 μl of the primary PCR product was used for the nested PCR in a final volume of 50 μl. The primers used for the second PCR run were NLO and SRO2A, generating a fragment of the omp1 gene of approximately 1070 bp. Both amplification conditions consisted of a denaturation step at 95 °C for 4 min, 45 cycles of 95 °C for 30 s, 45 °C for 30 s and 72 °C for 60 s, and a final elongation step at 72 °C for 6 min. The omp1 nested PCR products were visualized by 1 % agarose gel electrophoresis.

For RFLP analysis, 10 μl of the PCR products was digested with Alul (Promega). The restriction patterns of the products were analysed by 1.5 % agarose gel electrophoresis. Serovars belonging to group C were further digested with HindI, EcoRI and DdeI (Promega). The variants of serovar D were differentiated by restriction with CfoI (Promega).

In silico RFLP analysis. The sequences of known reference C. trachomatis strains, determined by searching GenBank, were analysed by using NEBcutter version 2.0 software (http://tools.neb.com/NEBcutter2/index.php). The following strain sequences were used: D/IC-Cal8 (GenBank accession no. X62920), E/Bour (X52557), F/IC-Cal3 (X52080), G/UW57 (AF063199), H/UW4 (X16007), I/UW12 (AF063200), J/UW36 (AF202457) and K/UW31 (AF063204). These reference sequences, appropriate for the omp1 PCR product, were pasted into NEBcutter version 2.0 software and the in silico-obtained length of the restriction bands were compared with the gel-based genotyping results.

Table 1. Primer sequences used for the C. trachomatis plasmid PCR and the omp1 PCR

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’→3’)</th>
<th>Position (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTP1*</td>
<td>GGCTATTTCCTGACACGACG</td>
<td>329–350</td>
</tr>
<tr>
<td>CTP2*</td>
<td>TACCTCCGATTTCCTCCACAG</td>
<td>804–824</td>
</tr>
<tr>
<td>NLO†</td>
<td>ATGAAAAACTCTTGGAATCG</td>
<td>1–21</td>
</tr>
<tr>
<td>NRO†</td>
<td>CTCACCTAACTCGGTATTTT</td>
<td>1108–1128</td>
</tr>
<tr>
<td>SRO2A†</td>
<td>TTTCATAGTTCATCTGTGT</td>
<td>1045–1064</td>
</tr>
</tbody>
</table>

*According to the sequence of the C. trachomatis cryptic plasmid gene (GenBank accession no. X06707).
†According to the omp1 sequence of C. trachomatis serovar E (GenBank accession no. X52557).
In our study, 75% of the positive specimens comprised the serovars D, E and F in agreement with other findings, indicating that most genital infections with Chlamydia trachomatis are caused by a small number of serovars. The present study had some limitations that should be acknowledged. The occurrence of false-positive results during the plasmid PCR should have been overcome, as all positive samples were confirmed with the omp1 PCR.

Among the positive samples, seven different serovars of C. trachomatis were identified by RFLP genotyping, showing the same restriction patterns as the control strains. Serovars D, E, F and G were differentiated by AluI digestion. The group C serovars with similar AluI restriction patterns were further distinguished by EcoRI digestion, which differentiates serovars H, I and J from serovar L3. Restriction digestion with Ddel and HinfI differentiated serovars H, I and J. The CfoI digestion of samples identified as serovar D showed a distinct restriction pattern in one sample, which referred to the D− variant identified by Lan et al. (1993).

In our study population, a heterogeneous distribution of C. trachomatis serovars was observed. All serovars belonging to the D–K urogenital group could be identified except for serovar K, which was not detected. Serovar D was the most prevalent (11/32, 34.4%), followed by E (7/32, 21.9%), F (6/32, 18.8%), G (3/32, 9.4%), J (3/32, 9.4%), H (1/32, 3.1%) and I (1/32, 3.1%). The in silico RFLP analyses of the GenBank reference sequences of C. trachomatis strains correlated well with the gel-based restriction patterns of the control strains as well as those of the samples. The actual trend in serovar distribution was similar to previous findings. According to these results, the three most frequently observed serovars varied among D, E, and the intermediate group serovars (F and G), according to different typing methods worldwide (Gao et al., 2007; Lan et al., 1993, 1995; Lima et al., 2007; Lima et al., 2003; Ward et al., 2004), this rate usually appears to be less than 10% in Europe (Belgium 7.3%, Czech Republic 5.5%, Germany 9.6%, Greece 5.9%, Spain 5.9%, UK 4.8%) as detected by nucleic acid amplification tests, and these findings correspond with our results.

Data on age were obtained from 275 individuals. The FSWs were aged from 18 to 59 years and the mean age was 27.5 years. A significant association was observed between FSW age and the prevalence of C. trachomatis infection (Table 2). Younger women had a higher prevalence rate, indicating that age is an important risk factor for chlamydial infection. These results are comparable to other studies showing the same age effect in relation to infection in different cohorts such as FSWs, pregnant women, patients of STI clinics and asymptomatically infected women (Deák et al., 1997; Lan et al., 1995; Lima et al., 2007; Papadogeorgaki et al., 2006).

**Genotyping of C. trachomatis**

For genotyping purposes, the C. trachomatis-positive samples were analysed further by an omp1 PCR. All samples positive for the plasmid PCR were successfully amplified by the semi-nested PCR method and were suitable for further genotyping analysis. The DNA controls for reference C. trachomatis strains were also successfully amplified and genotyped.

