Molecular epidemiology of Chlamydia trachomatis
Bom, R.J.M.

Citation for published version (APA):
Bom, R. J. M. (2014). Molecular epidemiology of Chlamydia trachomatis

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
1.

General introduction
1. General Introduction

*Chlamydia trachomatis* infections are the most prevalent bacterial sexually transmitted infections (STI) worldwide.¹ In the Netherlands, most infections are found among heterosexual adults under 25 years in age, certain migrant groups and men who have sex with men (MSM).² Among the latter, an outbreak of a more aggressive strain of *C. trachomatis*, which causes lymphogranuloma venereum, is seen since 2003.³ Although *C. trachomatis* infections are often asymptomatic, late complications, such as pelvic inflammatory disease, may occur, which can ultimately lead to infertility.¹ In addition, *C. trachomatis* infections may facilitate the transmission of HIV.¹ Due to the high prevalence, *C. trachomatis* infections are a large burden on society, from a public health perspective and from an economic perspective.⁴ A better understanding of the transmission of *C. trachomatis* may contribute to improved screening and prevention programs in the future and ultimately alleviate this burden.

**Biology of *Chlamydia trachomatis***

*C. trachomatis* is an obligate intracellular pathogen of eukaryotic cells; a trait shared with all other members of the phylum *Chlamydiae*.⁵,⁶ Therefore this evolutionary strategy must have already evolved in an ancestral bacterium within the *Planctomycetes-Verrucomicrobia-Chlamydiae* superphylum.⁷-⁹ Indeed this parasitic intracellular lifestyle is seen among other members of this superphylum, but mutualistic and commensal symbionts are also found, as well as free living bacteria with no relationship to an eukaryotic cell.⁷,⁸,¹⁰ As the *Chlamydiae* line already started diverging during the Precambrian period, the relationship between eukaryotic cells and *Chlamydiae* originates from this era when primordial eukaryotic protozoa became abundant.⁵,⁷,¹¹ At this moment in time, the dimorphic developmental cycle characteristic for *Chlamydiae* must have evolved as well.⁵,⁶,¹²

The chlamydial developmental cycle alternates between an extracellular and an intracellular phase (*Figure 1*).¹³-¹⁶ All

*Figure 1. The chlamydial developmental cycle. 1–2. Elementary bodies (black) invade the host cell and form inclusions (green). 3–4. In the inclusions, the elementary bodies differentiate into reticulate bodies (red) and replicate through binary fission. 4–6. The reticulate bodies differentiate to elementary bodies. 7. The host cell ruptures, releasing the elementary bodies. Adapted from Morais et al.¹⁶*
infections start with adhesion to and invasion of the eukaryotic host cell by the elementary bodies, the infectious, but metabolically inert spore like forms of the organism. Upon infection of the cell, the elementary bodies remain within membrane bound vacuoles, the inclusions, where they differentiate into reticulate bodies. These are the non infectious, but metabolically active chlamydial forms. The inclusions segregate from the endocytic pathway to avoid fusion with lysosomes and are transported to the peri Golgi region. Here, the reticulate bodies interact with trafficking pathways and host cell compartments to acquire host derived nutrients. The reticulate bodies replicate through binary fission up to the point that the inclusions contain about a 1000 reticulate bodies, which start differentiating to elementary bodies. The increasing size of the inclusions cause the host cell to rupture, releasing the elementary bodies to the extracellular environment, where they can find a new host cell.

At least 700 million years ago, the Chlamydiae phylum started diverging into multiple families (Figure 2). So far, eight families have been described: Chlamydiaceae, Clavochlamydiaceae, Criblamydiaceae, Parachlamydiaceae, Piscichlamydiaceae, Rhabdochlamydiaceae, Simkaniaceae, and Waddliaceae. However, more families are expected to be discovered in the coming years. Members of these families have adapted to a broad range of eukaryotic host cells and have interacted with their host during their evolution. Most chlamydial families still interact with simple unicellular eukaryotic protozoa and have a diverse host range in which they often show minimal to no pathogenic effects. One early branching family however, the Chlamydiaceae, has adapted to higher multicellular eukaryotic hosts and their interactions with their host became much more specific and pathogenic.

