UNIVERSITY OF SOUTHAMPTON

Faculty of Medicine, Division of Clinical and Experimental Science.

Molecular epidemiology of *Chlamydia trachomatis*: Evaluation, implementation and development of high resolution genotyping.

by

Clare Labiran

Thesis for the degree of Doctor of Philosophy

June 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE
DIVISION OF CLINICAL AND EXPERIMENTAL SCIENCE
DEPARTMENT OF MOLECULAR MICROBIOLOGY AND INFECTION

Doctor of Philosophy

MOLECULAR EPIDEMIOLOGY OF *CHLAMYDIA TRACHOMATIS*: DEVELOPMENT AND EVALUATION OF HIGH RESOLUTION GENOTYPING

By Clare Labiran

Chlamydia is the most frequently diagnosed bacterial STI worldwide. It is often an asymptomatic infection therefore individuals may not seek medical attention and continue to pass the bacteria on to various sexual partners. As the bacteria is passed on from person to person it continues to evolve and changes in the genome sequence continue to occur. The ability to sequence the bacteria in order to track its evolution and changes that may occur is highly important. However, whole genome sequencing takes a long time and has not yet been fully developed to obtain whole genome sequence data from clinical samples. This is a huge problem when dealing with bacteria such as chlamydia. The problem lies in the fact that chlamydia are obligate intracellular bacteria, only growing within cells and hence it is difficult to isolate chlamydia in cell culture to propagate the bacteria for whole genome sequencing purposes as whole genome sequencing currently requires large concentrations of DNA.
To get around the issues associated with whole genome sequencing, typing systems have been developed to sequence small sections of the genome and assign types to each sample based on sequence variation. The current typing systems include typing of the *omp*A gene (Types A-K, L1-L3), MLST typing and MLVA typing for *C. trachomatis*.

The purpose of this study was to evaluate current high resolution typing methods for *C. trachomatis* and to use these typing methods to study the epidemiology of *C. trachomatis* in Brighton and Southampton. To achieve the goals for this thesis, typing systems that offer the most discrimination were assessed for their suitability and the stability of the markers in the typing systems (Chp3) were determined. The epidemiology of chlamydia amongst HIV+ and HIV- men who have sex with men (MSM) in Brighton (Chp4) and a mainly heterosexual population in Southampton (Chp 5) were studied. MLVA-*omp*A typing of the samples from the two cohorts was also accompanied by demographic data from individuals such as their age, gender and postcodes to determine if there were specific MLVA-*omp*A types associated with each demographic factor. This data enhanced our understanding of chlamydia in these two cities. Additionally, chlamydia was isolated in cell culture from 37 rectal samples which will now be whole genome sequenced to compare whole genome data from rectal samples to pre-existing genome sequences from genital tract chlamydia.

Further analysis of the current high resolution typing methods for *C. trachomatis* has also shown that new typing systems are required that mimic the resolution offered by whole genome sequencing and with the availability of vast amounts of whole genome sequencing data new typing schemes can be developed for *C. trachomatis.* Additionally, there are regions left uncovered by the choice and location of the current markers used in these typing schemes which could be possible locations for new markers.

TABLE OF CONTENTS

ABSTRACT..... **I**

Table of Contents..... **III**

Figures..... **XI**

Tables..... **XIII**

Publications and Presentations..... **XV**

Declaration Of Authorship..... **XVII**

Acknowledgements..... **XIX**

Abbreviations. ..... **XXI**

CHAPTER 1 Introduction...... **1**

1.1 Historical perspective..... 1

1.2 Classification and phylogeny/taxonomy..... 1

1.3 Pathogenicity of Chlamydia spp. in humans..... 4

1.3.1 *Chlamydia pneumoniae.* 4

1.3.2 *Chlamydia psittaci.* 4

1.3.3 *Chlamydia trachomatis.* 4

1.4 Pathogenesis of *Chlamydia trachomatis* sexually transmitted infection in men and women 8

1.5 Chlamydial morphology 8

1.6 Chlamydial developmental cycle 10

1.7 Host Responses to *C. trachomatis.* 13

1.8 Persistent sexually transmitted chlamydial infections. 16

1.9 Antibiotic treatment for *C. trachomatis.* 16

1.10 Vaccine development for *C. trachomatis.* 16

1.11 Risk factors..... 17

1.11.1 Age and gender..... 17

1.11.2 Race and socioeconomic status (SES)..... 18

1.11.3 Sexual behaviour..... 18

1.11.4 Marital status..... 18

1.11.5 Transmission 19

1.12 Neonatal and infant infections. 19

1.13 Epidemiology of *C. trachomatis.* 19

1.13.1 STI incidence in England..... 21

1.13.2 Epidemiology of *C. trachomatis* in Europe 24
1.13.3 Epidemiology of *C. trachomatis* in the USA. 25
1.14 Chlamydia trachomatis diagnosis 27
1.14.1 Isolation of *C. trachomatis* in cell culture. 27
1.14.2 Enzyme Immuno-assay (EIA) Tests. 27
1.14.3 Direct Fluorescent Antibody (DFA) Tests. 27
1.14.4 Nucleic Acid Hybridisation (Nucleic Acid Probe) Tests (NAH tests) 28
1.14.5 Serological tests 28
1.14.6 Nucleic Acid amplification tests (NAATs) 28
1.14.7 *C. trachomatis* point of care tests. 29
1.15 Non molecular typing methods for *C. trachomatis* 30
1.16 General overview of genotyping systems in bacteriology. 31
1.16.1 Pulsed-field gel electrophoresis (PFGE). 31
1.16.2 Restriction fragment length polymorphism (RFLP). 31
1.16.3 Random amplification of polymorphic DNA (RAPD). 32
1.16.4 Amplified fragment length polymorphism (AFLP) 32
1.16.5 *omp*A PCR and sequencing. 32
1.16.6 Multilocus sequence typing (MLST). 32
1.16.7 Variable number tandem repeat (VNTR) 33
1.17 Aims of project 35
CHAPTER 2 Material and Methods 37
2.1 Chemicals and solutions. 37
2.1.1 Water. 37
2.1.2 Sterilisation. 37
2.1.3 Buffers, reagents and solutions. 37
2.1.4 Antibodies. 39
2.1.5 Plastic and Glassware. 39
2.2 Bacterial growth medium and solutions. 39
2.2.1 Media. 39
2.2.2 Antibiotics. 39
2.3 Chlamydia host cell line. 39
2.3.1 Cell culture. 39
2.4 Detection of mycoplasma by PCR. 40
2.5 Chlamydia species. 41
2.6 Infection of cell monolayer with *C. trachomatis.* 41

2.7 Harvesting of *C. trachomatis* isolates from wells and flasks. 41
2.8 Titration. 41
2.9 Fixing cells. 43
2.10 X-Gal staining. 43
2.11 Isolation of nucleic acids. 43
2.11.1 DNA extraction using Promega wizard genomic DNA purification kit. 43
2.11.2 DNA extraction using Nucleospin tissue kit. 44
2.11.3 DNA extraction using QIAamp Viral RNA. 44
2.12 Analysis of nucleic acids. 45
2.12.1 TAE Buffer (tris-acetate EDTA buffer). 45
2.12.2 Agarose gel electrophoresis. 45
2.12.3 Visualisation of DNA bands. 46
2.12.4 Estimation of DNA size and concentration. 46
2.12.5 Amplification of DNA by polymerase chain reaction (PCR). 48
2.12.6 PCR conditions. 48
2.12.7 Determinating the concentration of nucleic acids using the NanoDrop 1000™ spectrophotometer. 50
2.12.8 Purification of PCR products using Wizard SV gel and PCR clean-up system (Promega). 50
2.12.9 DNA Sequencing. 50
2.13 Whole genome sequencing. 51
2.14 Phylogentic trees. 54
2.15 Data analysis. 54
2.15.1 DNA sequence analysis: Nucleotide BLAST search tool used to search for similar genomes that have been deposited online. 54
2.15.2 DNA sequence analysis: Alignment and annotation using Lasergene. 54
2.15.3 Whole genome analysis using Artemis. 54
2.15.4 Comparison of whole genomes using ACT. 55
2.16 Statistics. 55
2.16.1 Simpsons index of diversity. 55
2.17 Bionumerics. 55
2.18 Ethics. 55
2.18.1 Ethics Research Governance Online (Ergo). 55
2.18.2 NRES. 56
2.18.3 NHS/HSC RESEARCH AND DEVELOPMENT OFFICES- (R AND D APPROVAL). 56
2.18.4 NIHR CRN PORTFOLIO APPLICATION. 56

CHAPTER 3 *Chlamydia trachomatis* marker stability study. 57

3.1 Introduction. 57

3.2 Aims: 58

3.3 Methods. 59

3.3.1 Experimental design. 59

3.3.2 Samples 60

3.3.3 Stability of the MLVA-*omp*A markers through adaptation of *C. trachomatis* to cell culture 63

3.3.4 Stability of the MLVA-*omp*A and the MST markers after long term passage in cell culture. 65

3.4 Results 69

3.4.1 Determination of the stability of the MLVA-*omp*A markers after eight passages in cell culture. 69

3.4.2 Determination of the stability of the MLVA-*omp*A markers in the Swedish new variant following long term passaging. 69

3.4.3 Determination of the stability of the MST markers in the Swedish new variant following long term passaging. 69

3.4.4 Whole genome sequencing of two isolates at Passage 36 and passage 72 of the Swedish new variant 74

3.5 Discussion 77

3.6 Future work 81

CHAPTER 4 Diversity of genital tract strains in HIV positive and HIV negative men who have sex with men (MSM) in Brighton. 83

4.1 Introduction. 83

4.2 Aims of project. 85

4.2.1 Primary aim 85

4.2.2 Secondary aims. 86

4.3 Study design. 86

4.4 Swabs chosen for the isolation of *Chlamydia trachomatis* in cell culture and for genotyping. 87

4.4.1 BD ProbeTec™ swabs and transport media chosen for the storage of *C. trachomatis* samples before DNA extraction and typing. 88

4.4.2 Virus, chlamydia and mycoplasma (VCM) swabs and transport media chosen for the isolation of *C. trachomatis*. 88

4.5 Determination of the optimal storage conditions for recovery and isolation of *C. trachomatis* from the VCM transport medium. 88

4.5.1 4°C and -20°C stability study. 88

4.5.2 -70°C swab stability study to determine the viability of *C. trachomatis* in VCM transport media. 91
4.6 Typing of samples from MSM in Brighton. 95
4.6.1 Methods used to assign genotypes to the samples. 95
4.6.2 Genotyping results. 98
4.7 Simpsons index of diversity. 102
4.8 Typeability of markers. 102
4.9 Minimum spanning tree. 102
4.10 Comparison of *omp*A and MLVA-*omp*A genotypes with HIV status, symptomology, co-infection, previous *C. trachomatis* and other STI infections. 105
4.10.1 HIV status. 105
4.10.2 Symptoms and signs. 107
4.10.3 Previous CT and previous STIs. 107
4.10.4 Co-infection. 107
4.11 Participants with both urethral and rectal infections. 107
4.12 Isolation of *Chlamydia trachomatis* from samples collected from MSM in Brighton. 110
4.12.1 Methods used for isolation of *C . trachomatis* from the VCM swab and transport media system. 110
4.12.2 Results of the isolation of *C. trachomatis* from positive rectal samples collected from MSM in Brighton. 111
4.13 Discussion. 116
4.14 Future work. 122
4.15 Conclusions. 122
CHAPTER 5 Epidemiology of chlamydia in Southampton. 125
5.1 Introduction 125
5.2 Aims: 126
5.2.1 Primary aim 126
5.2.2 Secondary aims 126
5.3 Study hypothesis. 127
5.4 Study objectives and study design. 127
5.5 Collection and storage of specimens. 131
5.6 Sexual networks. 131
5.7 Study population and typing of samples. 135
5.8 *Omp*A distribution. 136
5.8.1 *omp*A genotype B3/IU-FQ279. 136

5.8.2 *omp*A genotype Ia.....136
5.8.3 *omp*A genotype L2b.....136
5.9 Genotyping using the MLVA-*omp*A typing scheme.....138
5.10 Simpsons index of diversity for the MLVA-*omp*A typing system with this sample set. 140
5.11 Typeability of the four MLVA-*omp*A markers.....140
5.12 Minimum spanning tree. 140
5.12.1 Generation of the minimum spanning tree.....140
5.12.2 Location of *omp*A genotypes on the Minimum spanning tree.....141
5.12.3 Partioning of the minimum spanning tree to create groups for which demographic data was analysed. 141
5.13 Analysis of patient demographic information.....147
5.13.1 Methods used to analyse demographic information. 147
5.13.2 Age. 147
.....148
5.13.3 Gender. 149
5.13.4 Ethnicity. 149
5.13.5 Occupation.....149
5.13.6 Residential postcodes of participants: are there MLVA-*omp*A types of chlamydia which circulate in certain parts of Southampton? 151
5.13.7 Comparison of demographic data between those who were positive and those who were negative. 157
5.14 Symptoms presented by participants.....158
5.15 Determining if there was any significant difference between the month the specimen was collected and the MLVA-*omp*A type. 160
5.16 Patients whose partners also had *C. trachomatis* positive samples. 166
5.17 Participants who attended the GUM clinic more than once with repeated infections during the study period. 166
5.18 Co-infection with other STIs. 166
5.19 Comparison of genotypes between samples collected in Southampton in 2009 and samples collected in Southampton in 2012/2013.....167
5.20 Discussion 171
5.21 Future work. 180
5.22 Conclusions.....181
CHAPTER 6 Analysis of current high resolution typing methods and the development of a new typing scheme for *C. trachomatis.* 183
6.1 Introduction.....183

6.2 Aims: 184
6.3 Rationale for study 184
6.4 Study design. 185
6.5 Construction of a phylogenetic tree from whole genome sequencing data of 62 isolates from Southampton. 187
6.5.1 Methods used to construct the phylogenetic tree. 187
6.5.2 Analysing the phylogenetic tree 188
6.5.3 Comparison of the whole genome sequence data to the *omp*A types. 189
6.5.4 Comparison of whole genome sequencing data to the markers in the Pedersen MLVA-*omp*A typing system. 190
6.5.5 Comparison of whole genome sequencing data to the markers in the Klint MST typing system. 190
6.6 Analysis of genome data to investigate whether it was possible to design a new typing system for *C. trachomatis*. 191
6.6.1 Identifying potential SNPs which may be useful in a SNP typing scheme for *C. trachomatis.* 191
6.6.2 Identifying potential regions within samples with the same MLVA-*omp*A type which could be used in a new typing scheme for *C. trachomatis*. 195
6.7 Samples with the same MLVA but different *omp*A types 200
6.8 Discussion 200
6.9 Future work. 202
CHAPTER 7 Final discussion 205
APPENDIX 1: Marker stability paper 213
APPENDIX 2: Reviewer commentary on marker stability paper 221
APPENDIX 3: Information sheet for Brighton study (Chp 4). 223
APPENDIX 4 : Consent form for Brighton study (Chp 4). 225
APPENDIX 5 : BD ProbeTec data sheet. 227
APPENDIX 6 : VCM validation sheet for chlamydia 229
APPENDIX 7 : Information sheet for Southampton study (chp 5). 231
APPENDIX 8 : Consent form for Southampton study (Chp 5) 233
APPENDIX 9 : Qiamp information sheet 235
APPENDIX 10: Minimum spanning tree of samples collected in Brighton (red) and samples collected in Southampton (green) 237
References 239

FIGURES

Figure 1.1 Classification of *Chlamydiales* based on whole genome sequencing proposed by Stephens *et al* in 2009 (Stephens *et al.*, 2009). 3
Figure 1.2 Electron micrograph of a thin section of part of the contents of a mature *C. trachomatis* inclusion, 40 hours post infection of HeLa 229 cells. 9
Figure 1.3 The *C. trachomatis* developmental cycle. (Wilson *et al.*, 2003). 12
Figure 1.4 The cases of chlamydia per 100, 000 population in selected EU member countries between 1995 and 2009 (ECDC, 2011) 26
Figure 2.1 Process of titration in 96 well tray. 42
Figure 2.2 Ladders used to estimate size of bands. Hyperladder 1 (left), ... 47
Figure 2.3 Stages of the whole genome sequencing process. Starting with library preparation to analysis of Illumina reads. 52
Figure 3.1 Generalisation of the passaging process from patient swab to T75 tissue culture flasks**.** 64
Figure 3.2 A 1.5% (w/v) agarose gel showing the PCR products after amplification of the MLVA-*omp*A markers, CT1335, CT1299, CT1291 and *omp*A in the *C. trachomatis* Sweden new variant. 67
Figure 3.3 Flow diagram depicting the passage process of Sweden 2 and how the sample was processed to determine the stability of *omp*A and the VNTR markers. 68
Figure 3.4 Chromatograms showing the VNTR region CT1299 for Southampton isolate G4 70
Figure 3.5 Alignment of marker CT1335. 71
Figure 3.6 Alignment of marker CT1299. 72
Figure 3.7 Alignment of marker CT1291. 73
Figure 3.8 Sequence of the hypothetical protein SW2_8621 in Sweden 2. 76
Figure 4.1. Experimental plan to analyse the effect of storing samples for different periods of time at 4°C and -20°C. 90
Figure 4.2 Survival of *C. trachomatis* over a 28 day period. 94
Figure 4.3 The distribution of each *omp*A genotype from MSM in Brighton (Blue bar) and the number of samples isolated in cell culture (Red bar) for each *omp*A genotype. 99
Figure 4.4 Graph showing the distribution of MLVA types according to *omp*A genotype. 101
Figure 4.5 Minimum spanning tree using the MLVA data from samples from MSM. 104
Figure 4.6 Minimum spanning tree showing samples from individuals who are HIV+ and those who are HIV-. 106
Figure 4.7 Miminimum spanning tree showing culturable and non culturable isolates. 114
Figure 4.8 Phase contrast microscope image of cells infected with *C. trachomatis* in a 115
Figure 5.1 Flow diagram of process to gain ethical approval before a study requiring ethical approval can begin. 133
Figure 5.2 Flow diagram of the process by which samples were collected and subsequently analysed. 134
Figure 5.3 Distribution of *omp*A genotypes. 137
Figure 5.4 Minimum spanning tree generated using MLVA-*omp*A typing data of samples collected in Southampton. 143
Figure 5.5 Minimum spanning tree showing the location of the *omp*A genotypes within the clonal complex. 144

Figure 5.6 Minimum spanning tree of 302 samples from the Royal South Hants GUM clinic, Southampton (a).....145
Figure 5.7 Age range of participants. 148
Figure 5.8 Distribution of the ethnicities of people in Southampton according to National Statistics figures (a) and a pie chart of the distribution of participants in this study (b). 150
Figure 5.9 Regions where participants represented in this study reside. 153
Figure 5.10 The Southampton Ward Boundaries 155
Figure 5.11 The Index of multiple deprivation for Southampton in 2010.....156
Figure 5.12 Minimum spanning tree showing the months the samples were collected. 162
Figure 5.13 Minimum spanning tree generated using MLVA-*omp*A typing data of samples collected in Southampton in 2009 and in 2013.....170
Figure 6.1 A graphical representation of the circular genome of *C. trachomatis* (reference strain D/UW-3/CX).....186
Figure 6.2 Comparison of MLVA-*omp*A genotyping to whole genome sequencing (recombination removed). 193
Figure 6.3 Comparison of MST and *omp*A genotyping to whole genome sequencing (recombination removed). 194
Figure 6.4 Phylogenetic tree depicting the differences dictated by the number of SNPS between samples with the same MLVA-*omp*A types.....199

TABLES

Table 1.1 Number of people diagnosed with chlamydia by gender and age group at all sources (GUM clinics, NCSP, non-NCSP and GUM clinic returns from laboratories) of report between 2008 -2011 in England (HPA, 2012a). 7
Table 2.1 Buffers, reagents and solutions used for all projects and the final 38
Table 2.2 Concentration of agarose needed to visualise bands in a 1 x TAE electrophoresis 46
Table 2.3 Primers used for PCR of MLVA-*omp*A markers. 49
Table 2.4 Primers used for sequencing of MLVA-*omp*A markers. 53
Table 3.1 Isolates included in part one of the stability study. 62
Table 4.1 Variant numbers assigned to the three VNTR loci (CT1335, CT1299, CT1291)..... 97
Table 4.2 Distribution of different MLVA types according to *omp*A genotype (total of each type). 100
Table 4.3 Typing data from participants who gave a swab and a urine sample..... 109
Table 4.4 MLVA-*omp*A types of the 37 isolates from rectal swabs. 113
Table 5.1 Prevalence estimates and expected precision for different 130
Table 5.2 Showing the type of samples collected from two female and three male patients who gave multiple samples at one visit to the clinic. 135
Table 5.3 New types and variants of the variable number tandem repeats (VNTRs) found in this study..... 138
Table 5.4 Distribution of different VNTR types according to *omp*A genotype (total of each type). 139
Table 5.5 Participant *omp*A genotypes and demographics, stratified by three different segment analyses..... 146
Table 5.6 Shown are the postcodes, region in the UK and numbers of males and females who reside in each region who participated in the study and were positive for *C. trachomatis.* The districts and the wards for each postcode were sourced by typing the four digit post code provided by each participant into the website www.postcodearea.co.uk which provided information on every city in the UK. 154
Table 5.7 Symptoms found in male and female patients and the number of patients with those symptoms. Also shown are the P values for each segment analysis. 159
Table 5.8 Demographic data from four individuals who had the same MLVA-*omp*A type of chlamydia and who were all tested in November. 163
Table 5.9 Demographic data of 21 individuals who had the same MLVA-*omp*A type (8.5.1-E). 164
Table 5.10 Demographic data from individuals with chlamydia with single locus variants of a MLVA-*omp*A type all collected in October..... 165
Table 5.11 Comparison of the *omp*A genotypes recorded from 2009 and 2013. 169
Table 6.1 Southampton *C. trachomatis* isolates from women (Wang *et al.*, 2011b) who had more than one sample with the same MLVA-*omp*A type. 198

PUBLICATIONS AND PRESENTATIONS

Some of the work presented in this thesis has been published:

LABIRAN, C., CLARKE, I. N., CUTCLIFFE, L. T., WANG, Y., SKILTON, R. J., PERSSON, K., BJARTLING, C., HERRMANN, B., CHRISTERSON, L. & MARSH, P. 2012a. Genotyping markers used for multi locus VNTR analysis with ompA (MLVA-ompA) and multi sequence typing (MST) retain stability in Chlamydia trachomatis. *Front Cell Infect Microbiol,* 2**,** 68.

