Journal of Clinical Microbiology

Chlamydia trachomatis Strains Show Specific Clustering for Men Who Have Sex with Men Compared to Heterosexual Populations in Sweden, the Netherlands, and the United States

Linus Christerson, Reinier J. M. Bom, Sylvia M. Bruisten, Resha Yass, Justin Hardick, Göran Bratt, Charlotte A. Gaydos, Servaas A. Morré and Björn Herrmann
J. Clin. Microbiol. **2012, 50(11):3548. DOI: 10.1128/JCM.01713-12.**
Published Ahead of Print 22 August 2012.

JCM
Journals.ASM.org

Chlamydia trachomatis Strains Show Specific Clustering for Men Who Have Sex with Men Compared to Heterosexual Populations in Sweden, the Netherlands, and the United States

Linus Christerson,[a] Reinier J. M. Bom,[b] Sylvia M. Bruisten,[b] Resha Yass,[a] Justin Hardick,[c] Göran Bratt,[d] Charlotte A. Gaydos,[c] Servaas A. Morré,[e,f] and Björn Herrmann[a]

Section of Clinical Bacteriology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden[a]; Public Health Laboratory, Public Health Service of Amsterdam, Amsterdam, the Netherlands[b]; Division of Infectious Diseases, Johns Hopkins University, Baltimore, Maryland, USA[c]; Venhälsan, Stockholm South General Hospital, Stockholm, Sweden[d]; Laboratory of Immunogenetics, Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, the Netherlands[e]; and Institute of Public Health Genomics, Department of Genetics and Cell Biology, Research Institutes CAPHRI and GROW, Faculty of Health, Medicine & Life Sciences, University of Maastricht, Maastricht, the Netherlands[f]

High-resolution genotyping of *Chlamydia trachomatis* improves the characterization of strains infecting different patient groups and sexual networks. In this study, multilocus sequence typing (MLST) and *ompA* sequence determination were used for an analysis of *C. trachomatis* strains from 203 men who have sex with men (MSM) from Sweden, the Netherlands, and the United States. The results obtained were compared with data from 153 heterosexual women from Sweden and the Netherlands. The overlap in MLST/*ompA* profiles between MSM from Sweden and the Netherlands was 68%, while the overlap between heterosexual populations from these countries was only 18%. The distribution of genotypes in MSM from the United States was less similar to that in MSM from the European countries, with 45% and 46% overlaps for MSM in Sweden and the Netherlands, respectively. Minimum-spanning-tree analysis of MLST/*ompA* sequence types identified two large clusters that contained almost exclusively samples from MSM and comprised 74% of all MSM samples. Three other clusters were predominated by samples from women but also contained MSM specimens. Of 19 detected variants of the MLST target CT144, three variants were highly associated with MSM. Our study supports the hypotheses of both tissue tropism as well as epidemiological network structures as explanations for the linkage between specific genetic variants and sexual orientation.

Chlamydia trachomatis is one of the most common sexually transmitted diseases worldwide. If left untreated, serious sequelae may arise, such as ectopic pregnancy and infertility in women and epididymitis and proctitis in men. High-risk groups for *Chlamydia* infection are characterized by a large number of partners and an infrequent use of preventive measures, such as condom use. Among men who have sex with men (MSM), the transmission of infection is often via anal intercourse, but this practice has increased among heterosexual populations in recent years. More experimental sex behavior is being noted, and several transmission routes may occur for sexually transmitted diseases.

Genotyping is important for an understanding of the epidemiology of *C. trachomatis*. Until recently, genotyping has been performed by serotyping of the major outer membrane protein or by analysis of the coding gene *ompA* (28). The latter is now referred to as genovar determination. Genovars E, D, and F are highly prevalent in heterosexual populations worldwide (23, 26, 31, 33), while genovars G, D, and J predominate in MSM populations (2, 20, 22, 29, 32). Since 2003, clonal outbreaks of invasive strains of *C. trachomatis* causing lymphogranuloma venereum (LGV) with severe proctitis have been described among MSM in Europe, often in combination with HIV and other sexually transmitted diseases. This led to an increase in the incidence of genovar L2b in the last decade (6, 30). Another clonal outbreak occurred mainly in Sweden and was due to a new variant of *C. trachomatis*, leading to the generation of false-negative results by some molecular detection systems. This variant is a genovar E strain and has not been reported to occur in MSM (13, 16, 19).

The occurrence of different genetic variants in specific risk groups might be due to behavioral reasons and limited exchange between hetero- and homosexual networks. An alternative explanation is tissue tropism, where different *C. trachomatis* genotypes propagate preferentially either in the urogenital or in the anorectal tract (3, 15).

Previously, we developed a high-resolution multilocus sequence typing (MLST) system that was specifically designed for short-term epidemiology and outbreak investigations (18). This is in contrast to two conventional MLST schemes that are based on housekeeping genes and therefore are better suited to monitor evolutionary changes and slow genetic processes (8, 27). Our MLST system is based on the PCR amplification and DNA sequencing of five highly variable but stable genetic regions (21). It has been applied to specimens from heterosexual populations in a number of studies and has shown up to a 5-fold-higher resolution than *ompA* genotyping, revealing a large diversity of clinical strains missed by *ompA* genovar typing (7, 11, 16, 18). In a modified version, high resolution was retained but sensitivity was im-

Received 29 June 2012 Returned for modification 24 July 2012
Accepted 13 August 2012
Published ahead of print 22 August 2012
Address correspondence to Björn Herrmann, bjorn.herrmann@medsci.uu.se.
L.S. and R.J.M.B. contributed equally to the study.

doi:10.1128/JCM.01713-12

proved by targeting smaller parts of the polymorphic genetic regions (4).

Larger studies investigating specimens from MSM using high-resolution genotyping have not been conducted. In this study, the objective was to better understand the role of sexual networks as an explanation for the different *C. trachomatis* genotypes in MSM and heterosexual populations and also to determine if tissue tropism could be an alternative explanation. MLST analysis was applied for MSM cases from three countries, and genetic variation data were compared to previously obtained MLST data for women.