As a result of this more specialised lifestyle, the genome of Chlamydiaceae has reduced considerably. Whereas other chlamydial species have a genome of 2 to 3 Mb, members of the Chlamydiaceae family have a genome of about 1 Mb, which includes about 900 genes; this is one of the smallest genomes within the bacterial kingdom. The genome is highly conserved among all members within the Chlamydiaceae family, both in gene content, as in genomic synteny. During its developmental cycle, virtually every gene within the genome is expressed at some point, showing that the genome has almost no facultative capacity and that it has been minimised to an evolutionary optimum. The same is seen in the chlamydial plasmid, which is highly conserved among all lineages and has resulted from a single acquisition. As a result of its isolated lifestyle, virtually no horizontal gene transfer of plasmid or genomic content has occurred.

The Chlamydiaceae family comprises one genus, Chlamydia, in which nine species have been described so far, i.e. Chlamydia abortus, Chlamydia caviae, Chlamydia felis, Chlamydia muridarum, Chlamydia pecorum, Chlamydia pneumoniae, Chlamydia psittaci, Chlamydia suis, and Chlamydia trachomatis (Figure 2).
Members of the *Chlamydiaceae* family can infect amphibians, reptiles, birds and mammals.\(^{20}\) Although some species can be zoonotic, most of the host range diversity originates from co-evolution along the evolutionary radiation of their hosts during the Paleocene period.\(^{17,21}\) Therefore they are endemic to at least 469 species of birds, comprising 30 orders, and can be found in a broad range of mammals, including marsupials.\(^{20,22,23}\) Among humans, all scenarios of transmission can be found. *C. abortus* and *C. psittaci* are acquired zoonotically from ruminants and birds, and no transmission from human to human has been described.\(^{24}\) *C. pneumoniae* is transmitted from human to human and no transmission from animal to human has been documented.\(^{24}\) Phylogenetic analysis however, showed *C. pneumoniae* infections have been acquired from the large animal reservoir in which the pathogens reside.\(^{23,24}\) *C. trachomatis* is strictly a human pathogen and is thought to have co-evolved along the human evolution from primate to man.\(^{17,25}\) During this evolutionary trajectory, *C. trachomatis* has adapted to multiple ecological niches within the human body, causing distinct clinical manifestations between different variants of the pathogen.

**Pathogenesis and clinical manifestations of *Chlamydia trachomatis***

*C. trachomatis'* main target cells are the columnar epithelial cells of the mucosa.\(^{26-28}\) The infection spreads over the epithelium by the release of elementary bodies along the apical surface of the mucosa, which subsequently infect the neighbouring cells.\(^{26}\) The body reacts to the infection with the recruitment of neutrophils and mononuclear leukocytes, and with the secretion of cytokines, leading to inflammation of the infected site.\(^{26,28,29}\)
Upon clearance of the infection, fibrosis of the damaged and necrotic tissue can occur.\textsuperscript{28} While most initial \textit{C. trachomatis} infections have minor symptoms, repeated or persistent infections can lead to substantial scarring of the infected tissue and irreversible pathological damage of the infected organ.\textsuperscript{28,29} The clinical manifestations of these infections occur at different anatomical locations, depending on the tropism for a certain tissue of the chlamydial strain. For \textit{C. trachomatis}, three distinct biovars can be discerned, i.e. trachoma, urogenital infections and lymphogranuloma venereum (LGV).

Trachoma inducing \textit{C. trachomatis} strains preferentially infect the mucosa of the inner eyelids, the conjunctiva.\textsuperscript{30,31} For the initial episode of the infection, symptoms are usually mild, but repeated infections lead to scarring of the eyelids. As this scarring continues, the eyelids fold inwards, causing the eyelashes to rub the cornea. This leads to damage of the cornea, making it opaque with irreversible blindness as a consequence for the patient. Transmission occurs through direct contact of eyes or fingers, but can be facilitated through fomites, like face cloths, or through eye seeking flies.\textsuperscript{30,31}