DECLARATION OF AUTHORSHIP

I, CLARE LABIRAN,

declare that this thesis entitled **Molecular epidemiology of *Chlamydia trachomatis*: Evaluation, implementation and development of high resolution genotyping** and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published as: (Labiran *et al.*, 2012a)

Signed: …….

Date: …….

ACKNOWLEDGEMENTS

There are so many people who have supported me through the journey of completing my PhD and made this thesis possible and to them I would like to say thank you.

Firstly, I would like to acknowledge the Health Protection Agency (HPA) for the studentship that provided the funding for this PhD.

I would like to say a very big thank you to my supervisors, Professor Ian Clarke, Dr. Peter Marsh and Dr. Stuart Clarke who right from the beginning saw the potential in me, gave me this wonderful opportunity and have helped, guided and encouraged me to complete this PhD to the best of my ability. Many thanks are also due to Mrs Lesley Cutcliffe and Miss Rachel Skilton who were wonderful research technicians and were always there to teach me new techniques in the laboratory. I will also like to thank many staff who work for the HPA who have continually been a support system during my PhD, namely Dr Steve Green and Dr Simon Friar have been of help on many occasions.

I have collaborated with many people during my PhD and I would like to say thank you to all my collaborators in Southampton, Brighton, Sweden, Amsterdam and Cambridge who I have worked alongside to complete various projects in my PhD and who have taught me various techniques and have given me advice on different aspects in this project.

To my family and friends who have been there to support me through this journey, I would not have been able to do this without you. Mum and dad, you have always believed in me and taught me to follow my dreams. Thank you for being the best. To my brothers and sisters thanks for all your love and support over the years. I love you all.

To my Husband, Erese, thank you for being the support system I needed during my PhD and for the love and patience you've shown over the years. I would not have been able to do this without you.

And finally to God, thank you for making my dreams a reality.

ABBREVIATIONS

°C	Degrees centrigrade
3'	3 Prime
4SP	4 Sucrose Phosphate
5'	5 Prime
ACT	Artemis Comparison tool
ATP	Adenosine triphosphate
bp	Base pair
BLAST	Basic local alignment search tool
cm	Centimetre
CO2	Carbon dioxide
CRN	Clinical Research Network
dH2O	Deionised water
DFA	Direct Fluorescent antibody
DMEM	Dulbecco Modified Eagle medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyriobnucleic acid
EB	Elementary body
EDTA	Ethylenediaminetetra-acetic acid
ERGO	Ethics Research Governance Online
FCS	Foetal calf serum
g	gram
GTP	Guanosine triphosphate
GUM	Genitourinary medicine
H_2O	Water
HIV	Human Immunodefieciency Virus
ifu	inclusion forming units
IgG	Immunoglobulin G
IMD	Index of Multiple Deprivation
IMS	Industrial methylated spirit
INDEL	Insertion or deletion of bases
kb (s)	Kilobase(s)
kDa	Kilodaltons
L	Litre
LPS	Lipopolysaccharide
mg	milligram
$MgCl_2$	Magnesium chloride
ml	millilitre
MLVA	Multi locus variant
mm	millimeter
MLST	Multi locus sequence typing
MOI	Multiplicity of Infection
MOMP	Major outer membrane protein

MSM	Men who have sex with men
MSpT	Minimum spanning tree
MST	Multi sequence typing
NAATs	Nucleic acid amplification test
NAH	Nucleic Acid Hybridisation
NCSP	National chlamydia screening programme
NHS	National Health Service
NIHR	National Insititute for Health Research
nm	nanometers
NRES	National Research Ethics Service
ompA	Outer membrane protein A
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PID	Pelvic inflammatory disease
psi	pounds per square inch
RB	Reticulate body
RNA	ribonucleic acid
RNase	ribonuclease
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
STBRL	Sexually Transmitted Bacteria Reference Laboratory
STI	Sexually Transmitted Infection
TAE	Tris-acetate-EDTA
TBS	Tris buffered Saline
TEMED	N,N,N',N' tetramethylethylenediamine
Tm	melting temperature
Tris	Tris (hydroxymethyl) aminomethane
UHQ	Ulta high quality (water)
UV	Ultraviolet
VNTR	Variable number tandem repeat
WTSI	Welcome Trust Sanger Institute
x g	x gravity
X-GAL	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
µg	Microgram
µl	microlitre
µm	micrometre
µM	micromolar

CHAPTER 1 INTRODUCTION.

1.1 HISTORICAL PERSPECTIVE.

Chlamydiae are Gram negative bacteria which were discovered in 1907 by Halberstadter and Prowazek who were the first to observe the inclusions of the trachoma agent in infected ocular material and classified it as a protozoan. They also demonstrated how the disease is spread from person to person and it was then named *Chlamydozoa* meaning mantle bodies (Halberstaedter.L and von Prowazek, 1907). In 1935 Chlamydia was thought to be a virus as the infectious agent grew in living cells. The *Chlamydiae* were subdivided on the basis of their host and tissue tropisms and were found to be the causal agent in the 1929 psittacosis outbreak (Potter *et al.*, 1979). It was later discovered that *Chlamydiae* are in fact bacteria because they contained ribosomes, cell walls and both DNA and RNA (Moulder, 1966).

1.2 CLASSIFICATION AND PHYLOGENY/TAXONOMY.

Over the years there has been controversy surrounding the classification of *Chlamydia*. In the 1980s when the Approved Lists of Bacterial Names were published, the *Chlamydiaceae* had only two species; *Chlamydia trachomatis* and *Chlamydia psittaci* (Skerman V.B.D *et al.*, 1980). *C. trachomatis* strains were identified due to the accumulation of glycogen in their inclusions and their sensitivity to sulfadiazine. On the other hand, *C. psittaci* were usually resistant to sulfadiazine and did not accumulate glycogen in their inclusions. However, between the early 1970s and the early 1980s isolations were made that did not fit into these two species (Dwyer *et al.*, 1972, Darougar *et al.*, 1980, Forsey and Darougar, 1984). New techniques based on DNA-DNA reassociation studies in the 1980s revealed eight groups at the level of genus or species (Schleifer and Stackebrandt, 1983, Wayne *et al.*, 1987, Amann *et al.*, 1995). In addition, two new species were recognised known as *Chlamydia pneumoniae* (Grayston *et al.*, 1989) and *Chlamydia pecorum* (Fukushi and Hirai, 1992). In 1993, *C. trachomatis*-like *Chlamydia* from swine were identified by DNA sequence analysis and formed the ninth group. Many strains have been designated as *C. trachomatis*-like *Chlamydia* on the basis that they are resistant to sulfadiazine (Andersen and Rogers, 1998).

In 1999 Everett *et al* (1999) analysed the full-length 16S and 23S rRNA sequences from a selection of chlamydiae and suggested a reclassification based on the sequence similarity clusters of the 16S and 23S rRNA sequences. They suggested an update of the grouping where the family *Chlamydiaceae* would now be divided into two genera: *Chlamydia* and *Chlamydophila* and all members of the family would have 16S rRNA gene sequence that is >90% identical. The genus *Chlamydia* would contain the three species *C. trachomatis, Chlamydia muridarum* and *Chlamydia suis*, whilst the genus C*hlamydophila* would contain six species: *Chlamydia caviae, Chlamydia felis,*

Chlamydia pecorum, Chlamydia pneumoniae, Chlamydia abortus and *Chlamydia psittaci.* The novel chlamydia-like agents have >80% 16S rRNA gene sequence identity and are in various different families (i.e *Waddliaceae, Simkaniaceae* and *Parachlamydiaceae*) in the order *Chlamydiales* as they are obligate intracellular bacteria that have a developmental cycle which is similar to that of *Chlamydia* (Everett *et al.*, 1999). However, Everett *et al*'s proposal was and still is not accepted by all chlamydiologists (Everett *et al.*, 1999). In fact a letter was written to the editor from scientists at the University of California (Schachter *et al.*, 2001) two years post publication of the paper (Everett *et al.*, 1999), with an opening statement stating 'We, the undersigned, strongly object to the proposed reclassification of the order *Chlamydiales'* (Schachter *et al.*, 2001). The authors based their argument on the fact that Everett *et al* (1999) define that the 16S rRNA sequence should be >95% identical to be included in the same genus, yet *C. trachomatis, C. psittaci* and *C. pneumonia* only have a maximum of 5 to 6% difference between sequences so it is impossible to justify these species being in different genera.

There are several reasons why it is believed that *Chlamydia* and *Chlamydophila* should be in the same genus and these are because (1) the species in both genera cluster together, (2) the majority of the species have >97% 16S rRNA gene sequence similarity, (3) they share phenotypic characteristics and (4) each strain is exposed to common selective pressure due to the identical environments in which they live. It is therefore believed that the key is to analyse the whole genomes in order to observe the various differences and similarities between organisms and that groupings should be based on these findings (Stephens *et al.*, 2009).

Whole genome sequence analyses also reveal a nearly 80% conservation of genes and gene order between *C. trachomatis, C. abortus and C. pneumoniae*. The two new proposed genera also share biological attributes such as possessing a similar developmental cycle, which takes place within host cells and they all have the ability for horizontal gene transfer. On the basis of this evidence Stephens *et al* (Stephens *et al.*, 2009) proposed an end to the confusion in suggesting that scientists should adopt the one-genus approach where chlamydia is the only genus in the *Chlamydiaceae* family and has multiple species; *C. trachomatis, C. psittaci, C. pnuemoniae, C. pecorum, C. abortus*, *C. caviae. C. muridarum, C. suis and C. felis* (Figure 1.1).

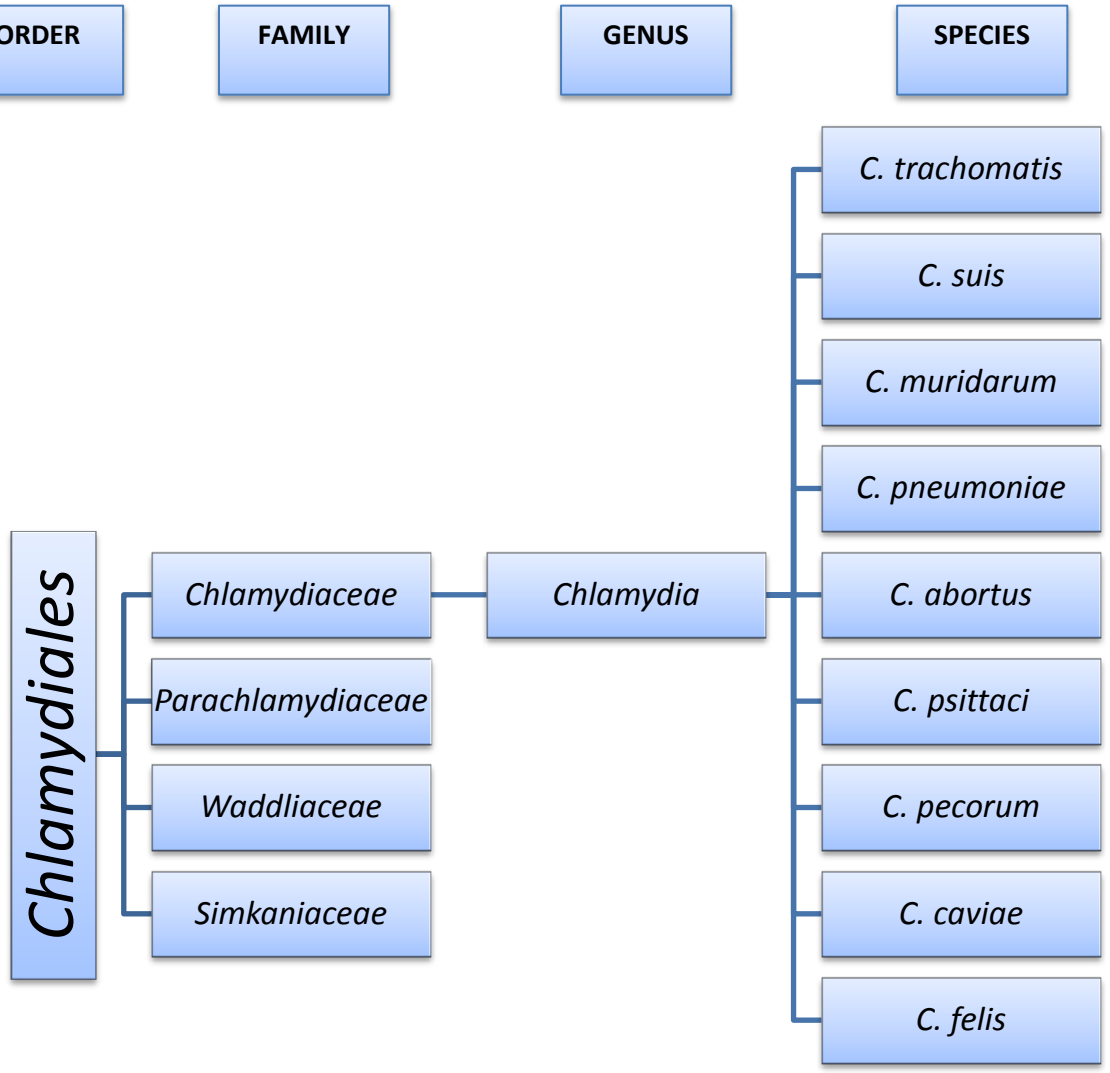


Figure 1.1 Classification of *Chlamydiales* based on whole genome sequencing proposed by Stephens *et al* in 2009 (Stephens *et al.*, 2009).

1.3 PATHOGENICITY OF CHLAMYDIA SPP. IN HUMANS.

1.3.1 *Chlamydia pneumoniae.*

Chlamydia pneumoniae is a pathogen of humans (Grayston *et al.*, 1989), which causes pneumonia and bronchitis, leading to chronic disease presentation resulting in reactive airways disease (Emre *et al.*, 1995), adult-onset asthma (Hahn *et al.*, 1996) and lung cancer (Laurila *et al.*, 1997). Previous studies have also suggested that *C. pneumoniae* could be associated with atherosclerosis (Kuo *et al.*, 1993a, Kuo *et al.*, 1993b, Kuo *et al.*, 1995, Mlot, 1996, Muhlestein *et al.*, 1996, Fong *et al.*, 1997). However, major clinical trials with antibiotics have not found a causal link between *C. pneumoniae* and atherosclerosis as antibiotics use in patients did not show any benefits with short and long term use (Gupta *et al.*, 1997, Muhlestein *et al.*, 2000, O'Connor *et al.*, 2003). *C. pneumonie* also causes pathogenicity in marsupials, frogs and horses (Mitchell *et al.*, 2010).

1.3.2 *Chlamydia psittaci.*

Chlamydia psittaci has been isolated from avian and mammalian hosts including humans. There have also been reports of *C. psittaci* in frogs (Newcomer *et al.*, 1982), fish (Wolke *et al.*, 1970) and tortoises (Vanrompay *et al.*, 1994). In humans *C. psittaci* causes respiratory psittacosis which is a disease contracted from bird such as parrots, pigeons and hens with disease presentation resembling that of a pneumonia infection (Vanrompay *et al.*, 1995).

1.3.3 *Chlamydia trachomatis.*

Chlamydia trachomatis is a human pathogen which is classified into a minimum of 15 serotypes, which are characterised by biological and clinical properties. These serotypes can be grouped into two biovars, the trachoma biovar comprising of strains which infect columnar epithelial tissue and the lymphogranuloma biovar, which is made up of strains infecting lymphatic tissue (van der Pol and Jones, 1992). The trachoma biovar includes those pathogens associated with causing trachoma (Serotypes A, B, Ba and C) and also pathogens associated with causing urogenital infections (serotypes D to K). The second biovar includes the serotypes that are associated with causing lymphogranuloma venereum [LGV (Serotypes L1, L2 and L3)] (van der Pol and Jones, 1992).

The classification of *C. trachomatis* is based on the sequence variation of the major outer membrane protein (MOMP) which is coded for by the outer membrane protein (*omp*) gene. The *omp*A gene encodes highly conserved protein structures that contain four evenly spaced variable domains (VDI, VDII, VDIII, VDIV), encoding the surface exposed epitopes (van der Pol and Jones, 1992). MOMP is the main target of host immune response and the various differences in the VDs

is believed to be due to immune selection (Wang *et al.*, 1973, Stephens *et al.*, 1982, Clarke, 2002). These differences in the VDs define the serotypes of *C. trachomatis*.

1.3.3.1 *C. trachomatis* ocular infection.

Trachoma is a disease that occurs in the eyes and is characterised by a roughening of the inner surface of the eyelids which eventually can lead to blindness. Trachoma is an ancient disease mentioned in the Ebers Papyrus, an Egyptian medical papyrus dating back to 1550 BC (Thygeson, 1962). The disease was spread to Europe by French and British troops. Before trachoma was eradicated in the United States any immigrant entering the country through Ellis Island in New York was examined to determine if they had trachoma and if it was discovered that they did then the individual was deported. Trachoma has now been eliminated from Europe and the United States through advances in hygiene and access to clean water (WHO, 2012); however, it is the leading cause of preventable blindness in the developing world. It is seen in poorer parts of Africa, the Middle East and Asia (WHO, 2012). The disease begins in children under the age of five and those that care for them and develops to blindness in older age, spreading via direct contact with eyes, nose and throat secretions from infected individuals. Flies, fomites and fingers can also be a source of transmission.

1.3.3.2 *C. trachomatis* genital tract infections.

Genital tract infections caused by *C. trachomatis* occur mainly in young people under the age of 24. The National Chlamydia Screening Program (NCSP) has been in place in the UK since 2003 to monitor and reduce the number of young people diagnosed with *C. trachomatis* by giving them access to free and confidential testing services and making them aware of chlamydia and its effects. However, despite the increasing efforts to reduce the number of young people diagnosed with *C. trachomatis*, this number continues to rise yearly. Table 1.1 shows figures from the 2012 report by the HPA of males and females that have been diagnosed with chlamydia between 2008 and 2011. Rates for 2011 have been calculated using Office Of National Statistics (ONS) population estimates for 2010 and are not the final figures (Final figures for 2011 will be published in 2013 when all the data is analysed by the HPA) (HPA, 2012a). The figures show that between 2008 and 2010 there was a 6.5% increase in the number of people diagnosed with chlamydia in England, showing that chlamydia may still be an increasing problem in England or be indicative of more people being tested. (As of the time this thesis was written, data for 2012 was available but was not comparable to data from previous years due to the introduction of Chlamydia Testing Activity Dataset (CTAD). CTAD was implemented in April 2012 and is a 'universal disaggregate dataset for the collection of data on all NHS and LA/NHS-commissioned chlamydia testing carried out in England' (HPA, 2013). CTAD replaces the NCSP core data return and the non-NCSP non-GUM aggregate data return).

1.3.3.3 *C. trachomatis* Lymphogranuloma venereum (LGV) infections.

Lymphogranuloma venereum (LGV) is caused by invasive serotypes of *C. trachomatis* – L1, L2 and L3. LGV affects the lymphatic system and the lymph nodes through breaks in the skin. It can also lead to proctitis, discharge and bleeding from the rectum (Spaargaren *et al.*, 2005). LGV was rarely seen in developed countries until 2003 when an outbreak was seen in Rotterdam, Netherlands (Nieuwenhuis *et al.*, 2004). Sequencing of the *omp*A gene revealed that the outbreak strain was a variant of the serotype, L2 , known as L2b. L2b is mainly seen in men who have sex with men (MSM) and has emerged as a significant problem among this group in Europe (Nieuwenhuis *et al.*, 2004, Herida *et al.*, 2005, Meyer *et al.*, 2005) and North America (Kropp and Wong, 2005). Many individuals infected with LGV strains of *C. trachomatis* have also been shown to be co-infected with HIV (Ward *et al.*, 2007). Since the outbreak there has been a total of 2397 LGV cases in the UK, with 395 cases recorded in 2012 compared to two cases in 2003 (Morre *et al.*, 2000).

	2008			2009			2010			2011		
Age group	**Male**	**Female**	**Total**	**Male**	**Female**	**Total**	**Male**	**Female**	**Total**	**Male**	**Female**	**Total**
<15	32	324	356	25	234	259	20	252	272	22	253	275
15-19	16,339	47,876	65,295	18,157	51,830	70,975	18,974	51,765	70,946	17,482	47,951	65,598
20-24	30,037	41,621	72,698	33,137	46,815	81,071	35,048	48,088	83,392	34,082	47,737	81,996
25-34	17,040	10,083	27,123	15,770	9,395	25,203	15,936	9,277	25,242	17,219	10,467	27,693
35-44	4,790	1,879	6,669	4,603	1,758	6,376	4,640	1,838	6,487	5,276	1,952	7,228
45-64	1,864	541	2,405	2,018	588	2,608	2,168	637	2,807	2,483	724	3,210
65+	82	13	95	84	10	94	113	18	131	128	16	144
Unknown	877	2,483	3,360	12	5	17	24	13	37	40	12	52
Total	**71,061**	**104,820**	**176,941**	**74,938**	**113,306**	**189,356**	**76,923**	**111,888**	**189,314**	**76,732**	**109,112**	**186,196**

Table 1.1 Number of people diagnosed with chlamydia by gender and age group at all sources (GUM clinics, NCSP, non-NCSP and GUM clinic returns from laboratories) of report between 2008 -2011 in England (HPA, 2012a).

1.4 PATHOGENESIS OF *CHLAMYDIA TRACHOMATIS* SEXUALLY TRANSMITTED INFECTION IN MEN AND WOMEN

C. trachomatis is effectively treated with antibiotics (azithromycin or doxycycline), however, in approximately 70% of women and 50% of men the infection is asymptomatic and therefore many individuals do not seek treatment for this condition and this can lead to further complications (Wilson *et al.*, 2002). Approximately 40% of women who have asymptomatic *C. trachomatis* then go on to develop Pelvic Inflammatory Disease (PID) and salpingitis as the infection spreads to the uterus or the fallopian tubes (Wilson *et al.*, 2002, Kalwij *et al.*, 2010). PID can cause infertility in approximately 20% of women and ectopic pregnancy in up 47% of women who have asymptomatic chlamydia. Females who are pregnant and become infected with chlamydia can also develop postpartum endometritis which is an infection in the uterus following delivery of the baby (Ismail *et al.*, 1987). In the remaining 30% of women that do develop symptoms these are in the form of abdominal pain, burning sensations during urination, painful intercourse, vaginal discharge and bleeding after intercourse (Kalwij *et al.*, 2010)

Symptoms in men include discharge from the penis, pain and/or burning when passing urine, a possible increase in the frequency of urination, burning and itching around the opening of the penis. Untreated chlamydia in males can lead to urethritis and epididymitis (Kalwij *et al.*, 2010).