### Table 2. Positive results for C. trachomatis in relation to age (n=275)

<table>
<thead>
<tr>
<th>Age group (years)*</th>
<th>Observation of positive test</th>
<th>Prediction of positive test†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Percentage</td>
</tr>
<tr>
<td>&lt;20</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>20–29</td>
<td>17</td>
<td>180</td>
</tr>
<tr>
<td>39–30</td>
<td>3</td>
<td>62</td>
</tr>
<tr>
<td>49–40</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>&gt;50</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>275</td>
</tr>
</tbody>
</table>

*No age data were available for 209 women.
†Age effect: P<0.001.

CI, Confidence interval.

### RESULTS AND DISCUSSION

The efficiency of the PCR-based RFLP method for the genotyping of different C. trachomatis serovars in clinical specimens has been proven by many studies (Gao et al., 2007; Lan et al., 1993, 1995; Morré et al., 2000; Van Duynhoven et al., 1998). To the best of our knowledge, this is the first study to determine the serovar distribution of C. trachomatis in a particular population in Hungary.

### Prevalence of C. trachomatis

A total of 484 endocervical specimens from FSWs was analysed in the study. All samples were successfully screened by β-globin PCR and no inhibition was detected. Of the 484 samples, 32 (6.6%) were found to be positive by C. trachomatis cryptic plasmid PCR. FSWs are regarded as a high-risk group for infection and transmission of C. trachomatis and other STIs. In a review article by Cwikel et al. (2008), it was pointed out that the prevalence of C. trachomatis in FSWs varies between 0.61 and 46.2% worldwide. According to these data and other studies (Folch et al., 2008; Mak et al., 2005; Papadogeorgaki et al., 2006; Rabenau et al., 2000; Resl et al., 2003; Ward et al., 2004), this rate usually appears to be less than 10% in Europe (Belgium 7.3%, Czech Republic 5.5%, Germany 9.6%, Greece 5.9%, Spain 5.9%, UK 4.8%) as detected by nucleic acid amplification tests, and these findings correspond with our results.

Data on age were obtained from 275 individuals. The FSWs were aged from 18 to 59 years and the mean age was 27.5 years. A significant association was observed between FSW age and the prevalence of C. trachomatis infection (Table 2). Younger women had a higher prevalence rate, indicating that age is an important risk factor for chlamydial infection. These results are comparable to other studies showing the same age effect in relation to infection in different cohorts such as FSWs, pregnant women, patients of STI clinics and asymptomatically infected women (Deák et al., 1997; Lan et al., 1995; Lima et al., 2007; Papadogeorgaki et al., 2006).

Statistical analysis. Logistical regression analysis was used to assess the effect of age on the probability of C. trachomatis infection. A comparison between predicted and observed probabilities was carried out to evaluate the validity of the fitted model. Data analysis was performed with SAS version 9.1 software.

### Table 2. Positive results for C. trachomatis in relation to age (n=275)

<table>
<thead>
<tr>
<th>Age group (years)*</th>
<th>Observation of positive test</th>
<th>Prediction of positive test†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Percentage</td>
</tr>
<tr>
<td>&lt;20</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>20–29</td>
<td>17</td>
<td>180</td>
</tr>
<tr>
<td>39–30</td>
<td>3</td>
<td>62</td>
</tr>
<tr>
<td>49–40</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>&gt;50</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>275</td>
</tr>
</tbody>
</table>

*No age data were available for 209 women.
†Age effect: P<0.001.

CI, Confidence interval.
performed for genotyping. Nevertheless, the true prevalence of C. trachomatis in the study group may have been rather underestimated. Although the possibility of the presence of inhibitors in the samples was well controlled with the bet-globin PCR, the plasmid PCR was not confirmed using other methods, e.g. Southern blot hybridization, and so could have resulted in false-negative results. However, the obtained prevalence rate in this study was similar to those observed in FSW populations in other European countries. Furthermore, our findings on serovar distribution correspond to other studies from different geographical regions worldwide. The prevalence of different C. trachomatis serovars has been examined in different cohorts comparing it to risk groups (FSWs and non-FSWs), gender and clinical symptoms, and the most prevalent serovars were also D, E and F/G (Gao et al., 2007; Geisler et al., 2003; Morré et al., 2000a; Van Duynhoven et al., 1998). In addition, the in silico analysis of RFLP patterns of the reference strains showed that this method could be a good alternative tool for characterizing C. trachomatis serovars if control strains are not available in the laboratory.

In conclusion, to the best of our knowledge, the present study is the first description of C. trachomatis serovar distribution in a high-risk group in Hungary. The predominating serovars were D, E and F, as has been reported from other countries. A higher prevalence of C. trachomatis infection was found among younger women. This PCR-based RFLP analysis is a valuable molecular method for investigating the prevalence of C. trachomatis serovars in epidemiological surveys. Further studies are required to compare the distribution of C. trachomatis serovars in different cohorts as well as in the general population, and to investigate the correlations between serovars and clinical manifestations in Hungary.

ACKNOWLEDGEMENTS

The authors thank Dr Hamid Jalal and Dr Janet Burton (Addenbrooke’s Hospital, Cambridge, UK) for providing the DNA controls of the reference C. trachomatis serovars.

REFERENCES