(The results of this study were presented in part at the 7th Meeting of the European Society for Chlamydia Research, Amsterdam, Netherlands, 2012 [5].)

MATERIALS AND METHODS

Inclusion of patient samples. Non-LGV chlamydia-positive samples from MSM were selected from three different countries: Sweden, the Netherlands, and the United States. From Sweden, 114 anogenital samples (75 rectal swabs and 39 urine specimens) were collected in 2006 at the Venhälsan Gay Clinic at South General Hospital in Stockholm. From the Netherlands, 107 rectal swab samples were collected at the Public Health Service in Amsterdam in 2008 and 2009. The Dutch samples were previously described by Quint et al. (29). From the United States, 145 rectal swab samples from MSM were collected at Johns Hopkins University in Baltimore, MD, in 2010 and 2011, and samples from across the United States (Boston, New York, Washington, DC, Atlanta, San Francisco, and Los Angeles) were also collected.

Chlamydia-positive urogenital samples from 153 women from Sweden ($n = 60$) and the Netherlands ($n = 93$) were also included. From Sweden, 40 culture isolates were collected at the Örebro University Hospital in 2006. To minimize sampling bias, cases of the new variant were not included in the present study, since they represent a clonal outbreak (16). In addition, 20 swab specimens were sampled in Uppsala in 2007. From the Netherlands, 70 culture isolates were collected between 2001 and 2005 (7), and 23 urogenital samples were collected at the Public Health Service in Amsterdam in 2008 and 2009 (4). Full MLST data were available for these samples. The sexual orientation of the women was unknown, but it was assumed that the vast majority were heterosexual.

DNA extraction. DNAs from all samples from Sweden and the United States were extracted by using the QIAamp DNA minikit or the MagAttract DNA M48 minikit (Qiagen, Hilden, Germany) on a BioRobot M48 workstation (Qiagen), according to the manufacturer's instructions. The preparation of the Dutch DNA samples was done by isopropanol precipitation, as described previously (4, 29).

Amplification of MLST regions. With the exception of 44 Dutch specimens, the MLST target regions and *ompA* were amplified in Uppsala by using previously described primers (6, 12, 16, 18). These primers are summarized in Table 1. The PCR reactions were carried out by using HotStar *Taq* DNA polymerase (Qiagen). Cycling conditions were as follows: an initial denaturation step at 95°C for 15 min, followed by 40 cycles at 94°C for 45 s, 60°C for 45 s, and 72°C for 90 s. The amplification was ended by an elongation step at 72°C for 10 min, followed by cooling at 4°C. The PCR products were verified by gel electrophoresis. In cases where no PCR product was detected, a second PCR step was carried out under the same cycling conditions but in a seminested fashion where one of the primers was exchanged, resulting in shorter products (Table 1). For 44 of the Dutch samples, the MLST target regions and *ompA* were amplified in Amsterdam by a nested PCR, generating the original full fragment length.

DNA sequencing. In Uppsala, the PCR products were purified with exonuclease I (Fermentas, Burlington, Canada) and FastAP thermosensitive alkaline phosphatase (Fermentas), according to the manufacturer's instructions. Purified PCR products and sequencing primers (Table 1) were mixed and sent to Macrogen (Seoul, South Korea) for sequencing. In Amsterdam, the MLST regions were sequenced as described previously (4). The data were analyzed and assembled by using BioEdit 7.0.9 (Ibis Therapeutics, Carlsbad, CA) and DNA Baser v2.80.0 (Heracle-Software, Lilienthal, Germany). All novel mutations were reamplified and resequenced to ensure their authenticity.

Statistics, similarity, and cluster analyses. Differences between groups were tested univariately by using Pearson's χ^2 test and Fisher's exact test, when appropriate. A P value of <0.05 was considered statistically significant. Analyses were performed with the SPSS package, version 19.0 (SPSS Inc., Chicago, IL).

The obtained sequences were checked against the Uppsala *Chlamydia trachomatis* MLST database (http://mlstdb.bmc.uu.se/) and were given an allele number for each region. Samples resulting in a complete MLST profile were included in this study and were assigned a sequence type (ST). These STs were used to calculate the percent similarity (PS) as a measure of similarity between every group. For this, the following formula was used:

$$PS_{ij} = \sum_{k=1}^{S} \min(p_{ik}, p_{jk})$$

where S is the total number of STs in groups i and j combined, p_{ik} is the percentage of ST k in group i, and p_{jk} is the percentage of ST k in group j. The discriminatory power was determined for MLST with 6 target regions (MLST-6) and was described as Simpson's discriminatory index with a 95% confidence interval (CI).

Minimum-spanning trees were generated by an analysis of the full MLST profiles. Cluster analysis was performed on the minimum-spanning trees, allowing single-locus variance (SLV); i.e., strains that differ in only one target region cluster together. All cluster analyses were performed by using BioNumerics 6.6 (Applied Maths, Sint-Martens-Latem, Belgium) under the categorical coefficient of similarity and the priority rule of the highest number of SLVs.

Phylogenetic analyses. DNA sequences from all CT144 variants (674 bp or 683 bp) were analyzed by using a gamma-distributed ($\gamma = 0.35$) Tamura 3-parameter model from which a boot-strapped (2,000 replicates) neighbor-joining tree was drawn. The tree was rooted by using the murine *Chlamydia* CT144 homologue TC0421 (*Chlamydia muridarum* [GenBank accession number AE002160]) as an outgroup. All phylogenetic analyses were performed by using MEGA 5 software.

RESULTS

Study population and samples. In total, 366 anogenital samples from men who have sex with men (MSM) were included, of which 114 originated from Sweden, 107 originated from the Netherlands, and 145 originated from the United States. Of the included samples, 203 (55%) were successfully genotyped by MLST and *ompA* sequencing, and these samples were obtained in equal numbers for the three countries: 67 samples (46 rectal swabs and 21 urine specimens) from Sweden, 69 rectal samples from the Netherlands, and 67 rectal samples from the United States. For the groups of heterosexual women, 153 fully typed samples were included, of which 60 originated from Sweden and 93 originated from the Netherlands (4, 7, 16).

