Original Article

Molecular epidemiology and genotyping of *Chlamydia trachomatis* infection in a cohort of young asymptomatic sexually active women (18-25 years) in Milan, Italy

S. BIANCHI¹, E.R. FRATI¹, M. CANUTI¹, D. COLZANI¹, E. FASOLI¹, A. AMENDOLA¹,²,³, E. TANZI¹,²,³

¹Department of Biomedical Sciences for Health, University of Milan, Italy; ²Department of Biology, Memorial University of Newfoundland, Canada; ³CIRI-IT, Department of Health Sciences, University of Genoa, Italy; ⁴Coordinated Research Center “EpiSoMi”, University of Milan, Italy

Keywords

C. trachomatis infection • Sexually active asymptomatic young women • C. trachomatis genovars • Molecular epidemiology • Sexually transmitted diseases

Summary

Introduction. *Chlamydia trachomatis* (Ct) is the most common bacterial cause of sexually transmitted infections (STI) and is associated with severe long-term sequelae in female populations. In Italy Ct infections are not submitted to a screening programme, and its epidemiological profile is understudied. Even scarcer information is available about the genetic diversity on ompA gene, whose sequence defines 18 different genovars. This study aims at evaluating the prevalence of Ct infection in young sexually active asymptomatic women aged 18-25, and characterizing the molecular epidemiology of the different circulating genovars in this population.

Methods. Cervical samples collected from 909 sexually-active-young women (mean age 21.5 years) were analyzed through molecular assay for the detection of Ct infection. Phylogenetic analysis on the ompA gene was performed on Ct positive samples to identify the circulating genovars.

Results. The overall prevalence of Ct-infection was 4.4% (95%CI: 3.2-5.9%): 5.3% among women aged 18-21 years and 3.5% among those aged 22-25 years. Phylogenetic analysis has identified 5 different genovars: D, E, F, G, and H. The most common genovar was the E (46%), followed by genovar F and G (18.9% each), D (13.5%), and H (2.7%).

Conclusions. This study underlines the high prevalence of asymptomatic Ct-infections among young women. Overall, about half of the asymptomatic infections is sustained by genovar E. The introduction in Italy of a systematic screening program should be considered to allow a better understanding of Ct spreading and providing women with an opportunity for early treatment to protect their sexual and reproductive health.

Introduction

*Chlamydia trachomatis* (Ct) is an obligate intracellular Gram-negative bacterium that is the most commonly reported microorganism responsible for sexually transmitted infections (STI) in Europe and is the cause of considerable acute morbidity and long term reproductive health problems, particularly in young people [1, 2]. Many infections are asymptomatic and result in delayed diagnosis and uninterrupted transmission. This is of particular clinical relevance as untreated infections can ascend in the female genital tract and cause Pelvic Inflammatory Disease (PID), which includes any combination of endometritis, salpingitis, tubo-ovarian abscess and pelvic peritonitis. PID can also result in ectopic pregnancy, infertility and chronic pelvic pain [2-4].

The presence of few or no specific symptoms that, if untreated, could lead to an increase in female reproductive tract morbidity and the existence of an inexpensive treatment against all these infections make chlamydia screening widely recommended to all sexually active women aged 25 or less [5-7]. In fact, Ct is under epidemiological surveillance in several European regions and the reported prevalence in young sexually active women ranges from 5% to 10%. Infection rates are the highest in females below 20 years of age and decrease with increasing age [8, 9]; in 2007, the incidence of chlamydial infection was 4.5-fold higher in the age group 15 to 24 years than in the age group 25-44 years [9].

In Europe the incidence of genital Ct infections seems to have increased dramatically over the last 20 years. However, this is most likely the result of more extended testing rather than a true rise in the incidence. Targeted screening, opportunistic testing for asymptomatic infections, contact tracing and mandatory notification help to explain the high notification rates in the UK and the Scandinavian countries compared with other European states [10].

In Italy, a reporting system for Chlamydia infections is not in place and an organised screening in specific asymptomatic population groups is not available. Surveillance data
are collected centrally by the Istituto Superiore di Sanità (ISS) by a sentinel network system consisting of 13 microbiology laboratories located throughout the Country. The last report, regarding data until 2012, described an overall prevalence of 2.4% in women aged between 15 and 45 years, with a significantly increased prevalence in lower age groups [11]. Moreover, a recent study identified 5.2% of Ct endocervical infection prevalence in a large population of sexually active women aged 15-55 years attending an outpatient service of cervico-vaginal pathology unit in Rome over a 10-year period [12]. Notwithstanding the existence of these studies, knowledge on the prevalence and molecular epidemiology of Chlamydia genital tract infections, particular in young asymptomatic women, remains modest.