1.5 CHLAMYDIAL MORPHOLOGY

There are two forms of *C. trachomatis*; the extracellular infectious form, known as the elementary body (EB) and the intracellular replicating and metabolically active form, known as the reticulate body (RB) (Figure 1.2).

The EB is a small metabolically inert, electron dense, spherical body that is approximately 0.2-0.3µm in diameter. EBs are also rigid, an adaptation which allows these structures to survive extracellularly (Matsumoto, 1973, Eb *et al.*, 1976). The cell wall of the EB is made of a granular outer layer and an inner layer that is composed of hexagonal array structures (Manire, 1966) but lacks detectable peptidoglycan. The major component found in these structures is *omc*B (a surface exposed glycosaminoglycan dependent adhesin) which contributes to cellular stability in EBs (Abdelrahman and Belland, 2005). In addition, cysteine rich proteins found in the outer envelope which include *omp*A, *omc*A have inter and intramolecular cysteine bonds which also add to the structural rigidity and stability of EBs (Abdelrahman and Belland, 2005). RBs on the other hand are approximately 1.0µm and are adapted for intracellular replication, therefore, RBs are the more fragile forms (Matsumoto, 1982). The cell wall has an inner and outer membrane and the cell envelope is pleomorphic due to the absence of any forms of structural stabilisation (Matsumoto, 1982). RBs have a granular appearing cytoplasm by EM due to ribosome synthesis taking place within the cytoplasm (Ward, 1983).

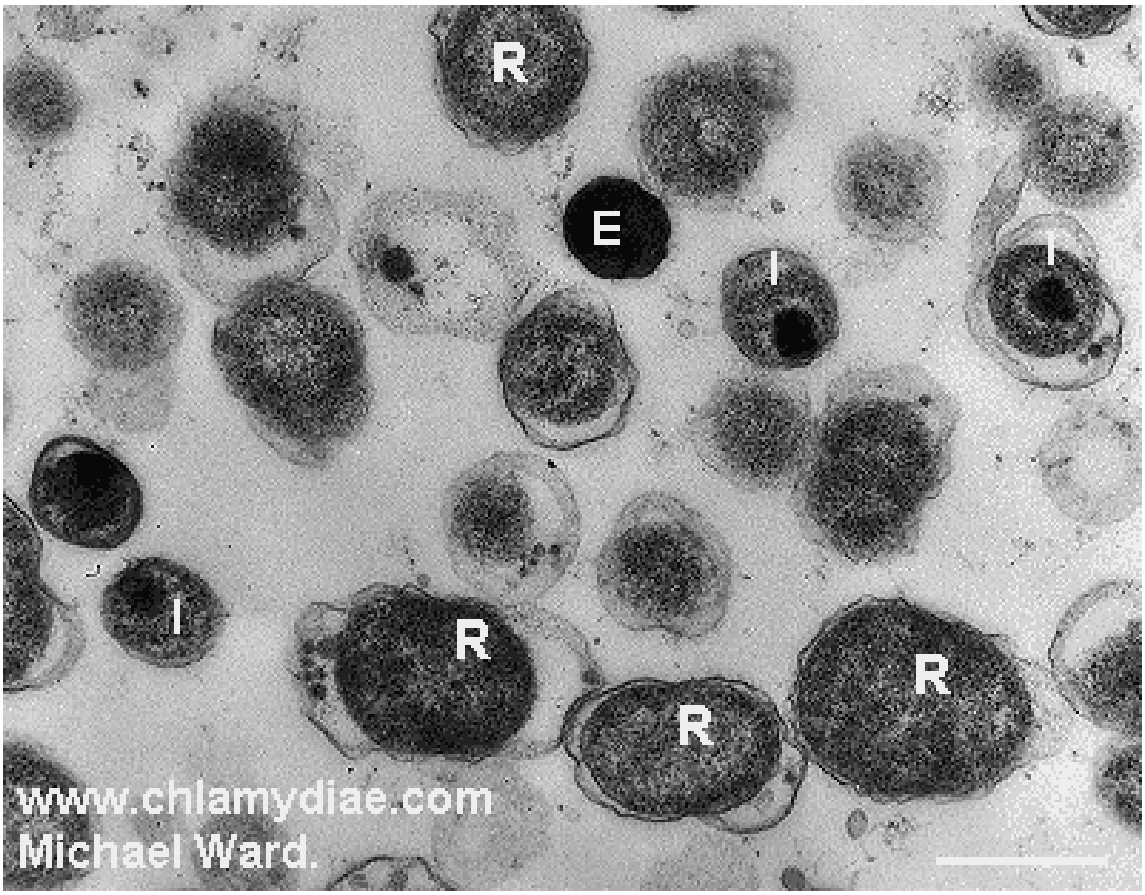


Figure 1.2 Electron micrograph of a thin section of part of the contents of a mature *C. trachomatis* inclusion, 40 hours post infection of HeLa 229 cells.
The image shows the reticulate bodies (R), the smaller intermediate bodies (I), and the elementary bodies (E). The bar represents 1 micron (Ward, 2004). Image taken by Professor M.E. Ward and included in this thesis with permission from Professor Ward.

1.6 CHLAMYDIAL DEVELOPMENTAL CYCLE

Chlamydiales are obligate intracellular bacteria and their developmental cycle takes place in eukaryotic cells (Figure 1.3). The developmental cycle takes place in several stages; the first stage is the attachment of the EB onto the host cytoplasmic membrane, followed by the entry of the bacteria into the cell where there is interaction of the bacteria with the host intracellular environment. Replication occurs within the host cell over 40-72 hours and ends when infectious EBs exit the cell. The EBs leave the cell ready to infect other cells and the life cycle starts all over again (Skilton *et al.*, 2009).

Initially, the EBs adhere to the host cell via projections on the host cell surface, there are several potential candidates that could make this possible and these include *omp*A, *omc*B, Hsp70 (Hackstadt, 1999) and *Pmp* (Everett and Hatch, 1995, Grimwood and Stephens, 1999, Grimwood *et al.*, 2001). There have also been suggestions that attachment of the EB to the host cell may be mediated by both electrostatic interaction between heparan sulphate containing glycosaminoglycans via the major outer membrane protein (through VS2 and VS4) and host-parasite interactions (Su *et al.*, 1996, Su and Caldwell, 1998, Stephens *et al.*, 2001, Taraktchoglou *et al.*, 2001).

After attachment, the EB is endocytosed through interactions between the host cell and *chlamydiae* either via a microfilament-dependent phagocytic mode of entry (Ward and Murray, 1984) or receptor mediated endocytosis (Soderlund and Kihlstrom, 1983). Within 6-9 hours after uptake the EB is re-organised into a larger structure (the RB) within an inclusion (an organelle in which the *chlamydiae* grow and divide). *Chlamydiae* evade phagolysosomal fusion and degranulation within the host cell to avoid host cell defences by a poorly understood process (Eissenberg and Wyrick, 1981).

Reorganisation also involves the loss of rigidity in the chlamydia cell membrane. In the inclusion the RB replicates by binary fission every 2 to 3 hours. The mature inclusion can contain up to 200-500 chlamydial particles (Wilson *et al.*, 2003) and therefore may occupy over three-quarters of the volume of the cell (Ward, 1983). Approximately 20 hours into the developmental cycle, some RBs undergo a reorganisation and condense back into the infectious form (the EB) which involves a reduction in size and the formation of rigid walls. Different species of intracellular bacteria leave the host cell in different ways. It has been demonstrated that some *chlamydiae* exit the host cell by lysis whilst others do not lyse the host cell at the end of replication, instead they are released into the environment by fusing with the plasma membrane and releasing the *chlamydiae* into the environment (Todd and Caldwell, 1985, Belland *et al.*, 2001).

Overall the chlamydial developmental cycle takes between 40-72 hours during which there is a rapid multiplication of RBs (8-10 divisions) with a constant increase in size of the inclusion body which at the end of the cycle contains RBs, EBs, intermediate forms and glycogen-like granules (Abdelrahman and Belland, 2005).

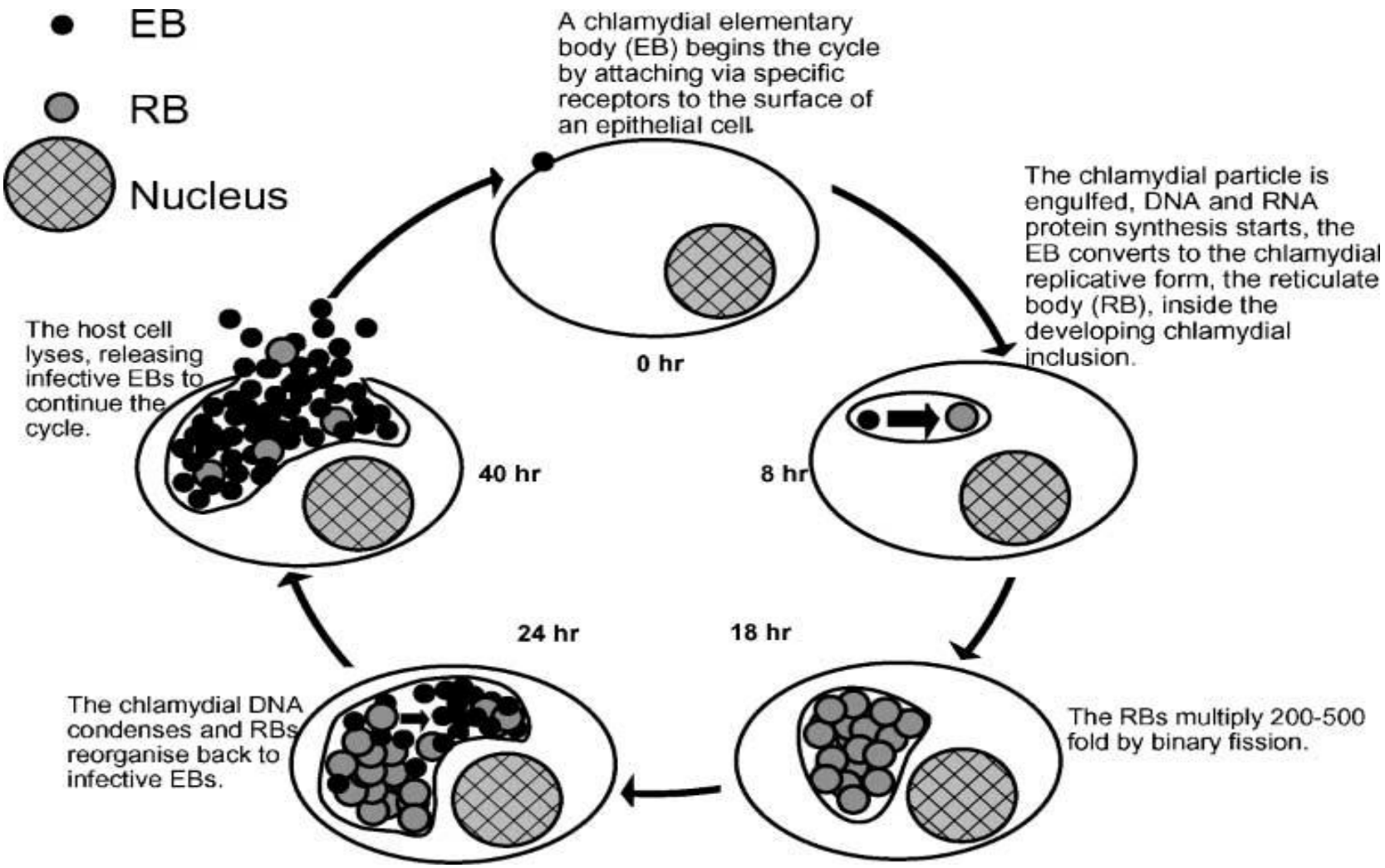


Figure 1.3 The *C. trachomatis* developmental cycle. (Wilson *et al.*, 2003).
Image included in this thesis with permission from Dr Wilson.

1.7 HOST RESPONSES TO *C. TRACHOMATIS.*

C. trachomatis primarily infects epithelial cells and has evolved mechanisms to survive intracellularly. Host cells have also devised mechanisms to regulate infection (Roan and Starnbach, 2008). The host cell response following *C. trachomatis* infection is established by innate immune cells, B cells and T cells.

The first line of defence following infection with *C. trachomatis* is provided by the mucosal barrier of the genital tract. Shedding of the endometrial epithelium can affect the ability of the bacterium to infect the cell (Tuffrey *et al.*, 1986, Ramsey *et al.*, 1999). Also present within the urethra of individuals who have been infected are defensins. Defensins have previously been shown to inhibit *C. trachomatis* infection *in vitro* (Porter *et al.*, 2005).

Toll-like receptors (TLRs) recognise pathogen-associated molecular patterns (PAMPs) and this in turn increases the production of Interferon-γ (IFN-γ) and other pro-inflammatory cytokines in response to *C. trachomatis* infection. TLR2 is the main candidate for signalling pro-inflammatory cytokine production (Darville *et al.*, 2003, O'Connell *et al.*, 2006), however, the chlamydia PAMP that signals through TLR2 remains unknown (O'Connell *et al.*, 2006). Once within the cell, innate immunity comes into play as the bacterium induces the production of pro-inflammatory cytokines. The markers induced include interleukin 1 (IL-1), interleukin 6 (IL-6), Granulocyte-macrophage colony-stimulating and tumour necrosis factor-α (TNF-α) (Rasmussen *et al.*, 1997, Johnson, 2004). In addition NK cells and phagocytes such as neutrophils, dendritic cells (DCs) and macrophages are recruited by interleukin 8 (IL-8) which is secreted by the infected epithelial cells (Parr and Parr, 1990). These immune cells can then produce more inflammatory markers such as TNF-α which plays a role in restricting growth of *C. trachomatis* within the cell (Dessus-Babus *et al.*, 2002). These cytokines may also play a role in the pathology which is seen in infected cells (Darville *et al.*, 2003). IFN-γ, another inflammatory cytokine has many mechanisms by which it can control chlamydia replication. IFN-γ can promote the engulfment and destruction of the extracellular EBs by upregulation of the phagocyctic potential of macrophages. Secondly IFN-γ can inhibit chlamydia growth within infected cells either by upregulation of nitric oxide synthase (Chen *et al.*, 1996, Igietseme *et al.*, 1997) or by the production of indoleamine 2,3-dioxygenase which catalyses the catabolism (the breakdown of large molecules into smaller units) of trypthophan (Gupta *et al.*, 1994), which is one of the amino acids that chlamydia requires from the host. Lastly Chlamydia growth may be hindered by IFN-γ down regulating the transferrin receptor thereby reducing the intracellular stores of iron available to the organism (Byrd and Horwitz, 1993, Freidank *et al.*, 2001). The cells involved in linking the innate immune response to the adaptive immune response

are known as dendritic cells (DCs), these cells process and present chlamydia antigen to T cells (Steele *et al.*, 2004).

Cells of the adaptive immune system provide a defence system for a further encounter with *C. trachomatis* following the initial infection. The cells of the adaptive immune system include B and T cells (Roan and Starnbach, 2008). Soluble antigens are presented to the B cells via the B cell receptor and these cells are important in retaining a memory of past infections with *C. trachomatis* by producing antibodies (Roan and Starnbach, 2008), however, this may not be the case for all *C. trachomatis* serovars (Vos *et al.*, 1995). It is believed that B cells work together with CD^+ T cells in eliciting immunity to *C. trachomatis* (Roan and Starnbach, 2008).

Once inside the cells antibodies are unable to neutralise or enhance phagocytosis of EBs and hence T cells come into action to clear chlamydia infection (Roan and Starnbach, 2008). Several groups have shown that $CD4^+$ and $CD8^+$ T cells are important for bacterial clearance and they provide protection against chlamydia infection *in vivo* (Starnbach *et al.*, 1994, Su and Caldwell, 1995, Roan *et al.*, 2006, Roan and Starnbach, 2006). Antigens from EBs and RBs broken down by proteases and the peptides produced are presented by the antigen presenting cells (APCs) to $CD4^+$ via the major histocompatibility complex (MHC) class II molecules (Trombetta and Mellman, 2005). Following this, $CD4^+$ T cells produce effector T cells, the two major subtypes are Th1 and Th2 cells (Roan and Starnbach, 2008). Th1 cells produce IFN-γ which in turn increases the presentation of antigens to both $CD4^+$ and $CD8^+$ T cells (Gaczynska *et al.*, 1993, Steimle *et al.*, 1994). Th2 cells produce IL-4, IL-2 and IL-13. These cells may inhibit the development of protective Th1 responses thereby enhancing chlamydia load, this response can lead to continuous production of inflammatory cytokines which can subsequently lead to tissue destruction (Roan and Starnbach, 2008). In contrast $CD8^+$ recognises chlamydia proteins that have access to the cytosol of infected cells (Roan and Starnbach, 2008). Antigens form a complex with MHC class I molecules on cells surfaces and this activates $CD8^+$ cells (Cresswell *et al.*, 2005). $CD8^+$ cells also produce IFN-γ and can kill target cells which have the MHC class I complex presented on the surface of them but not cells which have not been infected (Roan and Starnbach, 2008).

Although all these mechanisms described in this section exist to clear chlamydia infections by the host and prevent subsequent infection, the bacterium still manages to evade the immune system. The bacterium has developed several mechanisms by which it evades the immune system, including down regulating the expression of MHC molecules in the infected cells, which in turn will limit their recognition by chlamydia-specific $CD4^+$ and $CD8^+$ T cells (Zhong *et al.*, 2001). Secondly *C. trachomatis* encodes deubiquitinases (DUBs) whose functions are unknown; however

current speculation is that deubiquitinases remove ubiquitin from chlamydia proteins which could subsequently lead to the reduction in MHC class I presentation of antigens to $CD8^+$ T cells (Misaghi *et al.*, 2006). *C. trachomatis* have also managed to develop mechanism to evade apoptosis by inducing anti-apoptotic factors long enough to complete its replication cycle and then inducing apoptosis once the RBS have differentiated back into infectious EBs in order to avoid necrosis (Byrne and Ojcius, 2004, Miyairi and Byrne, 2006), which would stimulate inflammation causing *Chlamydia* specific immune responses (Roan and Starnbach, 2008).

In vivo it has been shown that previous exposure to chlamydia does not produce immunity against re-infection (Brunham and Rey-Ladino, 2005). This is possibly due to the fact that *Chlamydia*- specific T cells do not develop memory T cells (Roan and Starnbach, 2008).

1.8 PERSISTENT SEXUALLY TRANSMITTED CHLAMYDIAL INFECTIONS.

There is very little known about the persistence of chlamydia *in vivo* but there is some speculation that it could lead to asymptomatic urethritis in males and PID and cervicitis in females. These individuals present with non-culturable specimens with strong serological titres (Wyrick, 2010).

Persistent chlamydial development occurs when *Chlamydiae* remain viable in host cells but are culture negative (Beatty *et al.*, 1994). The persistence of chlamydia was first described by Moulder *et al* (1980) who saw altered morphological forms. In persistent chlamydial infections there are many features that can be seen *in vitro*. These include the RBs not converting back to EBs and so the cycle remains incomplete (Moulder *et al.*, 1980); the RBs become rather enlarged and abnormally shaped as they do not undergo binary fission when in a persistent state (Wyrick, 2010); the infectious titre is reduced due to the lack of EBs and in some cases there can be increased resistance to antibiotics (Hafner and Timms, 2005). Experiments have been conducted in cell culture altering the conditions in which chlamydia is grown. These include exposure to IFN-γ which causes chlamydia to become dormant within the cell (Beatty *et al.*, 1994, Pantoja *et al.*, 2001), exposure to some antibiotics (Kutlin *et al.*, 1999), phage infection (Hsia *et al.*, 2000), limiting tryptophan, herpes simplex infection and lowering the amino acid concentration (Wyrick, 2010). In persistent infections, the infectious EBs are only recovered after the removal of IFN-γ or the removal of stress causing agents, which allows for the developmental cycle to resume (Beatty *et al.*, 1994). This shows that persistence is reversible *in vitro* and this may also be the case *in vivo* (Roan and Starnbach, 2008).

Once the immune response to chlamydia reduces or stops, the persistent chlamydia may reactivate, which in turn causes the immune system to become alert and inflammatory cells return to the site of infection, which then may drive the bacteria back into persistence, which may lead to further scarring of the tissue (Roan and Starnbach, 2008).

1.9 ANTIBIOTIC TREATMENT FOR *C. TRACHOMATIS.*

Treatment for *C. trachomatis* is readily available in the form of antibiotics, once the disease has been detected. The recommended treatments include azithromycin which is a single dose of 1g or doxycycline, here a 100mg dose twice daily for seven days is given. Treatment for pregnant and breast feeding women include azithromycin, erythromycin and amoxicillin (Kalwij *et al.*, 2010).

1.10 VACCINE DEVELOPMENT FOR *C. TRACHOMATIS.*

Previous attempts to develop a vaccine against *C. trachomatis* have proved unsuccessful. These approaches include using live attenuated or modified chlamydia organisms which had been passaged several times (Schautteet *et al.*, 2011). During this passaging one or more mutations may arise in the genome. These mutations may result in a non-virulent attenuated strain. This

approach produces a vaccine that elicits a humoral and cell mediated immunity although there is a possibility of the strain reverting to the wild-type strain and therefore resulting in disease or persistence (Schautteet *et al.*, 2011). Attempts to use live attenuated vaccines against trachoma in human and primates have proved unsuccessful (Longbottom and Livingstone, 2006). The recent development of a transformation system for *C. trachomatis* (Wang *et al.*, 2011a) may lead to further research being conducted into live attenuated vaccines as long as the immunopathological effects can be prevented (Longbottom, 2003).

In the early 1960s, trials were conducted in mice using inactivated whole organism, the results of the study were shocking and resulted in some individuals presenting with disease severity greater than that of those who were not vaccinated (Grayston *et al.*, 1963). Studies also show that there is no long term protection and the vaccination is serovar specific (Karunakaran *et al.*, 2010).