Distribution of *ompA* genovars. The distribution of identified genovars per country and per gender is shown in Fig. 1. Within the 203 MSM samples, we identified the following 8 *ompA* genovars: genovars B, D, E, F, G, I, J, and K. The most predominant genovars were genovars G (37%), D (27%), and J (18%), whereas genovars E and F were uncommon (7% and 5%, respectively). For genovars B, I, and K, only 12 cases were identified among MSM, and genovar H was lacking in all 3 countries. The genovar distributions for Sweden and the Netherlands were almost identical among the MSM samples (Table 2 and Fig. 1). Significant differences were noticed for the American distribution compared to the distributions for both Sweden ($P =$

TABLE 1 Primers for the MLST system and *ompA*[a]

Region	Primer	Function(s)	Sequence
hctB (step 1)	hctB39F	PCR, sequencing	5′-CTCGAAGACAATCCAGTAGCAT-3′
	hctB794R	PCR, sequencing	5′-CACCAGAAGCAGCTACACGT-3′
hctB (step 2)	CT046NR3	PCR, sequencing	5′-CCCCAAATATGCAACAGGAT-3′
	CT046NF	PCR, sequencing	5′-AACTCCAGCTTTTACTGCTA-3′
CT058 (step 1)	CT222F	PCR, sequencing	5′-CTTTTCTGAGGCTGAGTATGATTT-3′
	CT1678R	PCR, sequencing	5′-CCGATTCTTACTGGGAGGGT-3′
	CT811F	Sequencing	5′-CGATAAGACAGATGCCGTTTTT-3′
	CT1022R	Sequencing	5′-TAAGCACAGCAGGGAATGCA-3′
CT058 (step 2)	CT058NF	PCR, sequencing	5′-AGGTGGCTGCGTTAAGATAACT-3′
	CT1678R	PCR, sequencing	5′-CCGATTCTTACTGGGAGGGT-3′
CT144 (step 1)	CT144:248F	PCR, sequencing	5′-ATGATTAACGTGATTTGGTTTCCTT-3′
	CT144:1046R	PCR, sequencing	5′-GCGCACCAAAACATAGGTACT-3′
CT144 (step 2)	CT144:248F	PCR, sequencing	5′-ATGATTAACGTGATTTGGTTTCCTT-3′
	CT144NR	PCR, sequencing	5′-CCTAAACATACGGCTATTCC-3′
CT172	CT172:268F	PCR, sequencing	5′-CCGTAGTAATGGGTGAGGGA-3′
	CT172:610R	PCR, sequencing	5′-CGTCATTGCTTGCTCGGCTT-3′
pbpB1 (step 1)	pbpB1F	PCR, sequencing	5′-TATATGAAAAGAAAACGACGCACC-3′
	pbpB823R	PCR, sequencing	5′-CAGCATAGATCGCTTGCCTAT-3′
pbpB1 (step 2)	CT682NF	PCR, sequencing	5′-TCATCACTTTGCGTATATGGCA-3′
	pbpB823R	PCR, sequencing	5′-CAGCATAGATCGCTTGCCTAT-3′
pbpB2 (step 1)	pbpB1455F	PCR, sequencing	5′-GGTCTCGTTTTTGATGTTCTATTC-3′
	pbpB2366R	PCR, sequencing	5′-TGGTCAGAAAGATGCTGCACA-3′
pbpB2 (step 2)	pbpB1455F	PCR, sequencing	5′-GGTCTCGTTTTTGATGTTCTATTC-3′
	pbpB2333R	PCR, sequencing	5′-GCAGATACTAACTTAAAAATAGAC-3′
ompA (step 1)	118F	PCR, sequencing	5′-ATTGCTACAGGACATCTTGTC-3′
	1163R	PCR, sequencing	5′-CGGAATTGTGCATTTACGTGAG-3′
	ctr200F	Sequencing	5′-TTAGG5*GCTTCTTTCCAATAYGCTCAATC-3′[b]
	ctr254R	Sequencing	5′-GCCAYTCATGGTARTCAATAGAGGCATC-3′
ompA (step 2)	MOMP87	PCR, sequencing	5′-TGAACCAAGCCTTATGATCGACGGA-3′
	RVS1059	PCR, sequencing	5′-GCAATACCGCAAGATTTTCTAGATTTCATC-3′

[a] "Step 1" indicates primers that were used in the first PCR step. In cases where no PCR products were detected, the "step 2" primers were used in an inner PCR step, resulting in slightly shorter products.
[b] 5* indicates inosine.

0.003) and the Netherlands ($P = 0.026$), as genovars D and G were less prevalent and genovar J was more prevalent in the United States. Interestingly, all samples of genovars B and I were of American origin, while genovar K was not seen among the Swedish samples.

For the 153 samples from heterosexual women, 9 different genovars were found, belonging to genovars B and D to K (Fig. 1 and Table 2). The most predominant genovars were genovars D (16%), E (29%), F (13%), and G (12%). The genovar distributions among women between the 2 countries were also significantly different ($P = 0.040$). Notable were the high prevalence of genovar E samples from Sweden compared to the samples from the Netherlands (42% versus 20%) and the higher prevalence of genovar H and I samples in the Netherlands than in Sweden.

Overlap of MLST sequence types. To characterize the *C. trachomatis* sequence types (STs) in the different MSM and heterosexual groups, we analyzed the full MLST profiles, including an allele type for *ompA* (MLST-6) (Table 2). The overlaps of STs between the MSM populations were 45% between Sweden and the United States, 46% between the Netherlands and the United States, and 68% between Sweden and the Netherlands. Most specimens (65%) had a genotype that occurred in all three MSM populations. The American MSM had more unique STs (46%) than did Dutch (13%) and Swedish (30%) MSM.