Molecular epidemiology studies are currently based on the analysis of differences in the Ct major outer membrane protein (MOMP), whose coding gene (ompA) contains four spaced variable domains. Genetic variability on ompA gene reflects the existence of 18 genovars, classified according to their pathogenic potential [13]. Genovars A, B, Ba, and C have been commonly associated with trachoma, D-K, Da, Ia, and Ja with urogenital infections, and L1-L3 with lymphogranuloma venereum [13]. In addition, on the basis of amino acid similarities, these genovars have been grouped into the following groups or classes: group B (B, Ba, D, Da, E, L1, and L2), group C (A, C, H, I, Ia, J, Ja, K, and L3), and the intermediate group (F and G) [14]. These type of studies are useful to obtain new information about Ct genovars distribution, which can be further on translated into improved strategies for Ct infection management, for the traceability of sexual contacts, and in developing strategies for vaccine design [15, 16].

To improve our knowledge about Ct epidemiology and to expand the available Italian data, we evaluate here the prevalence of Ct infections and the molecular epidemiology of circulating genovars in young asymptomatic sexually active women aged 18-25 in Milan, Italy, over a period of 6 years (2008-2013).

**Materials and methods**

**Study population**

This was a cross-sectional study, performed on female subjects who spontaneously visited gynecological centers of the Local Health Units (LHUs) of Milan for medical consultations for contraception. No evidence of clinical symptoms related to Ct infection were described by the physician. Cervical cytological samples were collected for routine test among women aged 18 to 25 years, from September 2008 to June 2013. Written informed consent was obtained from young women so as to store their samples for further anonymous research testing. Due to its design, ethical approval was not required for this study, in compliance with the international policy [17] and with the current Italian legislation [18, 19]. The database was anonymised before the analysis.

**Samples collection**

Cervical cytological samples were collected using a brush (Cytobrush Plus Medscand® Medical AB, Sweden), immersed and rinsed in a vial filled with 20 ml of PreservCyt® Solution (PreservCyt) and stored at room temperature (RT) until processing. A total of 10 ml of each PreservCyt® Solution containing cervical cells was centrifuged at 3800xg for 15 min at RT. After centrifugation the pellet was re-suspended in 1 ml of Phosphate Buffered Saline (PBS), transferred in a new 1.5 ml collection test tube, and stored at -20°C until nucleic acids extraction.

**DNA extraction and amplification**

DNA was extracted with the EasyMAG kit (NucliSens® easyMAG®, bioMérieux, France) from 500 µl of the re-suspended pellet with a final elution volume of 100 µl. The concentration and purity of the extracted DNA was evaluated through a spectrophotometer (NanoDrop ND-2000/200C, Euroclone®, Thermo Scientific, Wilmington, DE, USA). After DNA extraction molecular tests were promptly performed to limit the storage to a maximum of a week. DNA integrity was assessed by the amplification of a 268 bp fragment of the ubiquitous beta-globin gene using the primer pair GH20 and PCO 4 [20]. A nested PCR targeting a Ct cryptic plasmid was performed to screen each extracted sample. The primers used for the nested PCR were previously described by Jalal et al. [21] and amplify a fragment of 150 nt. Amplifications were performed on 50 ng of extracted DNA in a 50 µl reaction mix containing 5x PCR Buffer, 200 µM dNTPs, 25 pmol of each primer, and 1 U Taq (GoTaq® DNA Polymerase 5U/µl Promega, Madison WI). Each run was accompanied by positive and negative control samples. The cycling conditions were as follow: 94°C for 5 min, followed by 30 cycles of amplification during the first step and 25 cycles during the nested step consisting of 94°C for 30 sec, 50°C for 30 sec, 72°C for 30 sec, and a final 72°C for 7 minutes extension. The final amplification products were visualized using electrophoresis analysis on a 2% agarose gel containing ethidium bromide (0.5 mg/L) and compared with a standard (DNA Molecular Weight, Marker 100, SigmaAldrich, St. Louis, MO, USA).

All Ct-DNA positive samples were used to amplify a 395 bp fragment of the Ct ompA gene with previously described primer sets [21]. Amplifications were performed on 50 ng of extracted DNA in a 50 µl reaction mix containing 5x PCR Buffer, 200 µM dNTPs, 25 pmol of each primer, and 1 U Taq (GoTaq® DNA Polymerase 5U/µl Promega, Madison WI). The cycling conditions were: 94°C for 5 min, followed by 30 cycles of amplification during the first step and 25 cycles during the nested step consisting of 94°C for 30 sec, 50°C for 30 sec, 72°C for 30 sec, and a final 72°C for 7 minutes extension. The final amplification products were visualized using electrophoresis analysis on a 2% agarose gel containing ethidium bromide (0.5 mg/L) and compared with a standard (DNA Molecular Weight, Marker 100, SigmaAldrich, St. Louis, MO, USA).