Research into the development of a vaccine has now changed and a subunit vaccine approach is seen as safer as it cannot revert to a virulent form and antigens that cause inflammatory damage and immunopathology can be avoided (Olive *et al.*, 2001). The main candidate being researched for a subunit vaccine is MOMP - the major outer membrane protein that makes up approximately 60% of the membrane (Longbottom, 2003).

Recent approaches to the development of a vaccine for *C. trachomatis* also include research into microbial delivery systems and DNA vaccination, which has been shown to elicit protective immunity in animal models systems (Entrican *et al.*, 2001). However, this response has not been seen in mouse models for genital tract infections (Pal *et al.*, 1999).

It is important that research into a suitable vaccine for chlamydia continues despite previous attempts being unsuccessful. *C. trachomatis* is asymptomatic in several individuals, therefore even though there is treatment in the form of antibiotics it does not reach everyone who has the infection and this could lead to further complications and the spread of chlamydia. Antibiotic treatment is therefore unsuitable for controlling transmission through populations and is also expensive especially for developing countries.

1.11 RISK FACTORS.

There are several factors associated with increased risk of contracting chlamydia, these include:

1.11.1 Age and gender.

Age is one of the main risk factors of *C. trachomatis* acquisitions. Infection is more common in younger sexually active women, especially among adolescents. *C. trachomatis* is more frequently detected in young people mainly between the ages of 18-25 years of age. Overall the prevalence of *C. trachomatis* in females has been shown to be higher than in males. HPA figures show that in

2010, 30% more females than males were *C. trachomatis* positive (HPA, 2012a). A higher risk may be seen in females due to anatomy; for instance, younger females have columnar epithelium on the cervix which supports the growth of *C. trachomatis*, changes in the flora and mucus production during this period also aids in the growth of *C. trachomatis*. Older women may also have acquired partial immunity after being infected in the past. The higher rate in younger individuals is also seen due to differences in sexual behaviour such as having premarital intercourse and the early onset of having sexual intercourse has also led to young individuals having many sex partners which increases the risk associated with contracting chlamydia. Lastly the correct use of condoms is also not as widely practiced in younger individuals in comparison to older age-groups, greatly increasing the chances of infection spreading from individual to individual (Navarro *et al.*, 2002).

1.11.2 Race and socioeconomic status (SES).

SES is often associated with race, however, the relationship between race, SES and chlamydial infection remain unclear. It may also be the case that women of higher SES may be more likely to have routine examinations, and therefore detection of asymptomatic cases within this group of people may bias reporting of chlamydia. However, some studies have suggested that chlamydia is detected in higher rates amongst individuals of black Caribbean and Asian descent (Klausner *et al.*, 2001, Creighton *et al.*, 2003).

1.11.3 Sexual behaviour.

Having frequent sexual relationships and failure to use barrier contraceptives such as condoms, diaphragms and cervical caps also act as risk factors for *C. trachomatis.* The use of oral contraception may also increase the risk of chlamydia infection by inducing cervical ectopy and making the cervix more susceptible to infection (Navarro *et al.*, 2002). Users of oral contraception are also less likely to use barrier contraceptives and therefore this increases the chances of them contracting sexually transmitted infections (Navarro *et al.*, 2002).

The number of sexual partners may increase the likelihood of contracting a sexually transmitted infection as it increases the probability of having sexual intercourse with someone who is infected. Additionally, having new or casual sexual partners may also increase the risk of becoming infected (Navarro *et al.*, 2002).

1.11.4 Marital status.

Other risk factors include marital status; chlamydia is more likely to be detected in individuals who are single as they are more likely to have multiple partners and change partners more frequently therefore increasing their chances of contracting chlamydia (Navarro *et al.*, 2002).

1.11.5 Transmission .

Genital chlamydia can be transmitted via the mucosal route during vaginal, anal and oral sex. It can also be passed on by use of unwashed sex toys (Nilsson *et al.*, 1997). Due to the asymptomatic nature of chlamydia an individual who is infected may pass the bacteria on to another sex partner without knowing.

A female who has contracted chlamydia can pass it to her child during vaginal childbirth. The child can then go on to develop pneumonia or conjunctivitis (inflammation and redness of the thin layer of tissue that covers the front of the eye).

Men who have sex with men are also at risk of chlamydial infections as chlamydia can be passed on through the anus and during oral sex.

1.12 NEONATAL AND INFANT INFECTIONS.

Neonates who are infected with chlamydia during birth can develop conjunctivitis and/or pneumonia. Conjunctivitis can develop at 5-14 days post-delivery in one eye and then can affect the other eye after 2-7 days. Oedema and erythema of the eyelids is also common with the discharge. A more serious infection, pneumonia, can occur between 1-19 weeks after birth, causing a fever, coughing and an increased respiratory rate. In premature babies pneumonia can be more serious and lead to respiratory distress and apnoea. Infants who are infected with chlamydia have an increased risk of later developing chronic pulmonary disease such as asthma. Conjunctivitis and pneumonia in infants is treated systemically using erythromycin (Pellowe and Pratt, 2006).

1.13 EPIDEMIOLOGY OF *C. TRACHOMATIS.*

Sexually transmitted infections (STIs) are infections that can be passed on from individual to individual by unprotected sexual intercourse. Not everyone who contracts an STI presents with signs and symptoms associated with that particular infection. If left untreated, STIs may cause severe complications in later life. There is an estimated 448 million new cases of ''curable'' STIs occurring each year, however, this figure may be an underestimate as many STIs are asymptomatic and/or are not reported and are therefore very difficult to control (Gewirtzman *et al.*, 2011). STIs include those caused by viral, bacterial, mycological and protozoal agents (e.g. HSV, HIV, HPV, hepatitis B, syphilis, gonorrhoea, chlamydia and trichomoniasis). The burden of STIs is mainly seen in South East Asia, Sub-Saharan Africa, Latin America and the Caribbean. As well as being a cause of morbidity in adults, STIs, may also result in complications including male and female infertility, ectopic pregnancy, cervical cancer, premature mortality, congenital syphilis, low birth weights and prematurity (Gewirtzman *et al.*, 2011). The costs associated with STIs are also

rather high especially for developing countries where the economic burden resulting from STIs is constantly increasing.

C. trachomatis is a worldwide public health problem and is responsible for most cases of STIs caused by bacterial pathogens. It can easily be passed on unknowingly as most cases are asymptomatic. Aside from being passed on through sexual transmission this organism can also be transmitted by droplets, hands, contaminated clothing, flies and during vaginal childbirth from mother to child (Gewirtzman *et al.*, 2011).

In the developing nations of the world as well as *C. trachomatis* being a STI it also causes trachoma, which affects the inner upper eyelid and the cornea. Trachoma is prevalent in Africa, Asia and South America. Approximately 2.2 million people are currently infected with trachoma in 53 countries and trachoma has caused irreversible blindness in 1.2 million people in these countries. The Global Alliance for the Elimination of Trachoma by 2020 (GET 2020) was established by WHO in 1997 and a year later was endorsed by the World Health Assembly (WHA) (WHO, 2012). The aims of GET 2020 is to implement a SAFE strategy. SAFE is an acronym which stands for Surgery, Antibiotic treatment, Facial Cleanliness and Environmental improvement, which are the four areas the WHO want to implement for individuals with trachoma (WHO, 2012).

Lymphogranuloma venereum (LGV) is highly prevalent in Asia, Africa and South America, however, since 2003; there has been an epidemic of LGV in the UK, North America and Australia. The LGV seen in these regions are mainly detected in the MSM population, especially those who are HIV positive and also consists of high levels of the L2b variant of LGV which causes proctitis. A recent study has also shown that L2b can also cause infection in women (Peuchant *et al.*, 2011). In England in 2011 there were 402 cases of LGV in the UK, which was a 26% decrease in LGV rates from 2010. Although this decrease occurred, guidelines have been revised to encourage individuals at risk to use condoms during sexual intercourse, use different condoms for different sex partners, reduce sexual partners and avoid over-lapping sexual relationships. Individuals also need to be aware of symptoms and get tested regularly, even though no symptoms are visible.

Not many studies have been conducted on the epidemiology of *C. trachomatis*, as these types of studies are often difficult to organise and ethical permission is difficult to obtain especially in developing countries.

Table 1.2 shows the global prevalence of curable STIs including chlamydia. These data have been collated by the WHO and are prone to inaccuracies due to different recording methods. The data show that the highest incidence of curable STIs is in South and South East Asia, followed by sub-Saharan Africa and then by Latin America and the Caribbean. This trend in the prevalence of STIs by region is also mirrored in the cases of chlamydial infection in adults in the regions; again, the

highest prevalence is seen in South and South East Asia, Sub-Saharan Africa and then in Latin America and the Caribbean.

1.13.1 STI incidence in England.

It has been well documented that the number of new cases of chlamydia continues to rise yearly. The incidence of other STIs is shown in Table 1.3 and adapted from a table published on the PHE website. On the other hand,the incidence of syphilis was less than other bacterial STI in England, with figures up to ten times lower than that of chlamydia (Table 1.3). Table 1.3 also shows the total number of individuals who have been infected with various STIs in 2002 and 2011 and also the percentage change in new diagnoses over this period. As can be seen in table 1.3, the percentage change of new diagnoses of chlamydia increased by 135% and this was the largest increase seen amongst the STIs that were recorded in this table. Gonorrhoea was the only STI that the diagnoses fell between 2002 to 2011. So, the question still remains as to why the number of individuals diagnosed with chlamydia continues to rise over the years and possible reasons are discussed in section 1.11.

Region	Estimated new cases of curable STI among adults (Millions)	Estimated cases of curable STI among adults (Millions)	Estimated new cases of chlamydial infections among adults (Millions)
North America	14	3	4
Latin America and the Caribbean	38	18.5	9.5
Western Europe	17	4	5
Eastern Europe & Central Asia	22	6	6
East Asia & Pacific	18	6	5.3
South and South East Asia	151	48	43
North Africa and the Middle East	10	3.5	3
Sub-Saharan Africa	69	32	16
Australia and New Zealand	1	0.25	0.34
Global Total	340	116.5	92

Table 1.2 Incidence of curable STIs and new cases of chlamydia infections by region during 1999 adapted from the WHO website (Scieux *et al.*, 1993).

	2002	2011	% change 2002-2011
Chlamydia	79,271	186,796	135%
Gonorrhoea	24,123	20,965	-13%
Herpes	17,259	31,154	81%
Syphillis	1,560	2,915	87%
Warts	62,982	76,071	21%
Other new STI diagnoses	25,008	25,436	2%

Table 1.3 The rates of STI diagnosis in England in 2002 and 2011.
The table indicates numbers of new STI diagnoses in both 2002 and 2011 and the percentage change (increase or decrease between the two years) (Olive and Bean, 1999).

1.13.2 Epidemiology of *C. trachomatis* in Europe

Two-thirds of *C. trachomatis* infections in Europe are found in individuals aged between 15 to 24. Chlamydia cases have continued to rise, although, it is not clear whether the increase in cases is due to increased testing as opposed to just an increase in the level of chlamydia. In 2010, 344,491 cases were reported from 24 of the EU member states and there were 186 cases per 100,000 population (ECDC, 2012). The risk factors include being below the age of 24, having a new sexual partner, having multiple sexual partners and not using condoms during sexual intercourse. Figure 1.4 shows the trend in eight European countries. The highest rate of chlamydia is seen in Iceland and the lowest in Latvia.

The number of new diagnoses of sexually transmitted infected individuals in England rose by 2% in 2011 compared to 2010. There was also a 25% increase in gonorrhoea, 10% increase in syphilis, 5% increase in herpes and a 1% increase in genital warts in that time period.

Since 1999 the number of reported annual cases of chlamydia has more than doubled and in 2010 there were 189,612 new diagnoses of *C. trachomatis* in England.

The increase in STI diagnoses can be attributed to several factors including the increase in sexual health screenings, including chlamydia screening in GUM clinics and the National Chlamydia Screening Programme (NCSP), which aims to control chlamydia and reduce the sequelae therefore screening through this programme is recommended annually and after change of sexual partners. The use of nucleic acid amplification tests (NAAT) which offer more sensitive diagnostic testing can also be another reason why rates continue to rise.

In 2011, 2.1 million chlamydia tests were carried out in England among young adults aged 15-24 years, with 147,594 diagnoses being made in this age group. Southampton has one of the highest rates of sexually transmitted infection in the UK. The city has seen a 38% increase in *C. trachomatis* between 2000 and 2009 and was highlighted as having one of the worst infection rates by the Health Protection Agency (HPA) in 2010. Genital *C. trachomatis* is the most common STI diagnosed and treated in the UK.

In the past ten years, there have been two major incidents, which have occurred in Europe and caused the number of people infected with *C. trachomatis* to increase. The incidents occurred in Sweden and in the Netherlands.

Before 2003, LGV had not been a problem in the industrialised world. However, an outbreak of LGV was detected in the Netherlands amongst MSM who were co infected with HIV in 2003. Outbreaks in the UK and the United States followed shortly afterwards. The outbreak appeared to begin with a white bisexual man in Rotterdam, Netherlands who presented with LGV (L2). Upon

further investigation it was shown that the outbreak was caused by a newly identified variant that was called L2b (Spaargaren *et al.*, 2005).

In Sweden, figures of *C. trachomatis* had remained relatively low compared to other European countries until 2006, when a sharp rise in the number of people diagnosed with *C. trachomatis* was documented. The rise was attributed to a new variant of *C. trachomatis* (nvCT) which had evaded the routine diagnostic tests available at the time, due to mutations within the genome of serotype E, which is the most common serotype found in patients infected with *C. trachomatis* (Everett and Hatch, 1995). The new variant has a 377bp deletion in the plasmid DNA and it was this region that was used for nucleic acid amplification tests (Seth-Smith *et al.*, 2009). This gave rise to several thousand false-negative results (Klint *et al.*, 2011). In addition to the 377bp deletion in CDS1 there was also a 44bp duplication immediately upstream of CDS3 (Seth-Smith *et al.*, 2009). The rise of individuals diagnosed with *C. trachomatis* came as a shock as contact tracing has been mandatory in Sweden since 1988 (Osterlund *et al.*, 2005). Due to the evasion of the new variant *C. trachomatis* in Sweden of the current typing methods, a high resolution typing method to facilitate a detailed understanding of the spread of the disease and to provide the opportunity for monitoring the development of the bacteria will be useful.

1.13.3 Epidemiology of *C. trachomatis* in the USA.

According to the Centers for Disease Control (CDC) and Prevention a total of 1,307,893 chlamydial infections were reported across 50 states which corresponds to a rate of 426 cases per 100,000 population, which was a 5.1% increase compared to 2009 rates of chlamydia in the USA. Like England, the highest rates of chlamydia infection are seen in individuals aged between 15 to 24 and more women are diagnosed with chlamydia compared to men, suggesting that the sex partners of women are neither being screened for chlamydia nor are they attending GUM clinics. The rates of chlamydia in the USA are mainly in black men and women. The highest rates of chlamydia in black individuals were eight times the rate of positives seen in white individuals according to figures from 2010 (CDC, 2011). High rates of chlamydia were also seen in American Indians/Alaska Natives and Hispanics which were 4.3 and 2.7 times higher respectively than the rates detected in white individuals (CDC, 2011).

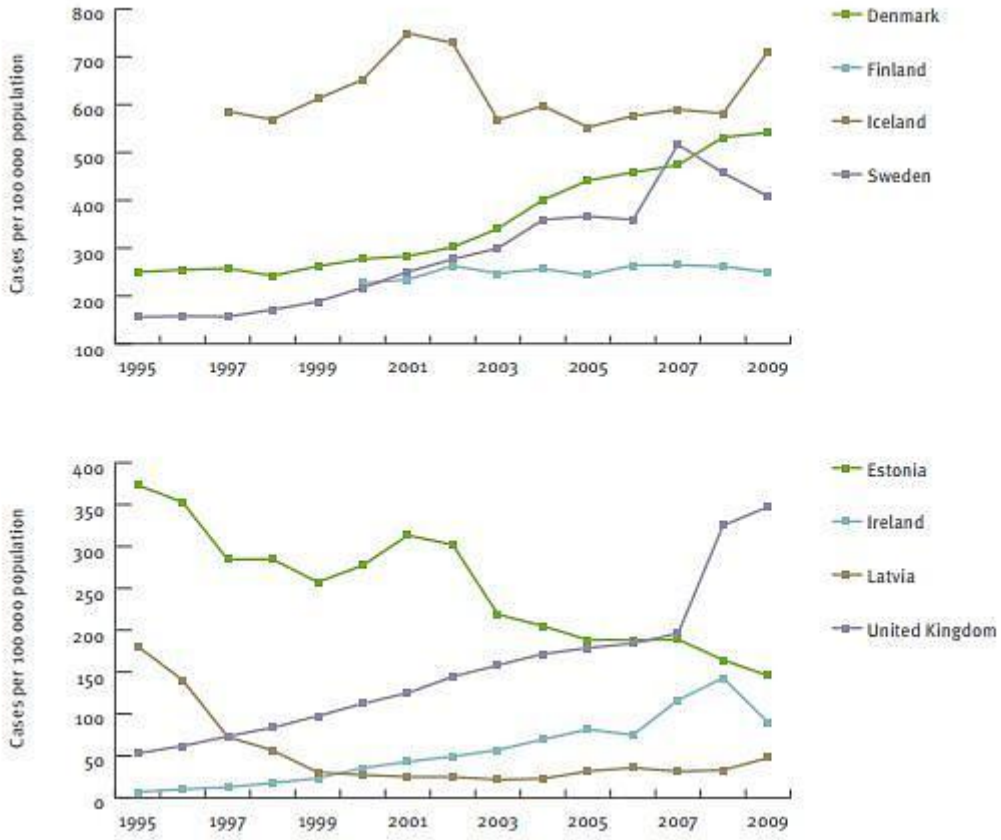


Figure 1.4 The cases of chlamydia per 100, 000 population in selected EU member countries between 1995 and 2009 (ECDC, 2011) .

1.14 CHLAMYDIA TRACHOMATIS DIAGNOSIS

1.14.1 Isolation of *C. trachomatis* in cell culture.

Culture is the only test that confirms the presence of viable organisms (Chernesky, 2005). The procedure involves inoculating a confluent monolayer of cells with a specimen that has been collected and transported appropriately to maintain the viability of the organism (Williams *et al.*, 1990). Inoculation involves centrifugation of the specimen onto cycloheximide treated monolayer cultures of cells. The centrifugation step is not required for LGV serotypes as these samples are able to infect cells without centrifugation. After 48 to 72 hours of growth at 35°C in 5% (v/v) CO_2, infected cells contain intracellular inclusions, which can be stained with a fluorescein-conjugated monoclonal antibody that is specific for the major outer membrane protein (MOMP) of *C. trachomatis*. There are several disadvantages of cell culture and these include its low sensitivity, high cost, long turnaround times and the stringent storage and transport requirements. The technique also requires staff with high technical skills to carry out the procedure (Williams *et al.*, 1990).

1.14.2 Enzyme Immuno-assay (EIA) Tests.

There are a number of commercial EIA tests available to detect *C. trachomatis* infection. *C. trachomatis* is detected using either a monoclonal or polyclonal antibody (labelled with an enzyme) against the Lipopolysaccharide (LPS). The enzyme converts a colourless substrate into a coloured product, which can then be detected using a spectrophotometer. One of the disadvantages with using EIA is that there is potential for false positive results caused by cross – reaction with LPS of other microorganisms. For these reasons, blocking assays have been developed where a monoclonal antibody is used that is specific for chlamydia LPS and verifies the initial positive result. The monoclonal antibody inhibits chlamydia-specific binding by the enzyme-labelled antibody. The results from this assay are verification assays of the initial EIA test carried out (CDC, 2002).

1.14.3 Direct Fluorescent Antibody (DFA) Tests.

Specimens can also be placed onto a specimen well of a slide, the slide is then dried and a fixative is applied. The specimen is then stained with a fluorescein-labelled monoclonal antibody that binds to *C. trachomatis* elementary bodies. Subsequently, this can be viewed by fluorescence microscopy. The antibody that is detected by DFA is either the MOMP or the LPS molecule. Only *C. trachomatis* organisms will stain with the anti-MOMP antibodies, whereas the anti-LPS monoclonal antibodies used in certain kits can cross-react with non-chlamydial bacterial species. However, DFA is very time consuming and only suitable for laboratories that have a fluorescent microscope and an expert microscopist who knows how to use the microscope.

1.14.4 Nucleic Acid Hybridisation (Nucleic Acid Probe) Tests (NAH tests).

There are currently two commercially available nucleic acid hybridisation assays; the Gen-Probe PACE 2 and the Diagene Hybrid capture. In the Gen-Probe hybridisation assays, a single-stranded DNA probe with chemiluminescent labels that are complementary to a specific sequence of *C. trachomatis* rRNA is implemented. The labelled DNA probes combine with *C. trachomatis* to form stable DNA: RNA hybrids. In the Diagene assay RNA probes which are specific for DNA sequences in *C. trachomatis* are used. There are also various versions of the test which can be used to detect both *C. trachomatis* and *N. gonorrhoeae*. These tests are advantageous as samples can be stored for up to seven days at room temperature before the test is carried out.

1.14.5 Serological tests

Serology is not recommended for *C. trachomatis* as the method cannot be used to determine between an old infection and a new infection, as previous infections often elicits long lasting antibodies. However, serology may still be used for neonatal cases where chlamydia is suspected.

1.14.6 Nucleic Acid amplification tests (NAATs)

NAATs are tests designed to amplify sequences that are specific to the organism being tested and for this, viable organisms are not required. Additionally, NAATs test are very sensitive and can produce a positive signal from a single copy of the *C. trachomatis* DNA or RNA. There are several tests available which use different reactions to test *C. trachomatis* samples. The Roche Amplicor® test uses polymerase chain reaction (PCR), the BD ProbeTec™ ET test uses strand displacement amplification, the Abbot LCx® test uses ligase chain reaction, whilst the Gen-Probe Aptima test uses transcription-mediated amplification (TMA) (Williams *et al.*, 1990, Chernesky, 2005). All the tests listed apart from the Gen-Probe Aptima TMA test amplify sequences of the cryptic plasmid (a plasmid with no known function), which is present in multiple copies of the EB. TMA is designed to detect a 23S rRNA target. Both the plasmid and the rRNA are found in multiple copies in *C. trachomatis.* The sensitivity of NAATs tests can be as high as 96% but because of sampling variability and the presence of inhibitors in reactions this can sometimes be lower (Chernesky, 2005). As long as cross contamination of the specimens is avoided, the reactions used for NAATs are also very specific as the primers are not known to cross-react with DNA from other bacteria found in humans (Williams *et al.*, 1990).