Between the heterosexual populations in the Netherlands and Sweden, this overlap in STs was much lower (18%) than that between the MSM populations.

Cluster analysis of MSM populations. A minimum-spanning tree was drawn for all 203 samples collected from MSM. The tree showed 4 STs that were highly prevalent, with 14, 34, 39, and 39 samples per ST, comprising 62% of the total samples (Fig. 2a).

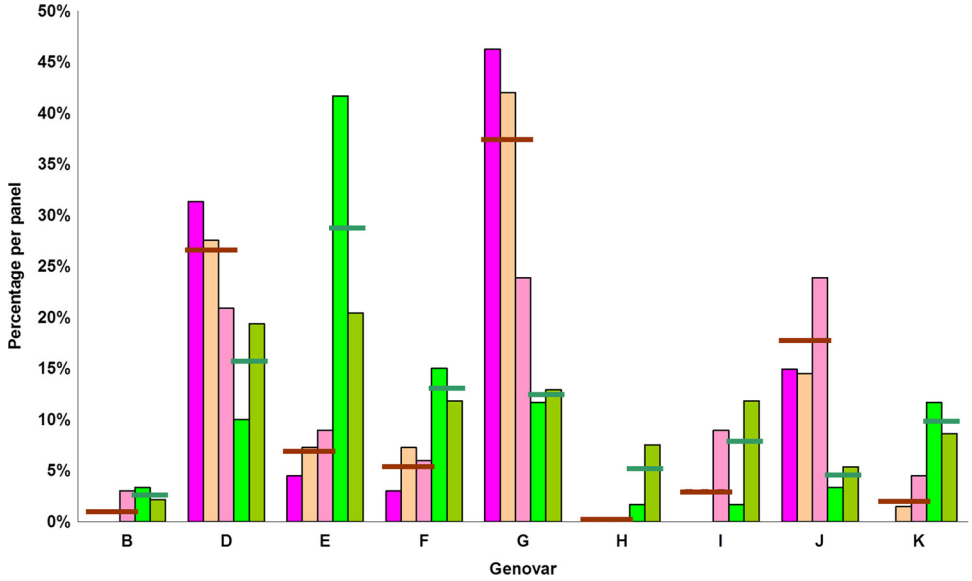

FIG 1 Genovar distribution given as percentages per group. The vertical bars give the distributions of the groups per country, with the following color coding: magenta, Swedish MSM ($n = 67$); salmon, Dutch MSM ($n = 69$); pink, American MSM ($n = 67$); green, Swedish heterosexual women ($n = 60$); light green, Dutch heterosexual women ($n = 93$). The horizontal lines give the average distributions per risk group, with the following color coding: red, MSM ($n = 203$); dark green, heterosexual women ($n = 153$).

When clustering with single-locus variance (SLV) was allowed, 3 clusters appeared, 2 of which were very large and contained 74% of all samples. Cluster I ($n = 43$) was centered on the large genovar D ST and consisted solely of genovar D samples. Cluster II ($n = 107$) contained mainly the other three highly abundant STs and presented both genovar G and J samples. Cluster III ($n = 10$) was much more heterogeneous and was comprised of genovar F samples. Although the range in calendar time between the collections of some samples was quite long (5 years), these distributions were very similar for the three different countries. Samples from the three countries contributed equally to all major STs and clusters, although the samples from the United States were more often singletons. This was reflected by the large number of STs found among this MSM group as well ($n = 33$) (Table 2).

TABLE 2 Numbers of samples in different groups and numbers of obtained genovars and sequence types[f]

Population	Collection yr	No. of samples	No. of *ompA* genovars[a]	No. of *ompA* genotypes[b]	No. of STs (MLST-5)[c]	No. of STs (MLST-6)[d]	MLST-6/*ompA* genotype ratio[e]	Simpson's diversity index (MLST-6) (95% CI)
MSM groups								
Sweden	2006	67	5	10	19	21		0.88 (0.83–0.92)
Netherlands	2008–2009	69	6	6	15	16		0.85 (0.81–0.89)
United States	2010	67	8	14	30	33		0.93 (0.90–0.97)
All MSM		203	8	17	53	57	3.4	
Heterosexual groups								
Sweden	2007	60	9	14	40	41		0.98 (0.96–1.00)
Netherlands	2001–2005 2008–2009	93	9	23	59	61		0.98 (0.98–0.99)
All heterosexual individuals		153	9	26	86	93	3.6	
Total		356	9	33	129	140	4.2	

[a] Defined by conventional genovar designations.
[b] Number of unique *ompA* sequence variants found.
[c] Includes five target regions in MLST (*ompA* excluded).
[d] Includes six target regions in MLST (*ompA* included).
[e] Ratio of MLST-6 STs and designated *ompA* genotypes.
[f] MSM, men who have sex with men; ST, sequence type; MLST, multilocus sequence typing.