Urogenital \textit{C. trachomatis} infections are mainly found in the urethra in males and in the cervix, vagina and urethra in females and these infections are often asymptomatic.\textsuperscript{27} Repeated or persistent infections however, can ascend in the genital tract in women, leading to inflammation of the uterus, fallopian tubes and ovaries.\textsuperscript{27} This is called pelvic inflammatory disease (PID), and the consequent extensive scarring of the fallopian tubes may ultimately lead to infertility.\textsuperscript{29,32} Also in men, the infection can ascend to the prostate, epididymides and testicles, but infertility due to scarring is rare.\textsuperscript{32} Anal intercourse can lead to infection of the rectal mucosa and oral sex might lead to infection of the nasopharynx, although this is not researched thoroughly.\textsuperscript{33,34} Urogenital \textit{C. trachomatis} strains can also cause infection of the eyes and respiratory tract of newborns upon birth from an infected mother.\textsuperscript{32} Urogenital \textit{C. trachomatis} infections are transmitted through direct sexual contact and urogenital secretions.

LGV is also induced by \textit{C. trachomatis} strains with a tropism for urogenital tissues, but these strains are invasive and have a more severe course of infection.\textsuperscript{35} Infections with LGV inducing strains begin in the urogenital or rectal mucosa, but in contrast to urogenital strains, these strains are able to exit the basolateral side of the epithelial cells, invade the underlying connective tissue, and spread subsequently to the lymph nodes.\textsuperscript{27} If these lymph nodes become abscessed and rupture, this will lead to fistulae and impaired lymph drainage.\textsuperscript{35} LGV inducing strains are transmitted through direct sexual contact and urogenital secretions. Distinction between infections with urogenital strains and LGV inducing strains is critical, as LGV requires a prolonged treatment regimen, due to the more invasive character of its-inducing strain.\textsuperscript{36}

**Epidemiology of \textit{Chlamydia trachomatis}**

The epidemiology of \textit{C. trachomatis} heavily
depends on the biovar to which the strain belongs. Trachoma was once a major health problem throughout the world, but has disappeared from high income countries, because of general improvements in living and hygiene standards. Nowadays trachoma is largely found in poor, rural communities in low income countries in sub Saharan Africa, but the disease is also endemic in the Middle East, Asia, Latin America and the Western Pacific.

Urogenital *C. trachomatis* infections are highly prevalent throughout the world. They are endemic to the general population, but some subpopulations have a higher prevalence. A major risk group are sexually active heterosexual adolescents and young adults, and this risk is related to sexual experience, changing sexual partners, and the number of new sexual partners. In addition, certain racial and ethnic groups are disproportionally affected as well. This is thought to be the result of differences in socio economic status and partnership structures. In many countries *C. trachomatis* infections are highly prevalent among female sex workers and their clientele. Lastly, among MSM the prevalence is high.

Like trachoma, LGV was considered a tropical disease, endemic to parts of Africa, Asia, Latin America, and the Caribbean. However, in 2004 a cluster of LGV cases was reported among MSM in Rotterdam, the Netherlands. These infections must have circulated in the Netherlands at least since 2002 and nowadays this outbreak is ongoing within mainly HIV positive MSM throughout Europe, North America and Australia.

**Typing of *Chlamydia trachomatis***

When it was discovered that *C. trachomatis* could be propagated in and isolated from yolk sacs of embryonated eggs, it became possible to study the pathogen in more detail. Injecting crude yolk sac suspensions in mice led to the discovery of differences in cross protectiveness between different strains. This implied that serological variation existed within *C. trachomatis*. After the development of cell cultures and serological tests, the 14 known serovars were characterised, i.e. A to K and L1 to L3. It was subsequently discovered that this serological variation was predominantly determined by only one membrane protein, called major outer membrane protein or MOMP, which on its turn was encoded by a ~1200 bp long gene, *ompA*. With the arrival of molecular amplification techniques, *C. trachomatis* samples could be more sensitively and specifically detected. In addition, *C. trachomatis* samples could be typed from direct patient material and cell cultures were therefore no longer needed. The first molecular typing methods of *C. trachomatis* strains were restriction fragment length polymorphism (RFLP) and sequence based typing techniques of the *ompA* gene. Although genetic variation could be found within the known serological MOMP variants, this observed variation is rare and the MOMP serovar/*ompA* genovar designation still stands as a reference until the present day.