Co-infection is often an issue in individuals who have contracted an STI; there is often a high possibility that they have contracted multiple STIs. This has been the case with chlamydia and gonorrohoea. Due to the high incidence of co-infection, diagnostic assays have been developed for the detection of both *C. trachomatis* and *Neisseria gonorrhoeae* as they are the two most prevalent bacterial STIs. This is also one of the main advantages with NAAT testing.

1.14.7 *C. trachomatis* point of care tests.

It has been shown that point of care tests can be less specific and less sensitive than other laboratory-based tests. The advantage of point of care tests are that they are rapid tests which can be performed whilst the patient waits and are useful in places such as clinics, hospitals and GP surgeries, where it is advantageous to have the test results whilst the patient is still present (Williams *et al.*, 1990). However, the cost of performing such tests are high in comparison to some laboratory based tests and therefore, point of care test for *C. trachomatis* is yet to become available (Welsh and McClelland, 1990). Additionally, point of care tests are currently less sensitive than NAATs tests.

1.15 NON MOLECULAR TYPING METHODS FOR *C. TRACHOMATIS*

Typing methods are being developed continuously and there are several reasons why it is advantageous to develop a typing scheme including:

- The ability to use a typing scheme to reveal transmission patterns in sexual networks and may be used as a tool in partner notification.
- It may also enable association with clinical manifestations and pathogenicity.
- Tissue or organ affinity for certain strains may be discovered, when typing data can be linked to patient data.
- Typing may help determine whether infections are persistent or new.
- For evolutionary surveillance of specific clones.
- It could play a role in cases of sexual abuse or assaults (Pedersen *et al.*, 2009).

In the early 1960s, it became apparent that for immunisation against trachoma to be possible, the different *C. trachomatis* strains would need to be characterised and the antigenic differences between strains would need to be known (Pedersen *et al.*, 2009). In 1963, Wang and Grayston published the development of an immunotyping method known as a mouse toxicity prevention test (MTPT) (Wang and Grayston, 1963). In this test, egg yolk sacs were inoculated with Trachoma Inclusion Conjunctivitis (TRIC) organism and were intravenously injected into mice. This triggered an immune response within the mice. The mice were then challenged with the same or different isolates of chlamydia. The mice that received the same isolate during the challenge were protected against the bacteria and some mice received cross-protection and so were protected even though they had received a different serovar during the challenge than that with which they were initially injected. The authors were able to classify 80 strains into six immunological types (Wang *et al.*, 1985); however, this took them seven years to do. This method was deemed tedious, expensive, and time-consuming and other methods needed to be developed for classification of *C. trachomatis*.

Subsequently a micro-immunofluorescent test was developed (Wang and Grayston, 1970, Wang *et al.*, 1973). Here, isolated strains were propagated in yolk sacs before use. Antiserum for strain typing was prepared by intravenous immunisation of mice with 1%(v/v) crude yolk sac suspension from infected eggs. In later years, due to the cross reactivity between related types, polyclonal antibodies were later replaced with monoclonal antibodies, which were more specific (Stephens *et al.*, 1982, Wang *et al.*, 1985). Antibodies to the MOMP of *C. trachomatis* were used to identify the different serovars before typing systems were used (Stephens *et al.*, 1982, Wang *et al.*, 1985, Barnes *et al.*, 1987).

Tests based on radio immuoassays (RIA) and fluorescent-antibody staining (FA) (van der Pol and Jones, 1992) were also developed and large numbers of *C. trachomatis* isolates could be typed using a dot-ELISA method (Barnes *et al.*, 1985).

These techniques need a lot of expertise and are time consuming as well as being expensive. Another disadvantage of such techniques is that serovars unrecognised by the antibody panels available might evolve and be left undetected (Wang *et al.*, 1985, Barnes *et al.*, 1987)

1.16 General overview of genotyping systems in bacteriology.

The ability to manipulate DNA has led to the development of many techniques that are currently used by biologists today:

1.16.1 Pulsed-field gel electrophoresis (PFGE).

Pulsed-field gel electrophoresis (PFGE) is very similar to standard gel electrophoresis. However, the voltage is switched in three directions during the procedure as opposed to just one direction in standard gel electrophoresis. PFGE is normally used in situations where the DNA molecule is longer than 25 kb and therefore will be poorly resolved using standard gel electrophoresis. This method requires large amounts of high quality DNA and normally rare cutter enzymes are used to result in strain-specific banding patterns. For *C. trachomatis* this method requires pure EBs to visualise DNA patterns related specifically to *C. trachomatis* without interference from host DNA (Pedersen *et al.*, 2009) and it is very difficult to obtain high quality chromosomal *C. trachomatis* DNA. However, despite PFGE's high discriminatory power when used to analyse other bacteria such as *E. coli* (Yuan *et al.*, 1989), when used for *C. trachomatis* isolates it is less discriminatory than serotyping (Rodriguez *et al.*, 1994). Another disadvantage of using PFGE is that it is time consuming (Olive and Bean, 1999).

1.16.2 Restriction fragment length polymorphism (RFLP).

Restriction fragment length polymorphism (RFLP) exploits the variations in DNA sequences. Briefly, chromosomal DNA is digested with a restriction enzyme and the fragments are separated by electrophoresis through an agarose gel. The DNA fragments within the gel are then transferred to a membrane (nylon or cellulose) by Southern blotting. Once the fragments have been transferred, the membrane then hybridises to a labelled probe homologous to the gene to be examined. The blotting technique has now largely been replaced by PCR-based RFLP (Olive and Bean, 1999). The loci of interest are amplified by PCR and this is followed by RFLP analysis. The DNA fragments, following restriction digest, are then separated on an agarose or acrylamide gel. For *C. trachomatis* this has shown to provide better results than PFGE as the results obtained from RFLP are comparable to the results obtained from serotyping (Rodriguez *et al.*, 1991).

1.16.3 Random amplification of polymorphic DNA (RAPD).

The use of short sequence primers (9 to 10 bases in length) which hybridise to chromosomal DNA sequences at low annealing temperatures, is known as Random Amplification of Polymorphic DNA (RAPD) (Welsh and McClelland, 1990, Williams *et al.*, 1990). The primers used for RAPD are made up of random sequences and are used to amplify small amounts of total genomic DNA by PCR (Bardakci, 2000). Products are then separated on an agarose gel and are subsequently visualised. The use of RAPD does not reflect serotyping results, however, it has been suggested that RAPD could be used to supplement *omp*A genotyping (Scieux *et al.*, 1993). RAPD is extensively used as there are no requirements for cloning or sequencing (Bardakci, 2000).

1.16.4 Amplified fragment length polymorphism (AFLP)

Amplified Fragment Length Polymorphisms (AFLP) are differences in restriction fragment lengths which are usually caused by SNPs or INDELs that create or eradicate restriction endonuclease recognition sites. Amplified fragment length polymorphism (AFLP) is a genome fingerprinting technique which uses restriction enzymes to cut genomic DNA (Vos *et al.*, 1995) and is based on the selective PCR amplification of restriction fragments. Adaptors which contain sequences homologous to a PCR binding site are subsequently ligated to the ends of the digestion fragments. The restriction fragments can then be amplified and visualised on gels. This technique produces reproducible results to *C. trachomatis* serotyping (Morre *et al.*, 2000). However this technique is not easy to handle as there is need to develop locus-specific markers from individual fragments and this can be difficult (Pedersen *et al.*, 2009).

1.16.5 *omp*A PCR and sequencing.

It has been shown that sequence determination of *omp*A provides a higher resolution than any of the above techniques and does not involve a culturing step. Primers are used to amplify the *omp*A region during a PCR reaction and the amplicon produced is then sequenced to determine the bases present in the amplicon.

*omp*A PCR and sequencing is still widely used for typing *C. trachomatis* strains, however the level of information this technique can offer is very limited. One of the disadvantages of *omp*A genotyping is that it has a low discriminatory power.

1.16.6 Multilocus sequence typing (MLST).

Multilocus sequence typing (MLST) is a typing system, which analyses 500-700 base pairs of 6-8 conserved housekeeping genes within a bacterial genome. The development and validation of MLST was first documented in 1998 where the bacterial pathogen of choice to develop this method was *Neisseria meningitidis* (Maiden *et al.*, 1998). This method takes advantage of housekeeping genes within the genome, which are utilised as they are stable and have a fundamental role in cell survival; therefore this typing scheme is intended for long term and

global epidemiology. The direct assignment of alleles based on nucleotide sequence determination distinguishes more alleles per locus and therefore offers high levels of discrimination between isolates and also offers a method which is comparable between labs (Maiden *et al.*, 1998). The genes chosen for this typing scheme should also be non-selectable, widely separated on the chromosome and should not be adjacent to the outer membrane protein or hypothetical proteins as these may be under diversifying selection (Maiden, 2006). Advantages of MLST include precision, allowing inter-laboratory comparisons, good discrimination between strains and it also offers buffering against the distorting effect of recombination on genetic relatedness (Pedersen *et al.*, 2009).

There have been several papers published that describe multilocus typing schemes for *C. trachomatis.* In 2007, Klint *et al.* analysed the MLST profile of 47 isolates and managed to detect 32 variants using the MLST system, whereas only 12 variants could be detected by use of *omp*A analysis (Klint *et al.*, 2007). This system, although highly discriminatory does not follow the rules of MLST. Firstly, the authors use genes (*hct*B and *pbp*B; these are DNA and penicillin binding protein respectively) which are under immune selection and secondly the PCR products produced by some of these genes are too large and therefore do not follow the rules outlined for this scheme to be used as an MLST typing system. On the other hand, Pannekoek *et al.* took a different approach and investigated the evolutionary relationships of the *Chlamydiacea* family and found that their MLST typing scheme provided almost no resolution for *C. pneumoniae,* whereas analysis of 26 *C. trachomatis* strains formed three clonal complexes consistent with that obtained when 16S rRNA and 23S rRNA genes were used in the phylogenetic analyses (Pannekoek *et al.*, 2008). Lymphogranuloma venereum isolates formed one cluster, whilst the more common *omp*A genotypes (D-F) formed the second cluster. The third cluster was formed by the non-prevalent genital tract strains (G-K) and those that cause trachoma (Pannekoek *et al.*, 2008). Findings by Pannekoek and colleagues were confirmed by work carried out by Dean *et al.* who analysed isolates from various continents of the world (Africa, Northern and Southern Europe, Asia, North and South America). They also found that by assigning genotypes to strains using seven different housekeeping genes (*gly*A, *mdh*C, *pdh*A, *yhb*G, *pyk*F, *lys*S and *leu*S) to the genes used by Pannekoek *et al* (*gat*A, *opp*A-3, *hfl*X, *gid*A, *eno*A, *hem*N and *fum*C), the strains in which they tested the typing scheme on formed three clusters (Dean *et al.*, 2009). This work showed that MLST can be used to compare isolates from different parts of the world and diverse and emerging *C. trachomatis* strains can also be detected by this method (Dean *et al.*, 2009).

1.16.7 Variable number tandem repeat (VNTR).

When tracking a strain in a community a high discriminatory power is required and this is why other methods have been developed. In 2008, Pedersen *et al* published a paper that described a

method for discriminating *C. trachomatis* using *omp*A and three variable number tandem repeats (VNTR). VNTR is defined as a region of nucleotide repeats or motifs and is used to type various strains of *C. trachomatis.* With this typing scheme various subtypes within an *omp*A genotype can be detected and thereby this increases the discriminatory power when assigning genotypes to *C. trachomatis* strains than is seen when typing only the *omp*A gene (Pedersen *et al.*, 2008, Wang *et al.*, 2011b). The three VNTR regions were designated as CT1335, CT1299, and CT1291 (Pedersen *et al.*, 2008). These three regions were chosen on the basis that amongst all the possible candidates assessed by the group, these regions were found to have very high discriminatory power, whilst other candidates showed genetic variation that was too low (Pedersen *et al.*, 2008). This MLVA-*omp*A typing system has been designed to study short term, local epidemiology in contrast to the previously developed MLST for chlamydia. Due to the presumed elevated error rate of DNA polymerase in these VNTR regions this typing system is considered suitable for short term, local epidemiological studies. This group found that the use of both *omp*A and the three VNTR regions gave a combined Simpsons index of diversity of 0.94 showing that this typing system offers a higher level of diversity than just typing with *omp*A or the VNTR markers alone (Pedersen *et al.*, 2008). The Simpsons index of diversity is a way of determining a method's ability to assign a different type to two unrelated strains sampled randomly from the population of a given species (van Belkum *et al.*, 2007). A Simpsons index of diversity of one shows an ideal typing system. However, a value in the order of 0.95 is seen as acceptable for a typing system (van Belkum *et al.*, 2007).

The reproducibility and resolution/typeability of this typing system was also evaluated during the study. Reproducibility is defined as the ability of the typing method to assign the same type to an isolate at different places, by different people and at different times. The reproducibility of the VNTR and *omp*A typing system was 100%. Typeability which is defined as the ability to assign a type to all isolates tested by the typing system was 100% for *omp*A and 99% for VNTR (Theunissen *et al.*, 1992).

Before the start of this thesis, colleagues in the Chlamydia group in Southampton evaluated this typing system by sampling a local population. Within all the *omp*A genotypes they found numerous multilocus VNTR subtypes, and in the most common genotypes there were defined MLVA subtypes, which may reflect different demographics such as location of people, age group and sexual networks (Wang *et al.*, 2011b).

There are two ways of achieving high levels of discrimination and these are to identify regions within the genome that are highly variable within the bacterial population, this is useful for short term epidemiological studies where it is necessary to identify micro variation in order to study

strains circulating within a geographical area. The second approach is to choose regions which accumulate variation very slowly in the population in order to study long term epidemiology and thereby can be used for the identification of lineages and global epidemiological studies (e.g. MLST markers). Therefore, depending on the scientific question the above methods will be useful in different ways. The Pedersen *et al* scheme, although it offers a high resolution typing method than that offered by typing solely the *omp*A gene, the loci chosen for this scheme are not evenly distributed across the whole of the genome and therefore any changes which occur in the area of the genomes that has not been covered by these markers can potentially be missed. Some users have also experienced some problems with this typing scheme where data have been ambiguous and interpretation of the data was user dependent (Bom *et al.*, 2011). This difficulty can be attributed to the type of polymerase used for experimental work.

1.17 AIMS OF PROJECT

- To evaluate the current typing methods available for typing *C. trachomatis* by testing the stability of the MLVA-*omp*A (Pedersen *et al.*, 2008) and the MST typing scheme for *C. trachomatis* (Klint *et al., 2007*).
- To apply the MLVA-*omp*A genotyping system to clinical samples to determine the diversity of types within men who have sex with men in Brighton.
- To apply the MLVA-*omp*A genotyping system to clinical samples to determine the diversity of types within a heterosexual population in Southampton and linking of the types to demographic information.
- To analyse whole genome sequence data from *C. trachomatis* strains and to identify key sequences and markers for strain discrimination to study local epidemiology. Key sequences and markers may also be useful for strain discrimination in terms of pathogenicity, tropisms and symptom-presentation.

CHAPTER 2 MATERIAL AND METHODS.

2.1 CHEMICALS AND SOLUTIONS.

2.1.1 Water.

Deionised water (dH_2O) was produced by reverse osmosis. The process was carried out using the SG reverse-osmosis-system euro™ (Triple Red LTD). Solutions for nucleic acid manipulation were prepared using ultra high quality water (UHQ H_2O). UHQ H_2O was prepared using the Barnstead/ Thermolyne NANOpure®Diamond™ life science (UV/ UF) ultrapure water system (Triple Red Ltd). This system further purifies dH_2O by reverse osmosis to a resistance of 18 mega-ohms and autoclaved at 1.05kg/cm^2 for 15 minutes.

2.1.2 Sterilisation.

Autoclaving: Pipette tips, media, bulk solution and glassware were autoclaved at 120°C (15Ibs/sq.inch for 20 minutes).

Filtration: Stock solutions of gentamicin (1mg/ml), cycloheximide (10mg/ml), 4SP, vancomycin (1mg/ml) and glucose were sterilised by filtration through a 0.22µm Tuffryn Millipore filter (Pall Life Sciences, Michigan, USA).

2.1.3 Buffers, reagents and solutions.

Buffers, reagents and solutions where commercially sourced and made using UHQ. A 0.2µM syringe filter was used to filter the buffer, reagent or solution if required.

Name	Final concentration
Phosphate buffer saline (PBS)	
Trypsin/EDTA (TE)	5% (v/v) Trypsin / 20% (v/v) EDTA
4SP	
Cycloheximide	1mg/ml
Gentamicin	10mg/ml
Fungizone	250µg/ml
Vancomycin	1mg/ml
Glucose	3M Stock solution
Orange G loading dye	6X

Table 2.1 Buffers, reagents and solutions used for all projects and the final Concentrations they were used at for studies in this thesis.

2.1.4 Antibodies.

2.1.4.1 Primary antibody.

The primary antibody used is called antibody 29. Antibody 29 is used to detect the LPS of chlamydia. This antibody was produced in the Molecular Microbiology department (University of Southampton) by injecting a mouse with hybridoma cells containing the clone. The clone was injected into the peritoneal cavity of the mouse. The hybridoma cells multiplied and produced fluid (ascites) in the abdomen of the mouse containing a high concentration of the antibody. This antibody was used at a dilution of 1/50.

2.1.4.2 Secondary antibody.

The secondary antibody used is the anti-mouse IgG, heavy and light chain specific (goat) β-galactosidase conjugate.

2.1.5 Plastic and Glassware.

Sterile plastic ware (i.e. universals, bijoux and stripettes) were purchased from various suppliers. Pipette tips were purchased from Alpha Laboratories Ltd (Eastleigh, Hampshire) and autoclaved at 18Psi for 30 minutes to reduce the possibility of contamination from nuclease.

2.2 BACTERIAL GROWTH MEDIUM AND SOLUTIONS.

2.2.1 Media.

Dulbecco Modified Eagle medium (DMEM) high glucose, 200mM glutamine (Invitrogen) supplemented with 10% (v/v) foetal calf serum (Invitrogen) was used for cell culture.

During the culture of clinical samples the media was supplemented with 10% (v/v) foetal calf serum (FCS) containing, 1mg/ml of vancomycin, 10mg/ml gentamicin (Sigma), 1mg/ml cycloheximide (Sigma) and 250µg/ml of fungizone (Gibco, UK).

2.2.2 Antibiotics.

Gentamicin (Sigma) and vancomycin were dissolved in UHQ H_2O to 10mg/ml and 1mg/ml respectively.

2.3 CHLAMYDIA HOST CELL LINE.

Chlamydia cell culture was carried out in McCoy B cells. McCoy B cells are fibroblast mouse cells and differ from McCoy A, cells which are of human origin and were reported to have originated from the synovial fluid in the knee joint of a patient suffering from degenerative arthritis.

2.3.1 Cell culture.

Monolayers of McCoy cells were grown in T25 or T75 tissue culture flasks with non-filter caps (NUNC). Confluent monolayers were achieved 24-48 hours after incubation at 37°C with 5% (v/v) CO_2 with DMEM supplemented with 10% (v/v) foetal calf serum. On the day of passaging the

cells, all necks of flasks and bottles were wiped with 70% (v/v) ethanol to prevent contamination. The media in the flask was discarded and the cells washed twice with phosphate buffer saline (PBS) to remove all traces of DMEM as the FCS may inhibit the trypsinisation process. Trypsin containing EDTA was added to the cells with any excess being poured off and then the cells left to detach for five minutes. New flasks where then labeled with the split ratio and the date and 5ml or 19ml of DMEM was added to a T25 and T75 respectively. After five minutes the flask containing the cells were tapped to ensure that all cells were now detached. The appropriate amount of DMEM was subsequently added to the flask depending on the split ratio required. The remaining cells were washed up and down the flask with the media by pipetting several times to break up the clumps of cells. A calculated volume was then added to the 'new' flasks dependent on the split ratio required. The caps of the tissue culture flasks were loosened before incubation at 37°C/5% (v/v) CO_2. Cells were passaged routinely twice a week.

2.4 DETECTION OF MYCOPLASMA BY PCR.

The McCoy cell line and the cultured *C. trachomatis* patient samples were tested for mycoplasma using the Lookout Mycoplasma PCR detection kit by Sigma-Aldrich (Sigma-Aldrich, Poole, UK). The mycoplasma test was set up in a final volume of 25µl. The required volume of rehydration buffer was pipetted into a clean amplification tube along with the required volume of DNA polymerase. For each sample, 23µl of DNA polymerase/Rehydration Buffer was added to the tube followed by the addition of 2µl of the sample to be tested. For the negative control 2µl of DNA-free water was added to the tube. In each reaction 0.5µl of JumpStart *Taq* polymerase was added to 24.5µl of Rehydration buffer. For the positive control 24.5µl of Rehydration buffer and 0.5µl of JumpStart *Taq* polymerase was added to the relevant tube. Once the required buffers and sample (or water in the case of the negative control) were in the amplification tubes, the tubes were incubated at room temperature for five minutes.

PCR was then performed on the samples. The thermal cycling program followed for the mycoplasma test was as follows:-

- 1 cycle at 94°C for 2 minutes
- 40 cycles at 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 40 seconds
- Cool down to 4-8°C

8µl of each PCR product was then added to a 1.2% (v/v) agarose gel to determine if mycoplasma contamination is present in the sample. Internal controls produced a band at 481bp in all of the lanes where the sample has been added. The positive control shows a band at 259bp and depending on the concentration of mycoplasma present an internal control will also be present at

481bp. If mycoplasma is present in the sample then a band will be present in the range of 260 ± 8bp.

The McCoy Cell line which was used for experimental purposes in this thesis was tested on a regular basis for mycoplasma and was shown to be negative for mycoplasma throughout all the projects.