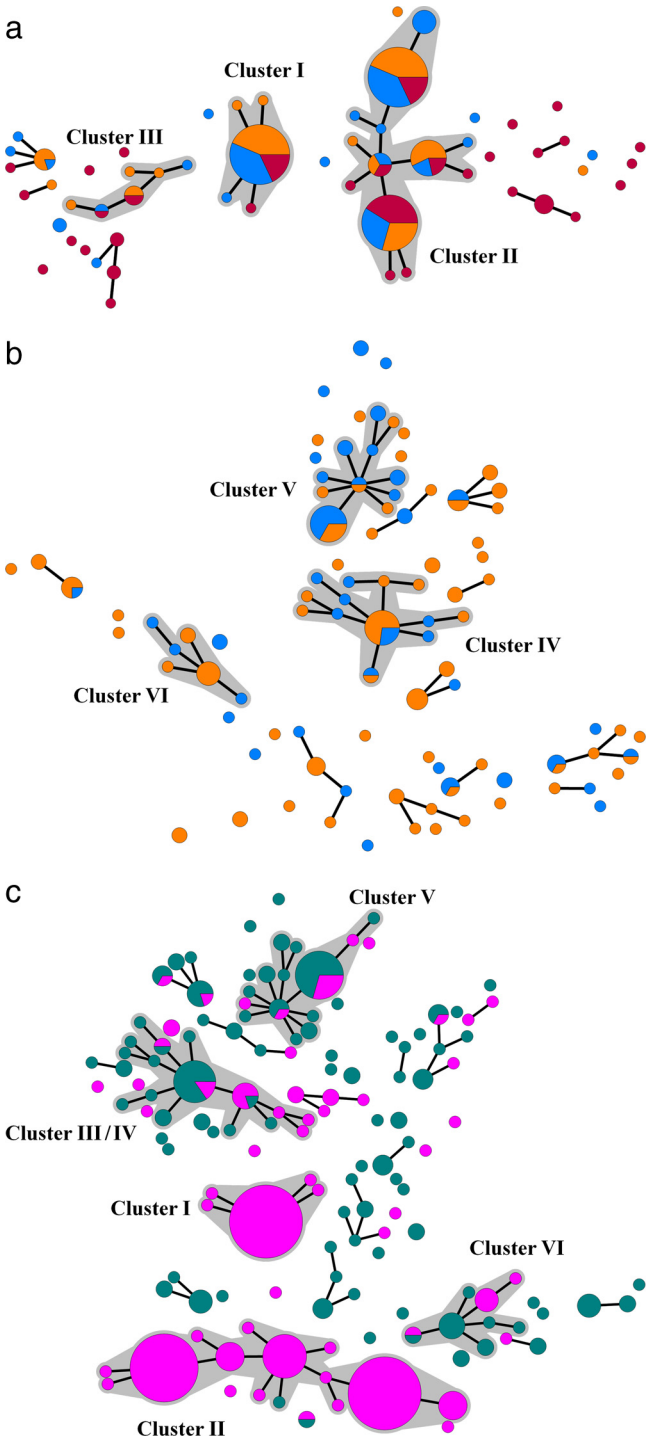

FIG 2 Minimum-spanning tree of 203 samples from MSM from Sweden, the Netherlands, and the United States (a); 153 samples from heterosexual women from Sweden and the Netherlands (b); and 356 samples from MSM and heterosexual women from Sweden, the Netherlands, and the United States (c). The sphere sizes indicate the numbers of samples in each sphere, branches show single-locus variants (SLV), and halos indicate clusters based on SLVs. The color coding for panels a and b is as follows: blue, Sweden; orange, Netherlands; red, United States. The color coding for panel c is as follows: pink, MSM; green, heterosexual women.

Cluster analysis of heterosexual populations. The distribution of the 153 samples from Swedish and Dutch heterosexual women was much more diverse (Fig. 2b). Two STs of moderate size ($n = 11$ and 12) were present, one of genovar E and the other of genovar F. The other samples contributed to 27 small STs ($n =$ 2 to 5) and 64 singletons, resulting in a total of 93 unique STs (Table 2). When clustering with SLV was allowed, 3 large clusters appeared, containing 40% of all samples. The clusters (clusters IV to VI) consisted of samples with genovar E ($n = 26$), genovars D and F ($n = 24$), and genovars I and J ($n = 11$). The samples from heterosexual populations were genetically much more diverse and clusters were more heterogeneous than those of MSM, where the two main clusters were dominated by just one or a few large STs. As a result of this diversity, the two *Chlamydia* populations from heterosexual populations in Sweden and the Netherlands were less overlapping.

Comparison of clusters of *C. trachomatis* strains of MSM and heterosexual populations. To investigate whether chlamydial clusters differ between MSM and heterosexual populations, we generated a minimum-spanning tree containing all 356 samples (Fig. 2c). Clusters were assigned to allow single-locus variance and sexual orientation to superimpose over the STs. In total, 5 large clusters appeared, containing 67% of all samples. When these clusters were analyzed, cluster III from the MSM group (Fig. 2a) proved to be the same as cluster IV from the heterosexual group (Fig. 2b). The other 4 clusters turned out to be clusters I and II from the MSM groups (Fig. 2a) and clusters V and VI, which were assigned to the heterosexual groups (Fig. 2b). About one-third of the samples from the MSM group that could not be assigned to a cluster now clustered with clusters V and VI from the heterosexual group.

When the distribution of the MSM and heterosexual samples over these clusters was analyzed, a clear separation was apparent. Two clusters were found almost exclusively among MSM and comprised 74% of all MSM specimens. The remaining MSM samples were distributed evenly over the other clusters that we identified in the minimum-spanning tree for heterosexual populations. The samples from heterosexual populations were almost all in heterogeneous clusters, and only one sample turned up in an MSM cluster (cluster II) (Fig. 2c).

Analysis of the CT144 region. CT144 is an open reading frame coding for a hypothetical 285-amino-acid protein. In our study, 19 different CT144 variants (allele types) were detected in the 356 specimens. Interestingly, three of these variants were strongly associated with MSM. Thus, clusters I and II (Fig. 2a) almost completely shared a single variant, CT144-5, while being discordant at all five other loci. A variant containing just one single-nucleotide polymorphism (SNP) (CT144-13) comprised the remaining six samples, and a third variant (CT144-31), again containing just one SNP, was found in a single MSM sample. These three CT144 variants were seen in only five samples outside the two MSM-associated clusters, belonging to samples from one woman and four MSM. These 3 variants were therefore highly linked to the MSM population and comprised 154 (76%) individuals of this group, while only 2 (1%) of the heterosexual individuals had any of these variants ($P < 0.001$). In a phylogenetic analysis performed on all 19 CT144 variants, these 3 MSM-linked variants were located in a significant monophyletic cluster (bootstrap support = 98), together with variant CT144-17, which was more distantly