All Ct-DNA positive samples were used to amplify a 395 bp fragment of the Ct ompA gene with previously described primer sets [21]. Amplifications were performed on 50 ng of extracted DNA in a 50 µl reaction mix containing 5x PCR Buffer, 200 µM dNTPs, 25 pmol of each primer, and 1 U Taq (GoTaq® DNA Polymerase 5U/µl Promega, Madison WI). The cycling conditions were: 94°C for 5 min, followed by 30 cycles of amplification during the first step and 25 cycles during the nested step consisting of 94°C for 30 sec, 50°C for 30 sec, 72°C for 30 sec, and a final 72°C for 7 minutes extension. The final amplification products were visualized using electrophoresis analysis on a 2% agarose gel containing ethidium bromide (0.5 mg/L) and compared with a standard (DNA Molecular Weight, Marker 100, SigmaAldrich, St. Louis, MO, USA).
sequencing on the ABI PRISM 3100 genetic analyzer (Applied Biosystem, CA, USA).

**Phylogenetic analysis**

Molecular characterization was performed by analysis of 395 bp amplicons of the ompA Ct gene (nucleotides 223-636 of Ct genovar A sequence, NC007429). All specimens presenting ompA variant sequences were confirmed by resequencing newly extracted DNA. All sequences obtained in this study were deposited into NCBI GeneBank Database [27], under accession numbers: KX449367-KX449406. The reference genovars used for the construction of phylogenetic trees were obtained from the NCBI GenBank Database (genovar A: NC007429, M58938; genovar B: AF063194, U80075, M33636; genovar C: M17343; genovar D: NC000117, X62920, X62918, X62919; genovar E: X52557; genovar F: X52080; genovar G: AF063199; genovar H: X16007; genovar I: AF063200; genovar J: AF063202; genovar K: AF063204; genovar L1: M36533; genovar L2: M14738; genovar L3: X55700). Sequences were aligned using ClustalX 2.1 multiple aligner [22] and then used for phylogenetic inference. A model selection was performed to identify the best model for distance estimation. The evolutionary history was inferred using the maximum-likelihood method [23] based on the Tamura 3 parameter model [24], identified as the best fitting model after the model test analysis, using MEGA 6.06 [25]. A discrete Gamma distribution was used to model evolutionary rate differences among sites (G = 0.8952) and phylogenetic trees were constructed with MEGA 6.06. A bootstrap test [26] with 1,000 replicates was performed to test the robustness of the analyses.

**Statistical analysis**

Data were expressed as mean (range) and percentages (95% confidence intervals, 95% CI) where appropriate. Comparisons between groups were performed using the chi-square test or Fisher’s exact test. A p-value < 0.05 was considered statistically significant (two-tailed test). All statistical analyses were performed using OpenEPI software, version 2.3.1 [28].

**Results**

**Prevalence of Ct infection**

A total of 909 samples collected from the same number of sexually active young women (mean age 21.5 years, range 18-25 years) were available for this study. All women were asymptomatic for Ct infection. The beta-globin gene was successfully amplified from all 909 cervical samples collected, confirming the suitability of the extraction method for these biological samples (extracted DNA: mean 20.1 ng/µl, range [3.3-45.7 ng/µl]).

The overall prevalence of Ct infections was 4.4% (95% CI, 3.2-5.9%). The prevalence of Ct infection was the highest among 20-21 year-old women with a value of 5.5% (95% CI, 3.0-9.3%) and decreased to 3.5% (95% CI, 1.5-6.7%) and 3.6% (95% CI, 1.6-6.9%) in age groups 22-23 and 24-25 years (Tab. I). Differences between infection rates in the different age groups and during the different years of the study period were not significant.

**Molecular characterization**

OmpA amplification was successful for 37 out of 40 positive samples (92.5%). Phylogenetic analysis of the ompA partial nucleotide sequences demonstrated 5 different genovars: D, E, F, G, and H (Fig. 1). In particular, 2 genovars (D and E) fell into subgroup B, 2 (F, and G) fell into the intermediate subgroup, and one, genovar H, fell into subgroup C. All circulating genovars are normally associated with infection of the urogenital tract. The most common genovar in our population was E (46%, 17/37), followed by genovars F and G (both 18.9%, 7/37), D (13.5%, 5/37), and H (2.7%, 1/37). No genovar distribution pattern was observed among the various age groups (Fig. 2), and during the different years of the study period (Fig. 3).