The variants found within *C. trachomatis*
were largely overlapping with the biovars, based on clinical manifestations.\textsuperscript{49} Trachoma is caused by $ompA$ genovar A, B and C strains and urogenital infections by genovars D to K, although a small proportion of urogenital genovar B infections are consistently found. LGV is induced by the L strains. Many molecular epidemiological studies, especially on urogenital infections, have therefore used MOMP or $ompA$ typing to discriminate between strains to elucidate transmission patterns or clinical symptoms.\textsuperscript{49} 

So far, these molecular epidemiological studies have resulted in little additional information. Although a considerable amount of antigenic variation exists between genovars, epidemiologically distinct risk groups may have identical $ompA$ genovars.\textsuperscript{59} More problematic is that $C.\ tractomatis$ has a nearly identical distribution of genovars in most populations, which seems to be independent of host risk group, geography, calendar time or clinical symptoms.\textsuperscript{49,59-63} In heterosexual populations, nearly always all different genovars are found, with

\begin{figure}
\centering
\includegraphics[width=\textwidth]{phylogenetic_tree.png}
\caption{Phylogenetic tree of $Chlamydia\ tractomatis$ using full genome sequences. Within this tree, four distinct clades of strains can be recognised: the LGV-inducing strains (yellow), the prevalent urogenital strains (dark blue), the trachoma-inducing strains (light blue) and the rarer urogenital strains (red). The letters on the right indicate the $ompA$ genovar of the sequenced strains. Adapted from Joseph et al.\textsuperscript{70}}
\end{figure}
genovar E consistently being the most prevalent, followed by genovars F and D. The only notable exception is the genovar distribution found among MSM. Within MSM populations throughout the world, genovars D, G and J are the predominant types, while the other variants are mostly absent or rare. Therefore, the need for higher resolution typing methods in molecular epidemiological research is apparent. A few years ago two methods with a degree of resolution needed for this kind of studies were published. In 2007, Klint et al. published a multilocus sequence typing (MLST) method for *C. trachomatis* that included five variable regions: *hctB*, CT058, CT144, CT172, and *pbpB*. A second technique, a multilocus variable number of tandem repeat (VNTR) analysis (MLVA) published by Pedersen et al. in 2008, combined *ompA* typing with analysing three highly variable single nucleotide repeats: CT1291, CT1299, and CT1335. Studies using these techniques have confirmed the clonal character of the LGV outbreak among MSM and the outbreak of the so called new variant *C. trachomatis* in Sweden.

Another problem with *ompA* typing, is that phylogenetic analysis of *ompA* subdivides the variants into three distinct clades: the B complex (genovars B, D, E, L1, and L2), the C complex (genovars A, C, H, I, J, K, and L3) and the intermediate complex (genovars F and G). This subdivision is incongruent with the biovar designation. Recent studies using full genome sequences of various *C. trachomatis* strains showed that this phylogenetic incongruence of *ompA* is the result of numerous homologous recombination events of the gene between the different strains. These whole genome analyses subdivide the *C. trachomatis* strains in four distinct clades (Figure 3). The first clade to branch off contains all LGV-inducing strains. The second branch, surprisingly, is a clade that contains the prevalent urogenital genovars E, D, F and J. The remaining tree is split into a clade containing the trachoma-inducing strains and a clade of rarer urogenital genovars G, H, I and K, but also some genovar D and J strains. As the urogenital biovar is split into two distinct clades, it is possible that biological differences exist between these two clades that have not been noticed before because of the incongruence of the *ompA* genovar designation. It has been speculated that clade of rarer urogenital strains has an increased affinity for rectal tissue. Although genovars D, G, and J are successfully propagated through anal intercourse within the MSM populations, no such relation exists for the genovars H, I and K. Therefore, more basic research on these biological differences is required to explore these new phylogenetic insights.

**Aims and outline of the thesis**

In Chapter 2, we report on an evaluation of the diagnostic performance of a newly developed *pmpH* real time PCR as a discrimination assay between LGV and non LGV-inducing *C. trachomatis* strains, by a comparison with a reverse hybridisation assay and *ompA* sequencing. In addition, we report the non LGV genovar distribution in...
rectal samples from MSM and investigate the occurrence of double infections in men infected with LGV and non-LGV-inducing strains.