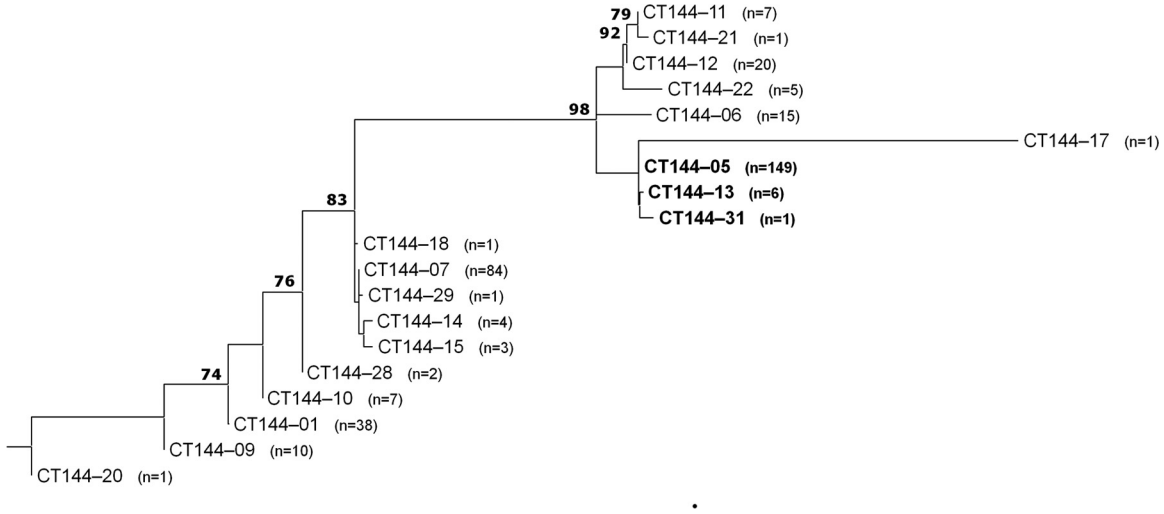

FIG 3 Consensus neighbor-joining tree of CT144 variants (674 bp or 683 bp). For all 19 variants, the numbers of samples per type are shown ($n = 1$ to 149). The 3 CT144 variants associated with MSM are given in boldface type. Bootstrap supports of ≥70% are given. The tree is rooted by using murine *Chlamydia* CT144 homologue TC0421 (*Chlamydia muridarum* [GenBank accession number AE002160]) as an outgroup.

located (Fig. 3). Variant CT144-17 was detected in one genovar E sample from a Swedish woman.

DISCUSSION

The development of a high-resolution genotyping method enables the improved characterization and understanding of the transmission of *C. trachomatis* infections. In this study, certain MLST profiles strongly predominated in MSM, while other profiles were found mainly in heterosexual women. These findings indicated that the transmission of *C. trachomatis* between groups of different sexual orientations is very limited. Some of our results, however, might support the hypothesis of tissue tropism, in which certain genetic variants of *C. trachomatis* would be associated with different types of epithelial cells. Both hypotheses are discussed here.

Our study certainly confirms previously reported findings that genovars D, G, and J are much more common among MSM than among heterosexual populations and also that genovar E, the predominating genovar in most heterosexual populations around the world, is only rarely found in MSM. The collection period for Swedish heterosexual populations included the year 2006, when the incidence of the new variant of *C. trachomatis* peaked (13, 16). The inclusion of this new variant would result in a proportion of 60% genovar E samples, but here, we found a proportion of 40% among Swedish heterosexual populations, which is a representative figure for Sweden in other years and also similar to data from reports from other countries (17, 24, 25, 31, 33). The small overlap in genovar distribution for heterosexual populations in Sweden and the Netherlands can be explained by the large diversity found for samples from heterosexual populations. However, there was also a selection bias for the Dutch group, which was selected to represent all different genovars in a study of associations with clinical symptoms (7).

For the characterization of *C. trachomatis* strains from MSM, the MLST system provided a 3- to 4-fold-higher discriminating capacity than conventional *ompA* sequencing (Table 2), and this is in accordance with the application of MLST to other study groups with different selection criteria (4, 7, 11, 14, 16, 18). The addition of *ompA* to the five MLST targets increased the resolution but only marginally (Table 2).

The use of MLST instead of genovars also resulted in the detection of genetic variants that were unique to specific study groups (countries), and here also, there was more variation for heterosexual women than for MSM. Comparisons between the three countries showed that between 45% and 68% of the MSM specimens had overlapping ST profiles. This percentage is considerably higher than what was noted for heterosexual women in Sweden and the Netherlands (18%) and even higher than what was reported in a Norwegian study of specimens from youths including geographic subpopulations that were collected in the same year (11). The large overlap between data for MSM from different countries probably reflects that this group has larger numbers of partners and more internationalized sexual networks than heterosexual populations (9).

The analysis of MLST variants identified two clusters (clusters I and II) (Fig. 2c) that were found almost exclusively in MSM and comprised 74% of all specimens from this group. The proportions of chlamydia samples derived from each country were equally distributed, although the difference in collection dates was up to 5 years. This indicates that specific MLST variants were associated with infections in MSM over time and space in a stable way.

Previous investigations of possible tissue tropism focused on associations with different *ompA* genovars, but other genes may be linked to infections of specific tissues. Since some MLST variants were found almost exclusively in the rectum of MSM, it may be argued that there is a tropism of certain chlamydia strains for epithelial cells in the rectum, while other strains are more adapted to the infection of genital or conjunctival epithelial cells. A recent genome sequencing study of *C. trachomatis* strains linked certain polymorphisms in open reading frames CT144, CT154, and CT326 with rectal tropism in genovar G isolates but did not find a linkage for genovar E and J strains (15). Thus, the reported linkage between CT144 and MSM may be restricted to genovar G by recombination linkage. Our analyses strengthen data from previous

reports and showed that 3 CT144 variants were found exclusively in *C. trachomatis* strains from MSM. The predominating allele 5 was not restricted to genovar G ($n = 63$) but was also frequently found in genovar D ($n = 49$) and J ($n = 37$) strains from MSM. Thus, the association between CT144 variants and MSM is independent of genovar G. Furthermore, these 3 variants clustered with a fourth variant, CT144-17 (Fig. 3), that was derived from a Swedish woman. In previous MLST studies, this specific CT144-17 type has been found almost exclusively in LGV strains from MSM (6) and only in 2 non-MSM specimens (genovar E, from a man and a woman in our study) out of 971 specimens in our database. On the other hand, the predominating MSM-linked variant, CT144-5, was previously found in 46 of 75 trachoma cases comprising genovars A and B in an unpublished study. This illustrates the complexity of genetic *Chlamydia* type linkage to tissue tropism.