2.5 CHLAMYDIA SPECIES.

All work was conducted using the*Chlamydia trachomatis.*

2.6 INFECTION OF CELL MONOLAYER WITH *C. TRACHOMATIS.*

McCoy cells were seeded on to wells or into flasks 24 hours pre-infection. Cells were grown overnight in Dulbecco's Modified Eagle medium (DMEM) containing 10% (v/v) FCS and then incubated at 37°C with 5% (v/v) CO_2. On the day of infection the medium was removed and replaced with inoculum containing the *C. trachomatis* isolates. This was then centrifuged at 754×g for 30 minutes. Subsequently, the inoculum was replaced with DMEM containing cycloheximide (1µg/mL) and gentamicin (20µg/mL). The cells were then incubated at 37°C for 48 hours. After 48 hours the inclusions were large enough to be visualised under a light microscope and the extent of infection could also be determined visually to decide whether a larger surface area was required for the next passage.

2.7 HARVESTING OF *C. TRACHOMATIS* ISOLATES FROM WELLS AND FLASKS.

Two days after infection, cells were harvested. For this a cell scraper was used to detach the cells from the plastic of the flask. The sample were then centrifuged at 2851×g for 10 minutes, the supernatant discarded and the pellet resuspended in 1:10 cold phosphate buffer saline (PBS). The sample was then added to glass beads and vortexed for 1 min to release the elementary bodies from the cell, centrifuged at 110×g for 5min to remove any cell debris and the supernatant was then added to 4 Sucrose Phosphate (4SP) and stored at -80°C. To harvest from wells the cells were scraped up into the medium using a sterilised pipette tip, the 10 minute centrifugation step was omitted as was the addition of PBS, instead the harvest was immediately beaded and then the protocol for harvesting was continued as for harvesting from flasks.

2.8 TITRATION.

McCoy Cells were seeded in a 96 well tray the day prior to the titrations being carried out. Titrations were carried out in a tray which did not contain any cells as in figure 2.1 and then transferred into the tray containing the cells. Dilution series were conducted in medium (DMEM +10% (v/v) FCS). Medium was removed in the trays containing the cells and then replaced with 100µl of the dilution series. The tray was then incubated at 37°C for 48 hours.

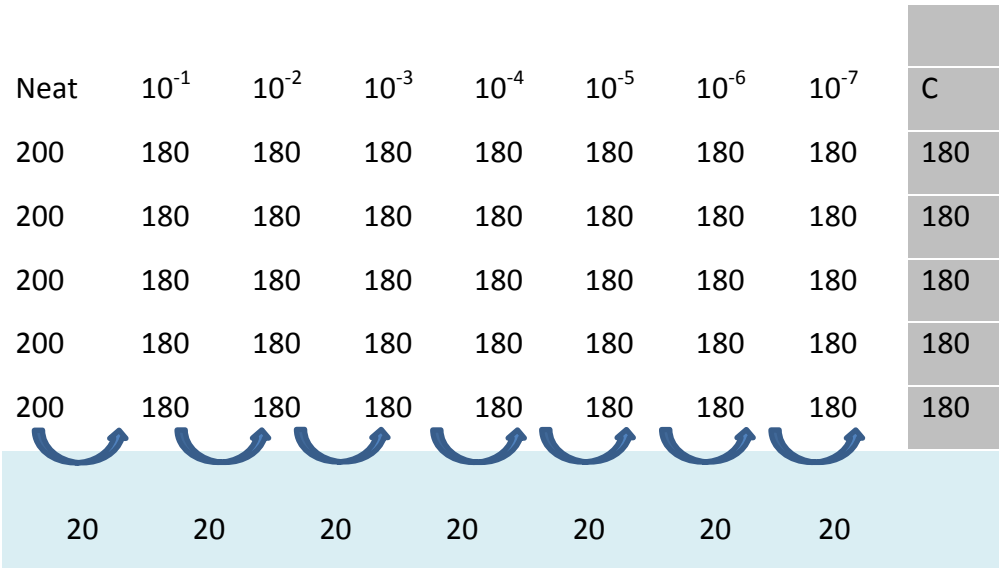


Figure 2.1 Process of titration in 96 well tray.
The process begins by pipetting 20µl of the sample under test (neat sample) into 180µl of media and then diluting across a 96 well tray as shown in the figure. 100µl of sample is then taken from the tray and used to infect cells in a 96 well tray seeded with McCoy cells 24 hours pre-infection. All figures are in µl.

2.9 FIXING CELLS.

The media was removed and the cells were washed twice with PBS. Cold methanol [100% (v/v)]was then added to the well and kept at -20°C for 20 minutes. The cells were washed twice again with PBS after which fixed cells were kept at 4°C with PBS covering the fixed monolayer in the well.

2.10 X-GAL STAINING.

Primary antibody (antibody 29) was added to the fixed cells at a dilution of 1/50. The PBS in the wells was replaced by 100µl of primary antibody. The tray was then left at 4°C overnight. The cells were then washed three times with PBS and 100µl conjugate (Anti-mouse IgG, Heavy and Light chain specific (goat) β-galactosidase conjugate) was then added at a concentration of 1/1000. The plate was then left at 37°C for one hour. Subsequently, the wells were washed three times for 2 minutes with PBS. 500µl of X-gal staining solution (see below for components) was then added to 10µl of X-GAL stock (see below for components) and 100µl of this solution was then added to the well and this was left at 37°C for 4 hours. Blue-stained inclusions were then visualised under a light microscope.

X-GAL staining solution
0.016g Potassium Ferricyanide
0.021g Potassium Ferrocyanide
0.04g Magnesium Chloride 6-hydrate
9.8ml PBS
Filter with 0.2µm filter

X-Gal Stock
0.05g X-GAL
1ml DMF
This solution is stored away from light

2.11 ISOLATION OF NUCLEIC ACIDS.

2.11.1 DNA extraction using Promega wizard genomic DNA purification kit.

DNA extraction was carried out using the Promega wizard genomic DNA purification kit (Promega, Southampton, UK). Extraction was carried out according to manufacturer's recommendations with some minor amendments made to the protocol. Samples were thawed and pelleted by centrifugation at 11,000 x g for five minutes and then resuspended in 480µl 50mM EDTA. 20µl of proteinase K was then added at 1mg/ml and subsequently incubated at 60°C for 1 hour. 600µl of nuclei lysis solution was then added until cells were resuspended and incubated at 80°C for five minutes to lyse the cells then left to cool at room temperature. 3µl of RNase solution was added

to the cell lysate and the tube inverted 2-5 times to mix. This was then incubated at 37°C for 20 mins to lyse the cells and left to cool at room temperature. 200µl of protein precipitate solution was added to the RNase treated cell lysate and the solution vortexed vigorously at 11,000xg for 20 secs to mix. The samples were incubated on ice for five min and 600µl of Isopropanol was added to two 1.5ml plastic tubes (Eppendorf) followed by the addition of 600µl of supernatant containing the DNA to each tube. The contents were gently mixed by inversion until the thread like strands of DNA formed a visible mass and subsequently were centrifuged at 11,000xg, for 2 mins. The supernatant was discarded and the tube drained on clean absorbent paper. 600µl of 70% (v/v) ethanol (kept at room temperature) was added and the tube gently inverted several times to wash the DNA pellet and centrifuged at 11,000xg for two minutes. The tube was again dried on clean absorbent paper and the pellet allowed to air dry for 10-15min. 50µl of DNA was thereafter added to the tube to rehydrate the DNA and then incubated at 65°C for one hour. Periodically the solution was mixed by gently tapping the tube. DNA samples were then stored at 4°C.

2.11.2 DNA extraction using Nucleospin tissue kit.

Another kit used for DNA extraction was the Genomic DNA from Tissue-Nucleospin Tissue kit by Macherey-Nagel (Fisher Scientific, Loughborough, UK). The samples were spun at 11,000xg for five mins, the supernatant was discarded and the 180µl lysis buffer (T1) and 25µl of proteinase K was added to the pellet and vortexed. The solution was then left overnight in a 56°C incubator. The following day an RNase digest was performed by adding 2µl of RNase A (20mg/ml) to the solution and incubating at 56°C for five mins after vortexing. Subsequently, 200µl of binding buffer (B3) was added to the solution, vortexed and incubated for 10mins at 70°C. 200µl of ethanol (96-100%) was added to the solution and vortexed vigorously before being transferred to a NucleoSpin Tissue Column and centrifuged for one min at 11,000xg. The filtrate was then discarded. The silica membrane was washed in two steps. For the first wash, 500µl of buffer BW was added to the membrane and then the column was centrifuged at 11,000x g for one min. The second wash involved addition of 600µl of buffer B5 and again this was centrifuged at 11,000xg for one minute. After both wash steps the flow-through was discarded. The silica was then dried by centrifugation at 11,000xg for two minutes. Lastly the DNA was eluted in a buffer (BE) which contained 5mM Tris/HCl and then used in upstream analysis of the DNA.

2.11.3 DNA extraction using QIAamp Viral RNA.

560µl of prepared Buffer AVL containing carrier RNA was pipetted into a 1.5ml micro-centrifuge tube. 140µl of urine or transport media from which the swab was placed was added to the Buffer AVL-carrier RNA in the micro-centrifuge tube. The contents were then mixed by pulse-vortexing for 15 s and incubated at room temperature (15–25°C) for 10 min. The tube was briefly centrifuged to remove drops from the inside of the lid. To the sample, 560µl of ethanol (96–100%)

was added, and mixed by pulse-vortexing for 15 s. After mixing, the tube was briefly centrifuged to remove drops from inside the lid. Carefully, 630µl of the solution was added to the QIAamp Mini column (in a 2 ml collection tube) without wetting the rim. The cap was closed, and the tube centrifuged at 6000 x *g* (8000 rpm) for one min. The QIAamp Mini column was placed into a clean 2 ml collection tube, and the tube containing the filtrate was discarded. The QIAamp Mini column was carefully opened and the previous step was repeated until all the solution had passed through the column. The QIAamp Mini column was carefully opened and 500µl of Buffer AW1 was added to the tube. The cap was closed, and the tube centrifuge at 6000 x *g* (8000 rpm) for one min. The QIAamp Mini column was placed into a clean 2ml collection tube, and the tube containing the filtrate was discarded. The QIAamp Mini column was carefully opened, and 500µl of Buffer AW2 was added. The cap was closed and the tube centrifuged at full speed (20,000 x *g*; 14,000 rpm) for three min. The QIAamp mini column was placed in a new 2ml collection tube and the old tube was discarded with the filtrate. The tube was then centrifuged at full speed for one min. The QIAamp Mini column was placed in a clean 1.5ml microcentrifuge tube and the old collection tube containing the filtrate was discarded. The QIAamp Mini column was carefully opened and 60µl of Buffer AVE equilibrated to room temperature was added to the tube. The cap was closed, and incubated at room temperature for one min. The tube was then centrifuged at 6000 x *g* (8000 rpm) for one min.

2.12 ANALYSIS OF NUCLEIC ACIDS.

2.12.1 TAE Buffer (tris-acetate EDTA buffer).

50x TAE buffer (Promega) was used at a working dilution of 1x for horizontal submerged agarose gels as the electrophoresis buffer.

2.12.2 Agarose gel electrophoresis.

Agarose (Fisher, Loughborough, UK) gel electrophoresis was used for the separation, visualisation, quantification and characterisation of DNA fragments. The concentration of agarose needed was determined by the size of the DNA fragment being resolved and by following rules that are outlined in the table 2.2. SYBR safe intercalating DNA gel stain was used for visualisation of DNA in the agarose gels and is a highly sensitive stain that is the less hazardous alternative to ethidium bromide that can be used for visualisation of DNA by UV excitation. The concentration of Sybr® safe used to TAE buffer was 1/20000. The agarose was dissolved in 1xTAE by heating to 100°C in a microwave oven. The dissolved agarose was cooled then poured into a gel mold, and a plastic comb was placed in the liquid. The liquid agarose was then allowed to set before being used. Once set the comb was removed carefully and the gel was submerged in 1xTAE. DNA samples containing Orange G gel-loading buffer were then loaded into the wells under the buffer. 5µl of a

DNA ladder (Bioline) (figure 2.2) was used as a marker for both the quantity and the size of the analysed DNA. Gels were electrophoresed at 90 volts for small gels and 100 volts for large gels until the Orange G dye front within the sample loading buffer had migrated to 1 cm from the end of the gel.

2.12.3 Visualisation of DNA bands.

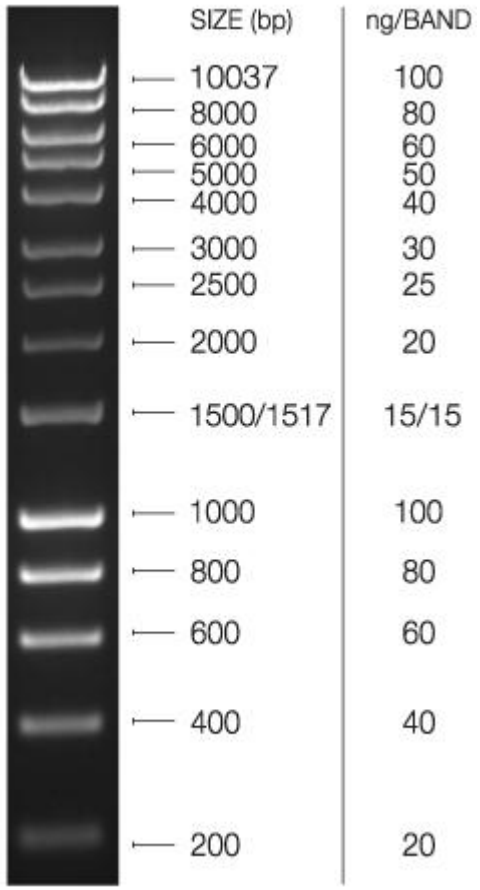
The bands on the gel were visualised by ultra violet illumination using a transilluminator (Ultra-violet products Inc) at 320nm. Photographic images of the gel were taken using a Polaroid 667 film.

2.12.4 Estimation of DNA size and concentration.

The sizes of DNA fragments were estimated by comparing their relative mobility through the gel matrix against DNA fragments of a known size. The DNA samples were routinely comparatively analysed using a DNA ladder; hyperladder I or hyperladder II (Bioline) (figure 2.2).

Agarose (w/v)	Size of DNA	
	Lower range (kb)	Upper range (kb)
0.6%	1	20
0.7%	0.8	10
0.9%	0.5	7
1.2%	0.4	6
1.5%	0.2	4
2.0%	0.1	3

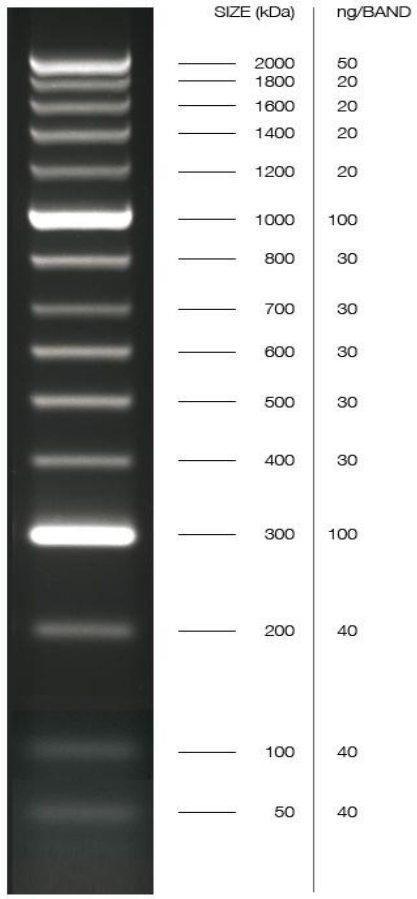
Table 2.2 Concentration of agarose needed to visualise bands in a 1 x TAE electrophoresis gel.



Hyperladder I:

1% agarose gel

5µl used per lane



Hyperladder II:

1.5% agarose gel

5µl used per lane

Figure 2.2 Ladders used to estimate size of bands. Hyperladder 1 (left), Hyperladder 2 (right).

2.12.5 Amplification of DNA by polymerase chain reaction (PCR).

VNTR and *omp*A sequences were amplified using PCR. *omp*A was amplified using primers P1F and CT5 (Table 2.3), whilst the three VNTR regions were amplified using primers outlined in the Pedersen *et al* paper (Pedersen *et al.* 2008); CT1335F and CT13335R, CT1299F and CT1299R and CT1291F and CT1291R (Table 2.3). PCR was conducted in 20µl reactions. The reaction components for the PCR included 10µl of Phusion™ high fidelity DNA polymerase mastermix (New England Biolabs UK, Hitchin, UK), which contains a high fidelity Taq DNA polymerase which ensures a relatively low error rate. The mastermix possesses 5→3 polymerase activity and 3 →5 primer exonuclease activity. The mastermix itself is a 2X mastermix containing Phusion DNA polymerase, nucleotides and optimised reaction buffer which includes $MgCl_2$. Therefore only templates and primers were added to the reaction. PCR primers are shown in table 2.3. 4µl of both the forward and reverse primers were added to the mastermix and 2µl of template DNA.

Amplification of *omp*A was carried out using P1F and CT5 primers or PCTM3 and NR1 to produce full length amplicons. If amplification of the full length PCR product could not be achieved then the primers OMPSeqF and OMPSeqR were used for amplification of *omp*A, which produced a shorter PCR product of 718bp that encompasses variable domains I, II and III.

2.12.6 PCR conditions.

PCR conditions for the three VNTR sequences and *omp*A were as follows:

1. Denaturation for 10 seconds at 98°C,
2. Thermal cycling: 35 cycles of 2 seconds at 98°C, 5 seconds at either 55°C for CT1335, 59°C for both CT1291 and *omp*A and 64°C for CT1299, and 10 seconds at 72°C.
3. Elongation step for 1 minute at 72°C.
4. The temperature was then maintained at 10°C.

The amplification was conducted using a Veriti™ Thermal Cycler (Applied Biosystems, UK).

Primer name	Primer sequence (5' – 3')	Nucleotide position[a]	Amplicon Size (bp)	Reference
CT1335 Forward	TCATAAAAGTTAAATGAAGAGGGACT	737,225–737,250	153	(Pedersen *et al.*, 2008)
CT1335 Reverse	TAATCTTGGCTGGGGATTCA	737,377–737,358		(Pedersen *et al.*, 2008)
CT1299 Forward	TTGTGTAAAGAGGGTCTATCTCCA	291,758–291,781	188	(Pedersen *et al.*, 2008)
CT1299 Reverse	AAGTCCACGTTGTCATTGTACG	29,1945–291,924		(Pedersen *et al.*, 2008)
CT1291 Forward	GCCAAGAAAAACATGCTGGT	195,536–195,555	225	(Pedersen *et al.*, 2008)
CT1291 Reverse	AGGATATTTCCCTCAGTTATTCG	195,760–195,738		(Pedersen *et al.*, 2008)
CT1335 Forward 2	AAAGCGTCCTCTGGAAGGG	737198..737198	208	This study
CT1335 Reverse 2	CCTTCTCCTAACAACTTACGC	737385..737405		This Study
CT1299 Forward 2	ATCGCTTAAGATTCTCGGAGG	291654..291674	342	This Study
CT1299 Reverse 2	AGGTTCTAGCTGAGCATGGG	291976..291995		This Study
CT1291 Forward 2	ATATAAAAAGAACCGTTGTTTCTG	291654..291674	329	This Study
CT1291 Reverse 2	CATCTTAGACATGCTCCGGC	291976..291976		
P1F Forward	ATGAAAAAACTCTTGAAATCGG	780039 - 780060	1167	(Frost *et al.*, 1991)
CT5R Reverse	ATGAGAGAGCTGCTCACGTAAAT	778897 - 778919		(Rodriguez *et al.*, 1991)
PCTM3 Forward	TCCTTGCAAGCTCTGCCTGTGGGGAATCCT	779,977–780,006	1,1019	(Lan *et al.*, 1994)
NR1 Reverse	CCGCAAGATTTTCTAGATTTC	778,988–779,008		(Lan *et al.*, 1994)
OMPSeqF	GGTGTGACGCTATCAGCATGC	779,880–779,900	718	(Wang *et al.*, 2011b)
OMPSeqR	GACCATTTAACTCCAATGTA	779,183–779,202		(Wang *et al.*, 2011b)

Table 2.3 Primers used for PCR of MLVA-*omp*A markers.
[a]According to D/UW-3/CX, accession number NC 000117.

2.12.7 Determinating the concentration of nucleic acids using the NanoDrop 1000™ spectrophotometer.

The concentration of nucleic acids was determined using the Nanodrop 1000™ spectrophotometer. Once the software for the machine was opened on the computer attached to the NanoDrop spectrphotometer, the nucleic acid application module was selected. The lower and upper pedestals of the nanodrop machine were cleaned with a lens cleaning tissue. To obtain a blank measurement, 1.5µl of H_2O was added to the lens and then the word Blank was selected. Once the blank measurement had been calculated the lens was cleaned again and then 1.5µl of the sample to be tested was added to the lens and the measure button was selected to measure the concentration of DNA in the sample. The purity of DNA was measured by determining the ratio of the absorbance at two wavelengths – 260nm and 280nm. A ratio of ~1.8 was accepted as pure for DNA.

2.12.8 Purification of PCR products using Wizard SV gel and PCR clean-up system (Promega).

Purification of the PCR products was done before sequencing commenced, to remove contaminating nucleotides and primers. The PCR products were processed initially using an equal volume of membrane binding solution (provided in kit) to the PCR reaction. The PCR product was then added to the mini column assembly, incubated at room temperature for one minute and then centrifuged at 16,000 x g for one minute. Flow-through was then discarded. 700µl of membrane solution (provided in kit) containing ethanol was added to the membrane and centrifuged at 16,000 x g for one minute. This step was then repeated with 500µl of membrane wash solution and centrifuged for a further two minutes. The collection tube was emptied and the column assembly was centrifuged for a further two minutes to allow evaporation of any residual ethanol. The mini column was then transferred to a clean 1.5ml microcentrifuge tube and 50µl of Nuclease-Free Water was added and incubated for one minute before being centrifuged at 16,000 x g for one minute. The DNA was subsequently stored at -20°C.

2.12.9 DNA Sequencing.

Sequencing of PCR products was carried out by the commercial company Geneservice (Nottingham, UK), who offer a Sanger sequencing service (Sanger *et al*, 1977). The DNA posted to Geneservice was required to be at a concentration of 1ng/µl/100bp for PCR products. The website for the company can be found at http://www.lifesciences.sourcebioscience.com/. The way in which their sequencing method works is as follows:

1) The DNA is denatured using heat into single strands of DNA.
2) The primer is attached to the DNA (either the primer or one of the nucleotides is fluorescently labelled so that the final product can be detected on a gel).