In Chapter 3, we investigate which of the methods is most suitable for molecular epidemiological analysis of *C. trachomatis* transmission patterns in sexual networks. We adapt the published high resolution typing methods to be more suitable for clinical samples. We assess both the minimal variation and the resolution of these typing methods compared with *ompA* sequencing. To test whether the typing methods are useful for molecular epidemiological research, a panel of samples from *C. trachomatis* infected heterosexual couples is selected.

In Chapter 4, differences in circulating *C. trachomatis* strains between MSM and heterosexuals are investigated using a modified MLST. Samples are collected from both MSM and heterosexual men and women visiting a single STI clinic in a relatively short time frame. We investigate the diversity of chlamydial genotypes and analyse epidemiological characteristics of *C. trachomatis* MLST clusters between the risk groups. To study geographical variation, we investigate samples from MSM from the Netherlands, Sweden, and the United States, and samples from women from the Netherlands and Sweden. We discuss the role of sexual networks as an explanation for the different *C. trachomatis* genotypes in MSM and heterosexual populations and examine whether tissue tropism could be an alternative explanation. Lastly, we assess whether circulating *C. trachomatis* strains are linked to certain subpopulations of MSM, as characterised by demographics, sexual risk behaviour, sexual partnerships, and lifestyle.

In Chapter 5, we assess whether Surinamese migrants in the Netherlands form a bridge population facilitating transmission of *C. trachomatis* between Suriname and the Netherlands. We investigate the sexual mixing with native Surinamese and native Dutch partners and compare the distribution of *C. trachomatis* genotypes found among Surinamese migrants with those found among the native Surinamese or native Dutch population. In addition, we elucidate determinants for *C. trachomatis* infections in Suriname, such as ethnicity and ethnic sexual mixing, and identify transmission patterns and sexual networks using molecular epidemiological network analyses. Lastly, we investigate the effect of geographical distance on the distribution of *C. trachomatis* genotypes by comparing strains found among heterosexuals from China with those found in the Netherlands.

In Chapter 6, we discuss our main findings and based on recent literature, we make recommendations for public health implementations and future research.

**References**

2. Triekens SCM, Koedijk
Table 1. Various panels of samples used in the studies of the thesis.

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Group</th>
<th>Country</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Women</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>MSM</td>
<td>The Netherlands</td>
<td>-</td>
</tr>
<tr>
<td>3.1</td>
<td>Women</td>
<td>The Netherlands</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>The Netherlands</td>
<td>-</td>
</tr>
<tr>
<td>4.1</td>
<td>Women</td>
<td>The Netherlands</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>The Netherlands</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>MSM</td>
<td>The Netherlands</td>
<td>-</td>
</tr>
<tr>
<td>4.2</td>
<td>Women</td>
<td>The Netherlands</td>
<td>Panels from Chapter 2.2 and 3.1 were included.</td>
</tr>
<tr>
<td></td>
<td>MSM</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>Sweden</td>
<td>Panels from Chapter 2.2 and 3.1 were included.</td>
</tr>
<tr>
<td></td>
<td>MSM</td>
<td>Sweden</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM</td>
<td>The United States</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>MSM</td>
<td>The Netherlands</td>
<td>Panel from Chapter 4.1 was included.</td>
</tr>
<tr>
<td>5.1</td>
<td>Women</td>
<td>The Netherlands</td>
<td>Panel from Chapter 4.1 was included.</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>Suriname</td>
<td>Panel from Chapter 5.1 was included.</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>Suriname</td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>Women</td>
<td>Suriname</td>
<td>Panel from Chapter 5.1 was included.</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>Suriname</td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>Women</td>
<td>The Netherlands</td>
<td>Panel from Chapter 4.1 was included.</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>China</td>
<td>Panel from Chapter 4.1 was included.</td>
</tr>
<tr>
<td></td>
<td>Heterosexual men</td>
<td>China</td>
<td></td>
</tr>
</tbody>
</table>


36. De Vries HJ, Smelov V,


69. Harris SR, Clarke IN, Seth-Smith