Another explanation for the association of certain MLST profiles with sexual orientation may be the limited contacts between MSM and the heterosexual population. Such an interpretation may also be supported by data from previous studies based on *ompA*. Genovar E predominates among heterosexual populations and comprises about 40% of the strains in most populations (23, 26, 31, 33). Although fewer studies are available, genovar E is also the most common genovar in rectal samples from women, and there was no difference between rectal and urogenital specimens in women (1, 3, 10). In contrast, genovar E is uncommon in MSM and typically comprises 4% to 10% of the strains, and no difference was seen previously when rectal and urethral/urine specimens were compared (2, 20, 22, 29, 32). The striking difference in genovar E distributions between MSM and heterosexual populations could be explained by a low level of interrelationship between the two groups rather than tissue tropism. Data on differences in genovar distributions between specimens from the rectum and urine/urethra from the same individual are very limited, However, in a previous study comprising 120 rectal and 97 urine/urethral samples from 203 MSM in Sweden, we could not find any difference in genovar distributions (20).

There are limitations in our study, and the first is that the inclusion of samples was done mainly retrospectively, and there was a selection bias in the inclusion of specimens from heterosexual individuals. To minimize sampling bias, we excluded samples that were known to have LGV genovars or the new variant genovar E. These two types of samples were also typed previously with the presently used MLST method and found to be completely identical in all 6 regions per genovar and were thus designated clonal isolates (6, 13, 16, 19). Another limitation is that the categorization of sexual orientation for the included persons may have been erroneous. A reanalysis of Swedish MSM with MLST profiles typical of heterosexual individuals showed in two cases that the individuals were actually bisexual. It was also not known if all women were truly heterosexual. In addition, only 55% of the tested samples resulted in complete MLST profiles, possibly leading to a biased population. The relatively low scores of complete MLST profiles could be explained by a retrospective analysis of previously tested specimens and low genome copy numbers, together with the fact that large gene sequences were chosen to be included for allele typing instead of shorter sequences (4). The use of shorter allelic regions in Amsterdam was shown improve the success rate of MLST typing to 80% to 90%, and the gender-based distribution differences then remained the same (data not shown). Also, the genovar distribution in our study is similar to those reported in previous studies (1, 20, 22, 29, 32), and therefore, it may be argued that the loss of chlamydia-positive samples was rather random and not a cause for biased results.

A strength of our study is that we demonstrated that specific genetic variants were clearly linked to different sexual orientations, although the specimens included were collected retrospectively over several years and at different geographical sites, despite the selection bias in samples from heterosexual populations.

In summary, we have shown that specific genetic variants are highly associated with different sexual orientations, and this is even more convincing when using MLST than when using conventional *ompA* typing. The available data cannot yet determine if the hypothesis of tissue tropism or that of epidemiological network structures is the best explanation for the linkage between specific genetic variants and sexual orientation. New studies are therefore needed to shed light on this intriguing question.

ACKNOWLEDGMENTS

We thank Anna Ståhlsten, Diem Nguyen, Kristina Ödman, Elaine Åstrand, Jenny Isaksson, Sahare Gashi-Ilazi, and Nadia Nassir Hajipour for help with the laboratory work.

The study was financed by local funds from Uppsala University Hospital, Sweden, and the Public Health Service of Amsterdam, Netherlands; by ZonMW grant nr 125010008 (the Amsterdam MSM Network Study), Netherlands; and by NIH NIAID grant HPTN U01 AI 068613.

REFERENCES

1. **Barnes RC, Rompalo AM, Stamm WE.** 1987. Comparison of *Chlamydia trachomatis* serovars causing rectal and cervical infections. J. Infect. Dis. **156**:953–958.
2. **Bauwens JE, Lampe MF, Suchland RJ, Wong K, Stamm WE.** 1995. Infection with *Chlamydia trachomatis* lymphogranuloma venereum serovar L1 in homosexual men with proctitis: molecular analysis of an unusual case cluster. Clin. Infect. Dis. **20**:576–581.
3. **Bax CJ, et al.** 2011. Analyses of multiple-site and concurrent *Chlamydia trachomatis* serovar infections, and serovar tissue tropism for urogenital versus rectal specimens in male and female patients. Sex. Transm. Infect. **87**:503–507.
4. **Bom RJ, et al.** 2011. Evaluation of high-resolution typing methods for *Chlamydia trachomatis* in samples from heterosexual couples. J. Clin. Microbiol. **49**:2844–2853.
5. **Bom R, et al.** 2012. Genotyping of *Chlamydia trachomatis* strains from men who have sex with men and women in the Netherlands, Sweden and USA: evidence for tissue tropism or not?, p 34–35. *In* Ouburg S, Morre S (ed), Proceedings of the 7th meeting of the European Society for Chlamydia Research, Amsterdam, Netherlands. VU University, Amsterdam, the Netherlands.
6. **Christerson L, et al.** 2010. Typing of lymphogranuloma venereum *Chlamydia trachomatis* strains. Emerg. Infect. Dis. **16**:1777–1779.
7. **Christerson L, de Vries HJ, Klint M, Herrmann B, Morre SA.** 2011. Multilocus sequence typing of urogenital *Chlamydia trachomatis* from patients with different degrees of clinical symptoms. Sex. Transm. Dis. **38**:490–494.
8. **Dean D, et al.** 2009. Predicting phenotype and emerging strains among *Chlamydia trachomatis* infections. Emerg. Infect. Dis. **15**:1385–1394.
9. **Fenton KA, Imrie J.** 2005. Increasing rates of sexually transmitted diseases in homosexual men in Western Europe and the United States: why? Infect. Dis. Clin. North Am. **19**:311–331.
10. **Geisler WM, Morrison SG, Bachmann LH.** 2008. Absence of lymphogranuloma venereum strains among rectal *Chlamydia trachomatis* outer membrane protein A genotypes infecting women and men who have sex with men in Birmingham, Alabama. Sex. Transm. Dis. **35**:856–858.
11. **Gravningen K, et al.** 2012. Multilocus sequence typing of genital *Chlamydia trachomatis* in Norway reveals multiple new sequence types and a large genetic diversity. PLoS One **7**:e34452. doi:10.1371/journal.pone.0034452.
12. **Harding-Esch EM, et al.** 2010. Multi-locus sequence typing: a useful tool