**Discussion**

*Chlamydia trachomatis* is the most common bacterium associated with STIs and genovars D-K cause genital tract infections in women (cervicitis and urethritis) and men (urethritis). These can also be responsible for sexually transmitted rectal and pharyngeal infections, be transmitted during labor, cause pneumonia and eye infections in infants, and be spread by close contact to cause eye infections in adults. Several data suggest that, even though young people aged 15-24 years represent only 25% of the sexually experienced population, they acquire nearly half of all new STIs [29]. Rates of reported chlamydial infection among persons aged 15-19 years and 20-24 years continue to increase. Overall, during 2009-2010, rates increased up to 2.8% and 7.5% in the age groups 15-19 and 20-24 years, respectively [30]. Compared with older women, sexually active adoles-

---

**Tab. I. Prevalence of C. trachomatis infections in different age groups of sexually active young women (18-25 years old).**

<table>
<thead>
<tr>
<th>Age group</th>
<th>Number of subjects</th>
<th>Ct-DNA +</th>
<th>Prevalence</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-19</td>
<td>220 (24.2%)</td>
<td>11 (27.5%)</td>
<td>5.0%</td>
<td>2.5-8.8</td>
</tr>
<tr>
<td>20-21</td>
<td>235 (25.9%)</td>
<td>13 (52.5%)</td>
<td>5.5%</td>
<td>3.0-9.3</td>
</tr>
<tr>
<td>22-23</td>
<td>230 (25.5%)</td>
<td>8 (20.0%)</td>
<td>3.5%</td>
<td>1.5-6.7</td>
</tr>
<tr>
<td>24-25</td>
<td>224 (24.6%)</td>
<td>8 (20.0%)</td>
<td>3.6%</td>
<td>1.6-6.9</td>
</tr>
<tr>
<td>Total</td>
<td>909 (100.0%)</td>
<td>40 (100.0%)</td>
<td>4.4%</td>
<td>3.2-5.9</td>
</tr>
</tbody>
</table>
cents aged 15-19 years and young adults aged 20-24 years are at higher risk of acquiring STIs for a combination of behavioural, biological, and cultural reasons [31].

This study focused on asymptomatic sexually active young women and describes the molecular epidemiology of Ct in Milan, Italy, over a period of 6 years. The prevalence of the infection was 4.4%, in line with what reported by other Italian studies [11, 12, 32, 33].

The most common genovars found in our populations were E (46%), F and G (both to 18.9%). Other recent studies described E (50-33%) and F (25-14%) as the most common genovars in symptomatic female populations [34-36]. A previous study, conducted in Italy in symptomatic and asymptomatic populations attending a Sexually Transmitted Disease (STD) outpatient clinic, described genovar E as the most prevalent in female (37.2%), followed by genovar G (32.6%), and genovar F

---

**Fig. 1.** Phylogenetic analysis of C. trachomatis partial ompA gene sequences (coded between nt 223-636 of C. trachomatis genovar A, accession number NC007429) obtained during this study and compared to reference sequences. The outcome of the bootstrap analysis is shown next to the nodes, and branch lengths are proportional to genetic distances as indicated by the scale bar. Genovars as well as subgroups are indicated on the right.
and J (both to 7%), regardless of presence or absence of related symptoms [37].

It has been demonstrated that genovars E and F have a biological advantage over the other genovars thanks to both their ability to escape the host immune response and the presence of specific virulence factors, which can facilitate the transmission and infectious processes [38, 39]. It is possible that genovars E and F are less immunogenic than other genovars and, therefore, remain the most prevalent strain in all populations,
regardless of the presence or absence of clinical signs. Although we could not detect any relationship between infection, related genovar, and absence of symptoms, the present study contributed to increase the current knowledge on genotype distribution of Ct in asymptomatic young women in northern Italy.

Conclusions

Our data indicate that Ct infections occur frequently in young sexually active women and that several different genovars are widely spread in Italy. Larger national and longitudinal studies are definitively required to better understand the spread of Ct infection and its impact on the young Italian population.

These results underscore the need to establish primary and secondary preventive measures and to allocate more resources for an adequate prevention of Ct infections. The introduction of an opportunistic screening program for Ct in Italy should be evaluated in order to achieve early treatment of infected subjects and reduce associated clinical manifestations. Finally, it seems essential to develop educational programs and information campaigns, particularly addressed to young people, about acquisition, risk factors and treatment of STIs with a particular focus on C. trachomatis infection. The Authors have no conflict of interest to declare.

Acknowledgments

The Authors thank Valentina Paladino for her helpful technical-scientific assistance. The Authors have no conflict of interest to declare.

Authors' contributions

SB: study design, project and protocol development, data analysis and manuscript writing. ERF: project and protocol development, phylogenetic analysis and manuscript writing. DC: phylogenetic analysis and manuscript writing. MC: sample and data collection, protocol development and critical revision of the manuscript. AA: study design, supervision of the study and manuscript writing. ET: conception, design, coordination and supervision of the study, manuscript writing.

References