3) DNA polymerase then binds to the primer, and polymerises a new strand of DNA by incorporating free nucleotides that are complementary to the target DNA.
4) The DNA polymerase continues to extend the strand until it incorporates a fluorescently labelled nucleotide. The fluorescently labelled nucleotides are chemically altered so that they terminate the DNA strand and the enzyme falls away.
5) This is repeated many times, generating a large number of different lengths that end in fluorescently labelled bases.
6) The fragments are floating freely in one of many tiny wells in a plate.
7) The samples are transferred into thin glass capillaries where an electric charge starts the negatively charged DNA molecules moving through a gel matrix
8) Longer DNA fragments move slower than the shorter fragments through the gel. The DNA fragments are therefore sorted by size from shortest to longest and a laser at the end of the capillary is used to excite the final fluorescent base, which is recorded as a coloured peak.
9) Each coloured base therefore represents the final base of each strand of DNA that was made in the sequencing reaction, ordered from the shortest to the longest.
10) The sequence of the original piece of DNA can therefore be decoded.

The sequencing primers sent to the sequencing company with the PCR products were as follow: CT1335InR for CT1335, CT1299InR for CT1299, CT1291InF for CT1291 and NR1 and OMPSeq2R for *omp*A (Table 2.4).

2.13 WHOLE GENOME SEQUENCING.

Whole genome sequencing was carried out by colleagues at the Sanger Centre in Cambridge using Illumina technology (illumina, 2012). The Illumina sequencing technology takes place in three stages, firstly the library is prepared, then the cluster is generated and finally sequencing takes place (figure 2.3).

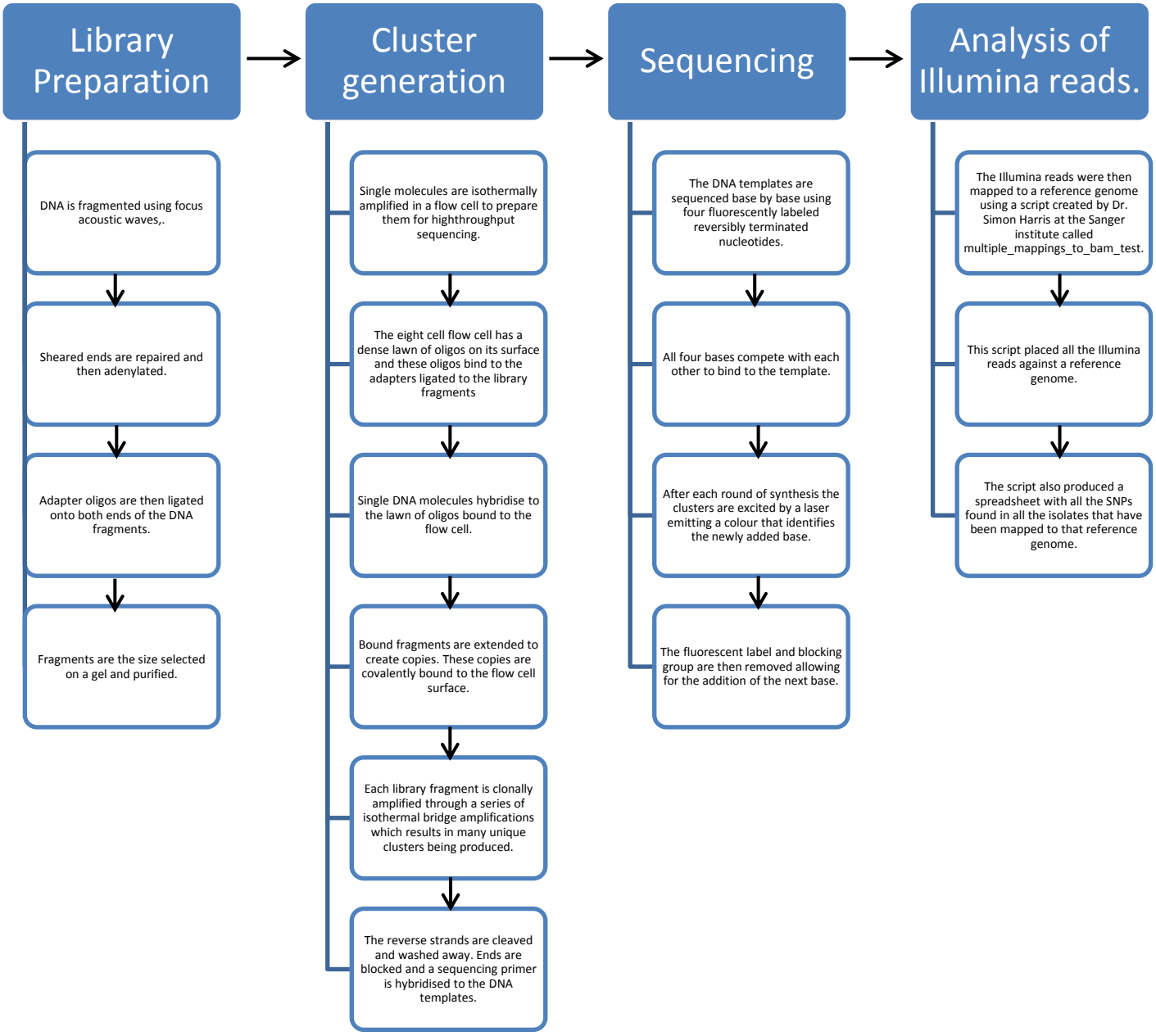


Figure 2.3 Stages of the whole genome sequencing process. Starting with library preparation to analysis of Illumina reads.

Primer name	Primer sequence (5' – 3')	Locations[a]	Region Sequenced	References
CT1335inR	GGATTCAACGATGATTAAGG	737,345 – 737,364	CT1335	(Wang *et al.*, 2011b)
CT1299inR	ACGAATCCTCTAAGTACGG	291,908 – 291,926	CT1299	(Wang *et al.*, 2011b)
CT1291inF	TACAAAAGTGTTGTGATAATTC	195,559 – 195,580	CT1291	(Wang *et al.*, 2011b)
NRI	CCGCAAGATTTTCTAGATTTC	778,988 – 779,008	*Omp*A	(Lan *et al.*, 1994)
OMPSeq2R	TATTGGAAAGAAGCICCTAA	779,345 – 737,364	*Omp*A	(Wang *et al.*, 2011b)

Table 2.4 Primers used for sequencing of MLVA-*omp*A markers.
[a]according to D/UW-3/CX, accession number NC 000117.

2.14 PHYLOGENTIC TREES.

Phylogenetic trees were created using the Randomised accelerated Maximum Likelihood (RAxML) script created by Dr Simon Harris at the Sanger institute. Files produced using this script were then viewed using the Fig Tree software.

2.15 DATA ANALYSIS.

2.15.1 DNA sequence analysis: Nucleotide BLAST search tool used to search for similar genomes that have been deposited online.

The basic local alignment search tool (BLAST) was used to find regions of similarity between sequences. The FASTA files of the sequences that were to be compared to that on the NCBI database, which is a database of genomes that have been sequenced and are now widely available for public use. The searches where all optimised for highly similar sequences using the megablast feature. The search then returns various entries that are similar to that entered initially into the database and generates scores of similarity between different sequences that have been entered into the database of available genomes. This aids in identification of the most closely related strain to the strain that was submitted for analysis. The BLAST can be found using the following link: http://blast.ncbi.nlm.nih.gov/Blast.cgi.

2.15.2 DNA sequence analysis: Alignment and annotation using Lasergene.

2.15.2.1 Seqman.

Seqman was used for alignment of MLVA-*omp*A and MST regions after they had been sequenced. Comparison of sequences was then conducted and SNPs between several samples were identified using this software.

2.15.2.2 Seqbuilder.

Seqbuilder was used to annotate key regions of the *C. trachomatis* genome such as the MLVA-3 (Pedersen *et al.*, 2008) and MLVA-5 (Peuchant *et al.*, 2012) and MST (Klint *et al.*, 2007) regions and present these as a circular genome.

2.15.3 Whole genome analysis using Artemis.

Artemis is a free DNA viewer and annotation tool and can be used for analysis of both prokaryotic and eukaryotic genomes. Artemis allows the user to view the sequence file generated from whole genome sequencing and manipulate the sequences. For instance genomes were annotated using this program, SNPs between strains were identified and the various reads generated by whole genome sequencing were also viewed and analysed. Artemis can be downloaded free of charge from http://www.sanger.ac.uk/resources/software/artemis/.

2.15.4 Comparison of whole genomes using ACT.

The Artemis Comparison Tool (ACT) was used for comparison of the various strains. Conserved unique regions in the *C. trachomatis* genome were identified using this software, which is freely available from the following website: http://www.sanger.ac.uk/resources/software/act/.

2.16 STATISTICS.

Statistical advice was sought from Mr. Scott Harris, a medical statistician at Southampton General Hospital. Differences between groups were tested by using the Pearson chi squared test and the Fishers exact test, when it was appropriate to do so. For continuous data, Mann-Whitney U and Kruskal-Wallis tests were used. A P value of ≤0.05 was considered statistically significant. Analysis was performed with the SPSS package, version 20.0.

2.16.1 Simpsons index of diversity.

The discriminatory power of the MLVA-*omp*A typing system was calculated using the Simpsons index of diversity as described by Hunter and Gaston (Stephens *et al.*, 2001). The discriminatory power was calculated using the following formula

$$D = 1 - \frac{1}{N(N-1)} \sum_{J=1}^{S} x_j(x_j - 1)$$

Equation 2-1 Formula used to define Simpson's index of diversity (D)

N is the number of unrelated strains tested, s the number of different types and x_j the number of strains belonging to the j^{th} type. D is a value between 0 and 1, were a value of 0 is equal to no diversity and a value of 1.0 means that there is complete diversity.

2.17 BIONUMERICS.

Minimum spanning trees were generated with Bionumerics 6.5 using the generated MLVA profiles.

2.18 ETHICS.

All applications except the ERGO application were completed on the IRAS website and submitted individually to the relevant bodies.

2.18.1 Ethics Research Governance Online (Ergo).

To enable the University of Southampton to grant the research team a letter on insurance and sponsorship for the study, the application for the Epidemiology of Chlamydia in Southampton had to go through an Ethics and Research Governance Online (ERGO) application which is a centralised online system designed to facilitate the process of gaining ethical, governance and insurance approval for research studies at the University of Southampton. As the Southampton study involves human participants it had to be registered on the University's electronic document management system.

2.18.2 NRES.

The National Research Ethics Service (NRES) is a service that enables and supports ethical research to maximise the benefits of research in the UK. NRES provides an efficient and robust ethics review service. Studies are reviewed by a Research Ethics Committee (REC) who work to safeguard the rights, safety, dignity and well-being of people participating in research in the National Health Service (NHS).

2.18.3 NHS/HSC RESEARCH AND DEVELOPMENT OFFICES- (R AND D APPROVAL).

The Coordinated System for gaining NHS Permission (CSP) is a standard process for adoption onto the National Institutes for Health Research (NIHR) Clinical Research Network (CRN) portfolio of Studies in order to access NIHR CRN Support and funding. To enable research to be conducted in the NHS, researchers must obtain NHS management permission for each research site. To gain permission an R and D application was completed on the IRAS website. A site specific form for each site involved in the research was also filled out. This form was submitted at the same time as the REC application as approval from the NHS can be a lengthy process.

2.18.4 NIHR CRN PORTFOLIO APPLICATION.

NIHR CRN stands for The National Institute for Health Research Clinical Research Network. The NIHR CRN portfolio is a database of high quality research studies that are eligible for support from the NIHR Clinical Research Network in England. Studies that are included on the portfolio have access to a local network of dedicated skilled research support staff, support to ensure that the study can be successfully undertaken in the NHS and access to experienced Research Management and Governance staff that can give advice on various matters to do with the study.

CHAPTER 3 *CHLAMYDIA TRACHOMATIS* MARKER STABILITY STUDY.

Part of the work described in this chapter has been published - (Labiran *et al.*, 2012a)- (Appendix 1).

3.1 INTRODUCTION.

Following the availability of whole genome sequences for chlamydia, many genetic typing methods have been developed for *C. trachomatis*. The benefits of typing systems are wide-ranging; prevalence rates for specific strains can be determined, they can be used to reveal sexual networks and therefore to track strains and can also be used for evolutionary surveillance of clones (Pedersen *et al.*, 2008). There are several genetic typing schemes available for *C. trachomatis* (Yuan *et al.*, 1989, Klint *et al.*, 2007, Pannekoek *et al.*, 2008, Pedersen *et al.*, 2008, Dean *et al.*, 2009, Bom *et al.*, 2011, Peuchant *et al.*, 2012), each of which offer different levels of discrimination and can be used to answer various epidemiological questions. For instanc,e the Pannekoek (2008) and Dean (2009) MLST systems use housekeeping genes and are suitable for long term and global epidemiological studies as these methods are used to analyse variations that accumulate within the *C. trachomatis* genome very slowly in populations and are likely to be selectively neutral (Maiden *et al.*, 1998). On the other hand, the Klint *et al* (2007) and the Pedersen *et al* (2008) schemes are better suited for analysis of short-term epidemiology and for partner tracing purposes, as the *loci* chosen for analysis in these schemes are selected on the basis that they provide high levels of rapidly evolving variation within the *C. trachomatis* population (Maiden *et al.*, 1998).

In this current study, the stability of the markers in two typing schemes- a MLVA-*omp*A (Pedersen *et al.*, 2008) and an MST typing system (Klint *et al.*, 2007)-were determined for *C. trachomatis*. These two typing schemes were chosen for analysis in this study on the basis that they are different schemes that offer the highest level of discriminatory power of the four typing schemes available for *C. trachomatis* at the time (Peuchant *et al* (2012) was not yet published). The MLVA-*omp*A scheme was designed to be used for local epidemiological studies and for this, three regions consisting of short polymeric tracts were chosen along with *omp*A (Pedersen *et al.*, 2008). Whilst the second scheme, although initially named MLST, actually defies the rules of a conventional MLST system as it includes genes which are not housekeeping genes but are genes that code for hypothetical proteins with no none function. Additionally, this scheme designed by Klint *et al* includes the *omp*A gene which is under immune selection (Klint *et al.*, 2007). This scheme is now better described as a multi-sequence typing (MST) scheme (Wang *et al.*, 2011b). Six target regions including *omp*A are used in the MST typing scheme (Klint *et al.*, 2007).

This marker stability study was conducted to determine if the DNA sequences of the markers for these genetic typing methods remain stable for the isolates after their primary isolation in cell culture and to determine whether new mutations may occur and to identify when the sequence of a marker may change, if at all. The reason for conducting this study was because one of the main aims of my PhD was to study the epidemiology of chlamydia in Southampton (later developments also meant that the epidemiology of MSM Brighton was also studied). To study the epidemiology of chlamydia in Southampton and in Brighton the samples collected had to be assigned genotypes. To do this there were only two options available; to either develop a new typing scheme or to use a pre-existing typing scheme for the purpose of assigning genotypes to samples. SNP typing would be the ideal approach because it offers high levels of discrimination. However, the development of such a typing scheme would require the need for a larger set of *C. trachomatis* whole genome sequences to determine which SNPs are important and should be included. This data were not available at the start of the project. Therefore, the approach to use a pre-existing typing scheme for these studies was considered most appropriate. The two schemes available to study the local epidemiology of *C. trachomatis* were independently developed by Klint *et al* (2007) and Pedersen *et al* (2008). However, what was missing from previous studies that have used these typing schemes was a detailed analysis on the stability of the markers, which would aid in making a decision on which was best suited for future studies. This was important as the markers used in a typing scheme must remain stable over the study period (van Belkum *et al.*, 2007). If the markers are constantly changing then it will be impossible to gain a clear picture of the epidemiology of *C. trachomatis* in any study in which the typing schemes are used for and may result in every sample possessing a different type according to the markers in the typing scheme.

To assess the stability of the *C. trachomatis* MLVA-*omp*A markers, the stability of the markers after eight passages was assessed. The reason for choosing eight passages was so that the time period from infection to symptomology *in vivo* was mimicked (Black, 1997). It has been observed that if symptoms appear then these can occur from anything between seven to twenty-one days after infection (Black, 1997), so the isolates were taken through 24 days in cell culture to cover this time period. Additionally, the stability of both the MLVA-*omp*A and the MST markers were investigated after continuous passaging of the Swedish new variant in cell culture.

3.2 AIMS:

- To assess two different typing systems (MLVA-*omp*A and MST) available for typing *C. trachomatis* positive samples. In doing so the suitability of these typing systems as a high discriminatory typing system for typing *C. trachomatis* was determined.

- To evaluate the stability of the markers used in the MLVA-*omp*A typing system as described by Pedersen *et al* (2008) for *C. trachomatis* through adaptation to cell culture (Section 3.3.3).
- To assess the stability of the MLVA-*omp*A (Pedersen *et al.*, 2008) and MST (Klint *et al.*, 2007) markers through adaptation of *C. trachomatis* to cell culture following long term passaging (Section 3.3.4).
- To determine the stability of the whole genome sequence of the Swedish new variant strain following long term passaging (Section 3.4.4).

3.3 METHODS.

3.3.1 Experimental design.

Firstly, the stability of the *C. trachomatis* MLVA-*omp*A (Pedersen *et al.*, 2008) markers under multiple passages in the controlled conditions of cell culture was assessed. For this process, seven urogenital strains were selected with the *omp*A genotypes D-K from a previous study carried out in Southampton (Wang *et al.*, 2011b). Seven strains were selected as there were seven different *omp*A genotypes identified in the Wang *et al* (2011b) study, therefore one sample with each of the *omp*A genotypes was included in my study. The samples were selected on the basis of them having different *omp*A genotypes because the *omp*A gene has been widely used for many years to study the molecular epidemiology of *C. trachomatis* (Jurstrand *et al.*, 2001, Falk *et al.*, 2003, Lysen *et al.*, 2004, Osterlund *et al.*, 2005, Mejuto, 2013). Therefore, determination of the stability of different *omp*A genotypes was considered important and there may also be differences detected in the stability of other markers between different *omp*A genotypes.

The study design was to take the chlamydia through eight passages in cell culture. The MLVA-*omp*A markers were sequenced before passaging began and after eight passages in cell culture to determine if the sequence of the markers had changed after passaging (i.e. after adaptation to cell culture following origination in human host-cells). By following potential changes through many developmental cycles it was possible to establish the stability of the individual markers within a numerical framework of bacterial divisions and this in turn informed us of the usefulness of using such a typing system for short and long-term epidemiology.

Subsequently, a study was conducted over the period of a year where the Swedish new variant (nvCT) also known as Sweden 2 (as this was the name assigned to the isolate sent from colleagues in Malmo to Southampton for research purposes) was passaged 72 times in cell culture as 72 passages was the number of passages that could be completed in a year. This timescale was selected as any future studies that would be carried out for this thesis using the markers in these typing schemes would take place within the period of one year. Therefore, this study was carried

out to investigate whether the markers would remain stable over the period of any future studies. For this study both the MST (Klint *et al.*, 2007) and the MLVA-*omp*A markers were assessed for their stability. These two typing schemes were selected as they offered the highest resolution for typing *C. trachomatis* samples and were also designed to study local epidemiology. As one of the main aims of this thesis was to study the epidemiology of *C. trachomatis* in two cities, Brighton and Southampton, it was therefore important to determine the stability of these typing schemes for use in later studies. To determine the stability of the markers in these typing schemes isolates were sequenced at every twelfth passage up to and including passage 72. It was decided to sequence the markers at every twelfth passage to increase the probability of detecting mutations occurring during the year that may have reverted back to the original sequence by passage 72.

3.3.2 Samples.

The residual material from seven swabs collected from patients who had attended a GUM clinic in Southampton, Hampshire, in 2009 were selected as these swabs were available at the time the study was carried out and therefore a prospective study to determine the stability of the markers was considered unnecessary. The transport medium in which the swabs had been placed were obtained from the Health Protection Agency (HPA) laboratory at Southampton General Hospital (SGH) after routine analysis had been carried out. The transport medium from the samples were used initially in a previous study (Wang *et al.*, 2011b). The study by Wang *et al* (2011b) evaluated the MLVA-*omp*A typing system (Pedersen *et al.*, 2008) by amplifying the four genomic regions that make up the typing scheme in samples from a population of women living in Southampton using PCR. The corresponding DNA for each sample which had been extracted using the Tecan DNA extraction system (Tecan, Männedorf, Switzerland) was also collected after routine analysis at the HPA. Following the completion of the study by Wang *et al* (2011b) the extracts and swabs were stored at -70°C so that denaturation of the DNA would be minimised and also to ensure that the samples remained infectious in cell culture if they were to be used in future studies such as is described in this chapter. Seven samples from this study (Table 3.1) were chosen as the baseline to establish the stability of these markers in a lab cell culture environment where mouse cells were used to propagate the bacteria. The basis for choosing the samples as isolates for the study was that they had to be capable of at least two passages in cell culture and that they had to have different *omp*A types. It was important to include a range of *omp*A types in case there were differences in the stability of the markers between different *omp*A genotypes. This was also a guard against cross contamination with other chlamydia *omp*A genotypes and ensured that during passages the samples had not been cross contaminated.

For the second part of this study, an isolate of the Swedish new variant which had been sent to the Molecular Microbiology lab in Southampton from colleagues in Malmo, Sweden was assessed

for the stability of its MLVA-*omp*A and MST markers following prolonged passage in cell culture. The sample was obtained from the urethra of a male patient who had contracted the new variant in 2006. nvCT is a variant of *C. trachomatis* which was first described in 2006 (Ripa and Nilsson, 2007). This variant evaded several of the commercial molecular diagnostic tests available at the time, which were designed to detect the presence of a specific plasmid sequence. However, a deletion of 377bp in coding sequence one (CDS1) in the plasmid DNA led to many tests producing negative results (Ripa and Nilsson, 2007). As the Swedish new variant was a mutable strain, it was considered an ideal strain to study marker stability.