for trachoma molecular epidemiology, p 55–58. *In* Stary A, et al (ed), Chlamydial infections. Proceedings of the Twelfth International Symposium on Human Chlamydial Infections, San Francisco, CA. International Chlamydia Symposium, San Francisco, CA.
13. **Herrmann B, et al.** 2008. Emergence and spread of *Chlamydia trachomatis* variant, Sweden. Emerg. Infect. Dis. **14**:1462–1465.
14. **Ikryannikova LN, Shkarupeta MM, Shitikov EA, Il'ina EN, Govorun VM.** 2010. Comparative evaluation of new typing schemes for urogenital *Chlamydia trachomatis* isolates. FEMS Immunol. Med. Microbiol. **59**: 188–196.
15. **Jeffrey BM, et al.** 2010. Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. Infect. Immun. **78**:2544–2553.
16. **Jurstrand M, et al.** 2010. Characterisation of *Chlamydia trachomatis* by *ompA* sequencing and multilocus sequence typing in a Swedish county before and after identification of the new variant. Sex. Transm. Infect. **86**:56–60.
17. **Jurstrand M, et al.** 2001. Characterization of *Chlamydia trachomatis omp1* genotypes among sexually transmitted disease patients in Sweden. J. Clin. Microbiol. **39**:3915–3919.
18. **Klint M, et al.** 2007. High-resolution genotyping of *Chlamydia trachomatis* strains by multilocus sequence analysis. J. Clin. Microbiol. **45**:1410–1414.
19. **Klint M, et al.** 2011. Prevalence trends in Sweden for the new variant of *Chlamydia trachomatis.* Clin. Microbiol. Infect. **17**:683–689.
20. **Klint M, et al.** 2006. Lymphogranuloma venereum prevalence in Sweden among men who have sex with men and characterization of *Chlamydia trachomatis ompA* genotypes. J. Clin. Microbiol. **44**:4066–4071.
21. **Labiran C, et al.** 2012. Genotyping markers used for multi locus VNTR analysis with *ompA* (MLVA-ompA) and multi sequence typing retain stability in *Chlamydia trachomatis.* Front. Cell. Infect. Microbiol. **2**:68. doi: 10.5923/j.microbiology.20120204.01.
22. **Li JH, et al.** 2011. Prevalence of anorectal *Chlamydia trachomatis* infection and its genotype distribution among men who have sex with men in Shenzhen, China. Jpn. J. Infect. Dis. **64**:143–146.
23. **Lysen M, et al.** 2004. Characterization of *ompA* genotypes by sequence analysis of DNA from all detected cases of *Chlamydia trachomatis* infections during 1 year of contact tracing in a Swedish county. J. Clin. Microbiol. **42**:1641–1647.
24. **Machado AC, et al.** 2010. Distribution of *Chlamydia trachomatis* genovars among youths and adults in Brazil. J. Med. Microbiol. **60**:472–476.
25. **Morre SA, et al.** 1998. Serotyping and genotyping of genital *Chlamydia trachomatis* isolates reveal variants of serovars Ba, G, and J as confirmed by *omp1* nucleotide sequence analysis. J. Clin. Microbiol. **36**:345–351.
26. **Morre SA, et al.** 2000. Urogenital *Chlamydia trachomatis* serovars in men and women with a symptomatic or asymptomatic infection: an association with clinical manifestations? J. Clin. Microbiol. **38**:2292–2296.
27. **Pannekoek Y, et al.** 2008. Multi locus sequence typing of Chlamydiales: clonal groupings within the obligate intracellular bacteria *Chlamydia trachomatis.* BMC Microbiol. **8**:42. doi:10.1186/1471-2180-8-42.
28. **Pedersen LN, Herrmann B, Moller JK.** 2009. Typing *Chlamydia trachomatis*: from egg yolk to nanotechnology. FEMS Immunol. Med. Microbiol. **55**:120–130.
29. **Quint KD, et al.** 2011. Anal infections with concomitant *Chlamydia trachomatis* genotypes among men who have sex with men in Amsterdam, the Netherlands. BMC Infect. Dis. **11**:63. doi:10.1186/1471-2334-11-63.
30. **Spaargaren J, Fennema HS, Morre SA, de Vries HJ, Coutinho RA.** 2005. New lymphogranuloma venereum *Chlamydia trachomatis* variant, Amsterdam. Emerg. Infect. Dis. **11**:1090–1092.
31. **Suchland RJ, Eckert LO, Hawes SE, Stamm WE.** 2003. Longitudinal assessment of infecting serovars of *Chlamydia trachomatis* in Seattle public health clinics: 1988-1996. Sex. Transm. Dis. **30**:357–361.
32. **Twin J, et al.** 2011. *Chlamydia trachomatis* genotypes among men who have sex with men in Australia. Sex. Transm. Dis. **38**:279–285.
33. **Yang B, et al.** 2010. The prevalence and distribution of *Chlamydia trachomatis* genotypes among sexually transmitted disease clinic patients in Guangzhou, China, 2005-2008. Jpn. J. Infect. Dis. **63**:342–345.