***omp*A type** (determined by comparison to NCBI database sequences using the BLAST search)	**Repeat sequence for each VNTR loci before culture**			**Variant number**[a]
	CT1335	**CT1299**	**CT1291**	
D/IC-CAL8	13T7A	11C	8C	8, 5, 2
D/UW-3	13T7A	11C	8C	8, 5, 2
E/Bour	13T7A	11C	8C	8, 5, 2
F/IC-CAL3	13T7A	12C	8C	8, 6, 2
G/392	10T8A	10CT3C	10C	3, 4b[b], 4
Ia/870	10T8A	11C	11C	3, 5, 5
K/UW-31	10T8A	3C2T6C	10C	3, 9[c], 4

Table 3.1 Isolates included in part one of the stability study.
All of the seven isolates were collected from the cervix of women who attended the GUM clinic in Southampton. The table shows the corresponding *omp*A NCBI database type for each isolate, the repeat sequences for the three VNTR loci and their associated variant number before culture.
[a] Based on the numbering system of Pedersen *et al* (2008).
[b] VNTR region modified by Wang *et al* (2011b).
[c] New VNTR region found by Wang *et al* (2011b).

3.3.3 Stability of the MLVA-*omp*A markers through adaptation of *C. trachomatis* to cell culture

The first part of this study was to investigate if there was an effect on the sequence of the markers by obtaining *C. trachomatis* from a human source and transferring the bacteria to mouse cells for isolation. The new environment in which the bacteria was being made to replicate could cause a change in the sequence of markers during adaptation to *in vitro* cell culture and therefore lead to marker instability which may not be the case in an *in vivo* situation. It is known from previous studies that chlamydia can grow well in McCoy cells (Croy *et al.*, 1975, Rota, 1977); however, what was not known is whether the shock of the changing environment affects key sequences in regions which are used for typing *C. trachomatis*.

3.3.3.1 Methods used to determine the stability of the MLVA-*omp*A markers through adaptation of *C. trachomatis* to cell culture.

Passaging of the seven isolates initially began with them being grown in a single well of a 24 well tray, then expanded to a single well of a six well tray and then into flasks (T25 and T75) (Figure 3.1). The passaging process is outlined in section 2.6. After 48 hours had elapsed in cell culture, the cells were harvested and passaged on, the level of infection observed before harvesting determined the size of the tray or flask the samples were put into next which depended on the area of cells that was required for the next passage. This was carried out for eight passages, at this point all isolates were growing well in T75 flasks. After eight passages, DNA extraction was carried out using the Promega wizard genomic DNA purification kit (Promega, UK)(Section 2.11.1). To confirm that the DNA had been extracted correctly, the products produced from the DNA extraction were separated by agarose gel electrophoresis. PCR was then carried out to amplify the MLVA-*omp*A regions and a clean-up of the PCR product was performed using a PCR clean up kit (Promega, Southampton, UK) (Section2.12.8). All PCRs were conducted using a high fidelity DNA polymerase to ensure that the sequence data obtained were of the best quality possible. The PCR clean-up products were then checked using agarose gel electrophoresis to ensure that the regions had been amplified correctly and that there was no contamination during PCR preparation (Figure 3.2). Subsequently the concentration of the DNA was determined using a nanodrop machine (Thermo Scientific, Southampton, UK). Once the PCR product had been checked and the product had been cleaned using a PCR clean up kit (Section 2.12.7), the product was sent for sequencing at a concentration of 1ng/µl/100bp following the requirements outlined by the sequencing company (Source Bioscience, Nottingham, UK). When the sequencing data had been received the marker sequences were aligned with the cognate sequence from before passaging began (Section 2.12.9). Alignments were performed using seqman (Section 2.15.2.1).

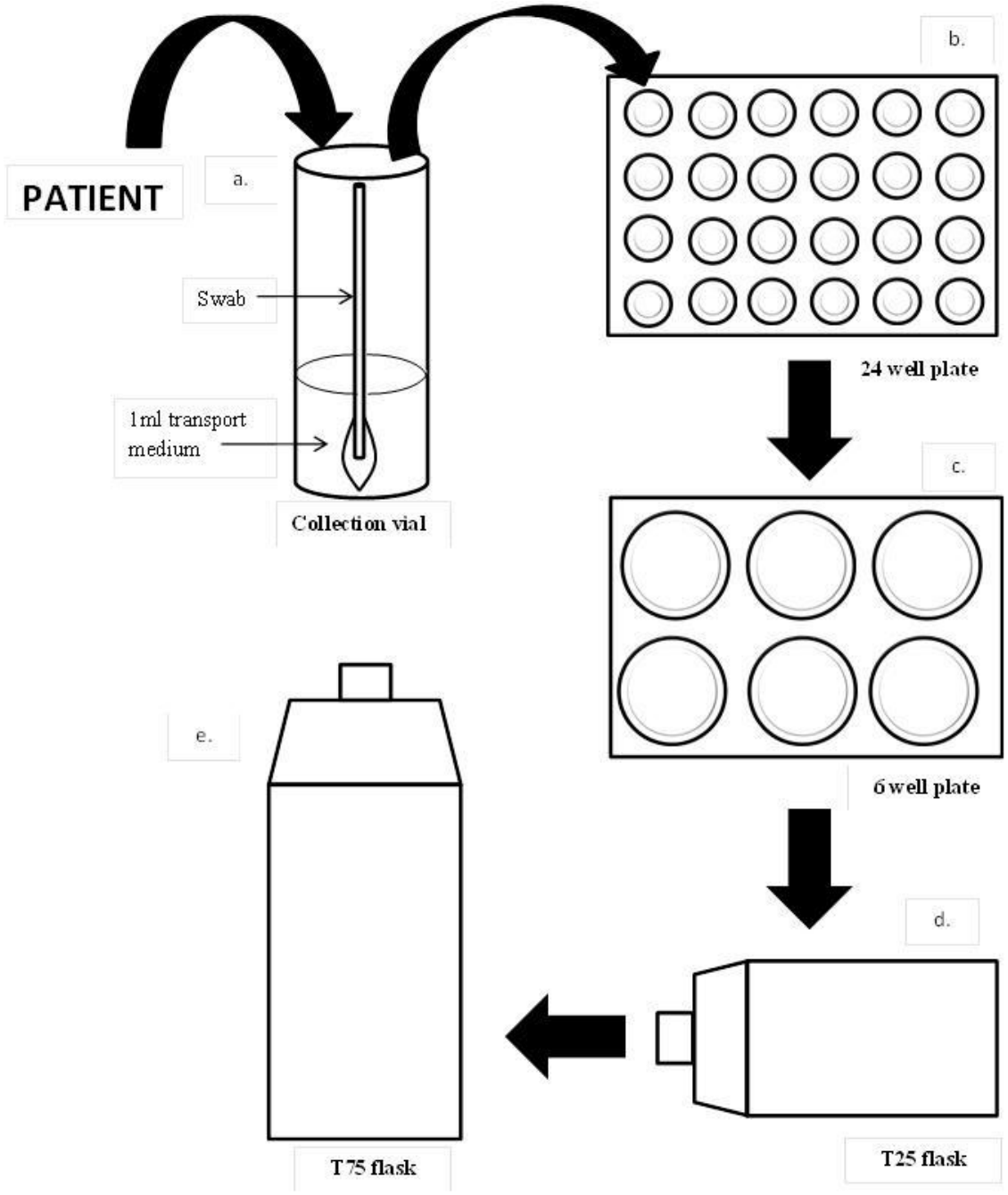


Figure 3.1 Generalisation of the passaging process from patient swab to T75 tissue culture flasks.

The swab was placed in transport medium (a) and this was briefly whirl mixed. 1ml of the transport medium was then taken and put into a single well in a 24 well plate (b). Once inclusions were present in ≥90% of the cells in the well the next passage was then continued in a single well of a six well plate (c). Again once ≥90% of the cells had been infected in a single well of a six well plate the next passage took place in a T25 (d) and was then passaged into a T75 flask (e).

3.3.4 Stability of the MLVA-*omp*A and the MST markers after long term passage in cell culture.

The Swedish new variant (serotype E with a 377bp deletion in the cryptic plasmid) isolated in Sweden was taken through 72 passages in cell culture. It was important to determine the stability of the MLVA-*omp*A and MST typing systems in the Swedish new variant as there is a possibility that this strain could have different properties in cell culture over the period of a year. There was also greater possibility that the MLVA-*omp*A and MST markers may have been unstable in this mutant because the strain had already mutated and could therefore be prone to further changes in the chromosome. The isolate was initially passaged nine times to get it to a level of infection where we were convinced that it would continue through the 72 passages and that this study was therefore worth continuing. After the initial nine passages there was enough inoculum to infect a T75 flask so that 100% of the cells were infected and therefore all the remaining passages were conducted in a T75 flask. Before the passaging process began, the variant number for the Swedish new variant was 8,8,1 for the three VNTR loci as the samples had the following sequences for the three loci; CT1335=13T7A, CT1299=14C, CT1291= 6C and was typed as having an *omp*A genotype E. Whilst the MST profile for the new variant before passaging began was 21-19-1-2-1 for the five MST markers.

3.3.4.1 Methods used to determine the MLVA-ompA and the MST marker stability in the *C. trachomatis* Swedish new variant after long term passaging.

After passaging of the Swedish new variant had been completed (Section 2.6), DNA was extracted from samples from every twelfth passage using a DNA extraction kit (section 2.11.1). The presence of DNA in the sample was checked by agarose gel electrophoresis. MLVA-*omp*A typing was carried out in the Molecular Microbiology department, Southampton. Whilst the MST was carried out by colleagues in Uppsala (Björn Herrmann and Linus Christerson, Section of Clinical Bacteriology, Department of Medical sciences).

For typing of the markers that make up the the MLVA-*omp*A scheme PCRs were done to amplify the three VNTR regions defined by Pedersen *et al* (2008) and also to amplify the *omp*A gene. After the PCR, amplification products were checked on an agarose gel to determine that the correct products had been amplified. PCR products were sequenced by a sequencing company as described in section 3.3.3.1. Sequences for every twelfth passage were determined up to and including passage 72 (passage 0, 12, 24, 36, 48, 60 and 72). Upon receipt of the sequencing data sequences received were aligned with the sequence data already obtained from before the isolates had been passaged using Seqman (an alignment software) (Section 2.10.2). Figure 3.3

shows the passage process and how the stability of the markers was determined using PCR and sequencing techniques.

To assign types to the MST markers a DNA sample of the isolated chlamydia at passage 72 of the Swedish new variant was sent from Southampton to Uppsala as part of a batch of samples that were sent to Uppsala and given random codes to prevent sequence analysis bias. MST was carried out as previously described (Klint *et al.,* 2007) with minor changes to the protocol. The *pbp*b and CT058 region were amplified using nested PCR as these regions are over 1000 base pairs in length. The use of two sets of primers for each of these regions was necessary to obtain accurate sequence reads.

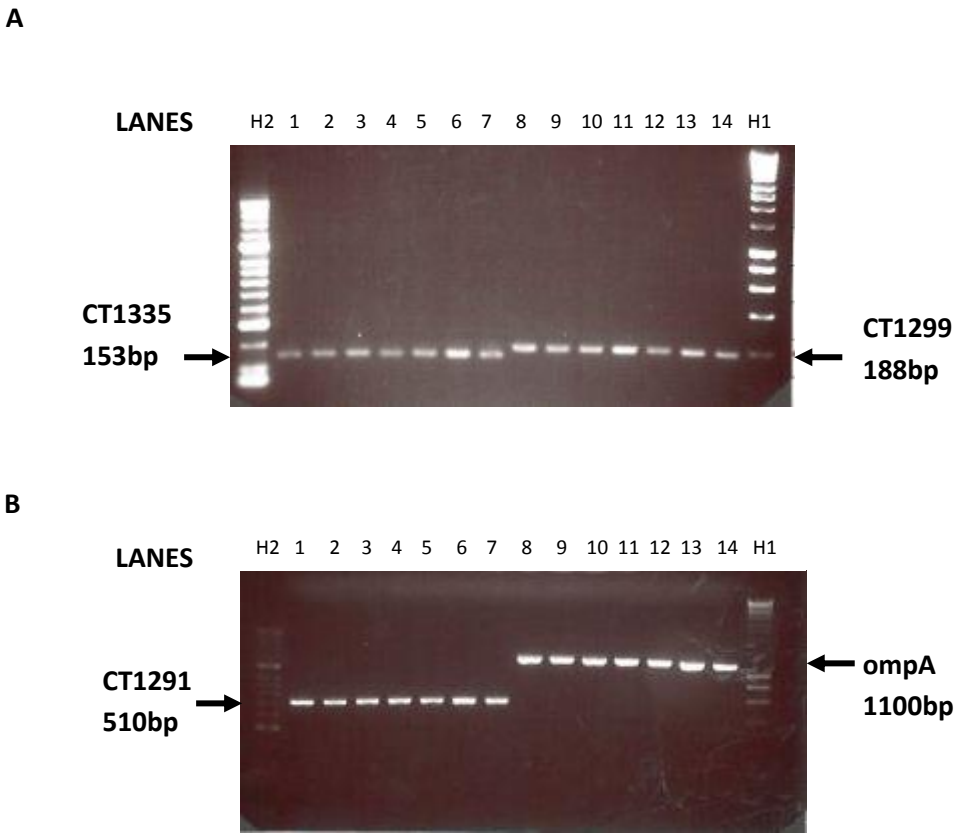


Figure 3.2 A 1.5% (w/v) agarose gel showing the PCR products after amplification of the MLVA-*omp*A markers, CT1335, CT1299, CT1291 and *omp*A in the *C. trachomatis* Sweden new variant.
The four regions were amplified in DNA extracted from every twelfth passage up to and including passage 72. The gel shows the PCR products before the passaging process began, passage 0, 12, 24, 36, 48, 60 and 72. Lanes 1-7 in gel A show the products from amplification of CT1335 in DNA from passage 0, 12, 24, 36, 48, 60 and 72. Lanes 8-14 in gel A show the products from amplification of CT1299 in DNA from passage 0, 12, 24, 36, 48, 60 and 72. Lanes 1-7 in gel B show the products from amplification of CT1291 in DNA from passage 0, 12,24, 36, 48, 60 and 72. Lanes 8-14 in gel B show the products from amplification of *omp*A in DNA from passage 0, 12, 24, 36, 48, 60 and 72. H1 and H2 represent hyperladder 1 and hyperladder 2 respectively (Figure 2.2).

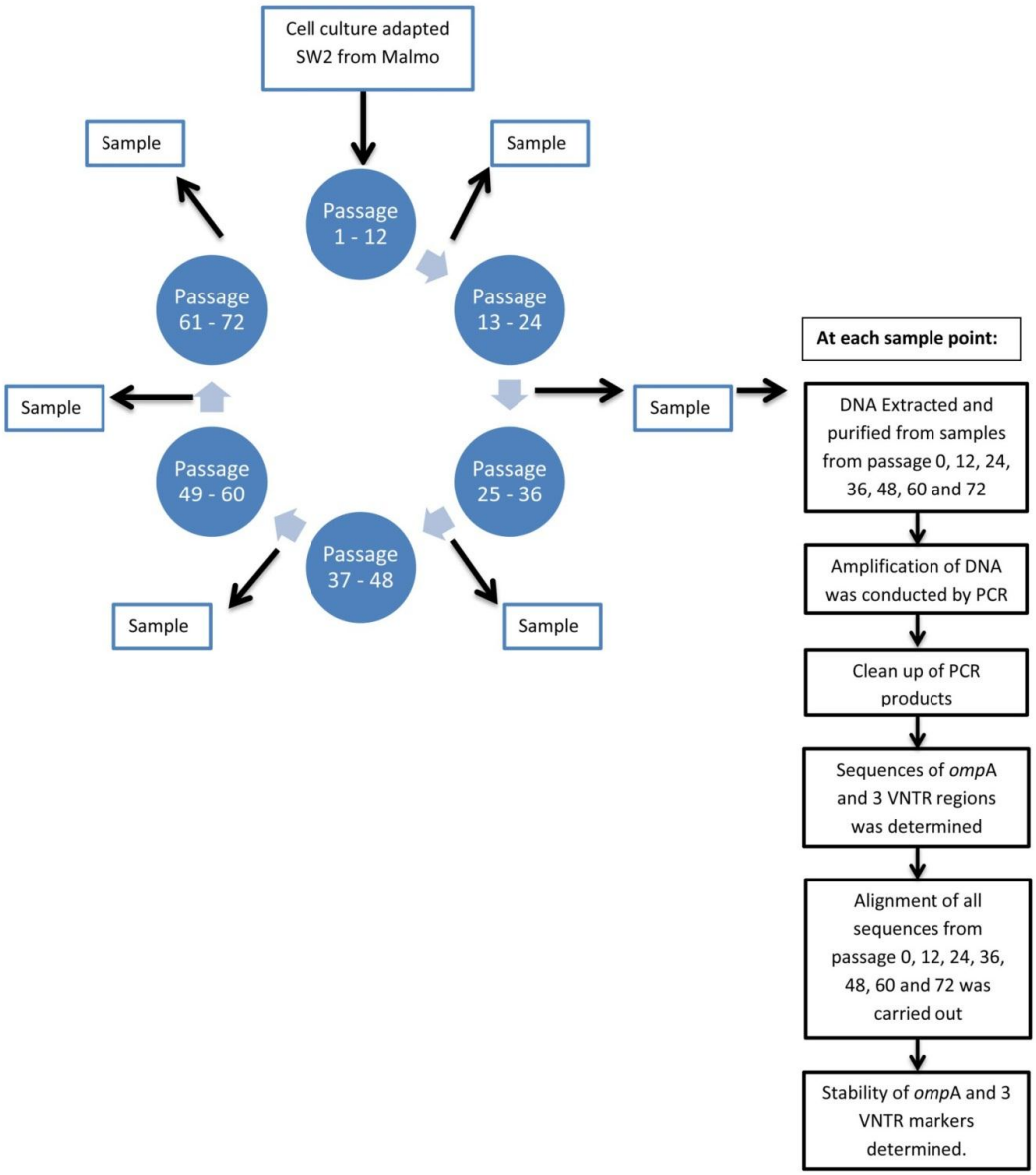


Figure 3.3 Flow diagram depicting the passage process of Sweden 2 and how the sample was processed to determine the stability of *omp*A and the VNTR markers.

3.4 RESULTS

3.4.1 Determination of the stability of the MLVA-*omp*A markers after eight passages in cell culture.

The sequences of the MLVA-*omp*A markers were aligned and analysed to determine the stability of the loci. The alignments showed that the MLVA-*omp*A markers are stable upon adaptation of the samples to cell culture using McCoy cells as the sequences of the markers were unchanged from the sequence before the samples had been passaged. Additionally, there was no change in the dominant sequence after the passaging process, which is the dominant peak in a sequence when there are two peaks at the same nucleotide position (Figure 3.4).

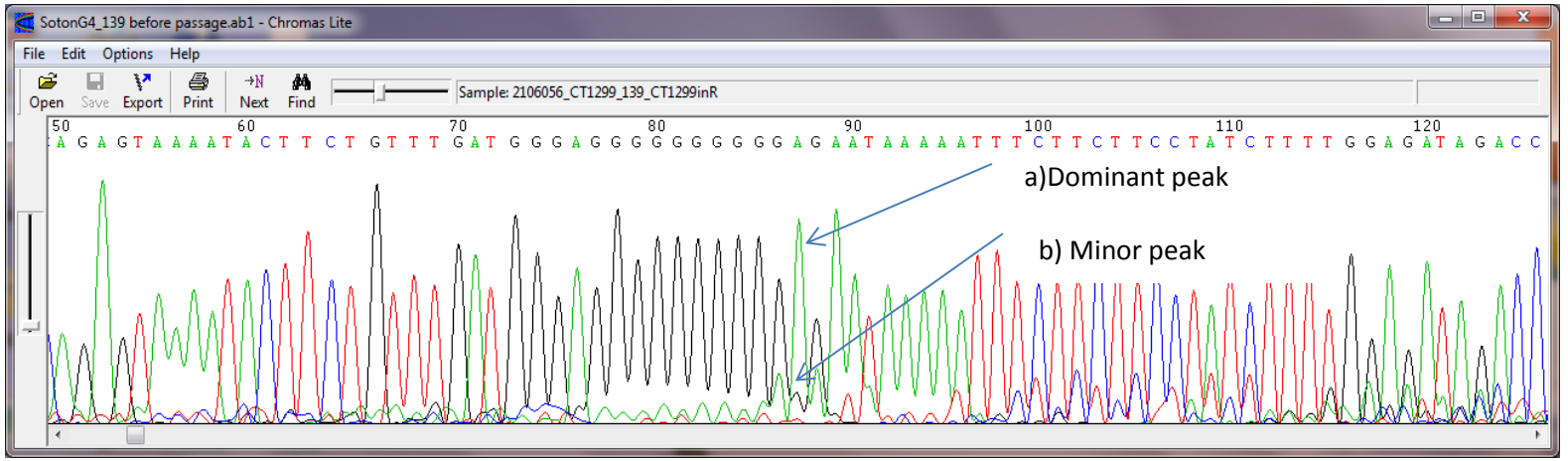
3.4.2 Determination of the stability of the MLVA-*omp*A markers in the Swedish new variant following long term passaging.

Sequences of the three MLVA-*omp*A markers were aligned for every twelfth passage up to and including passage 72. All the sequence data received from the sequencing company were aligned and there were no changes seen in the MLVA-*omp*A profile at any of the seven passages sequenced (Passage 0, 12, 24, 36, 48, 60, 72). The alignments of the chromatograms for the three markers – CT1335 and CT1299, CT1291 are shown in Figure 3.5, Figure 3.6 and Figure 3.7. The figures show that the sequence of the markers remain unchanged for all of the seven passage points that were sequenced. The *omp*A gene was also shown to be stable during the passaging process (chromatograms not shown due to the length of this gene).

3.4.3 Determination of the stability of the MST markers in the Swedish new variant following long term passaging.

Five MST markers were also analysed following long term passaging. These regions included: *hct*B, CT058, CT144, CT172 and *pbp*B. The five regions were sequenced and the sequences were aligned with sequence data from the reference genome for the Swedish new variant. The MST profile obtained for the regions was 21-19-1-2-1 which was identical to the original MST profile of the reference genome for the Swedish new variant. This shows that the five MST markers in the Swedish new variant are stable over long term passaging. *omp*A is also part of the MST scheme but its stability had already been determined in the Swedish new variant when the markers in the MLVA-*omp*A typing scheme were assessed as *omp*A is also part of that typing scheme (section 3.4.2).

A



B

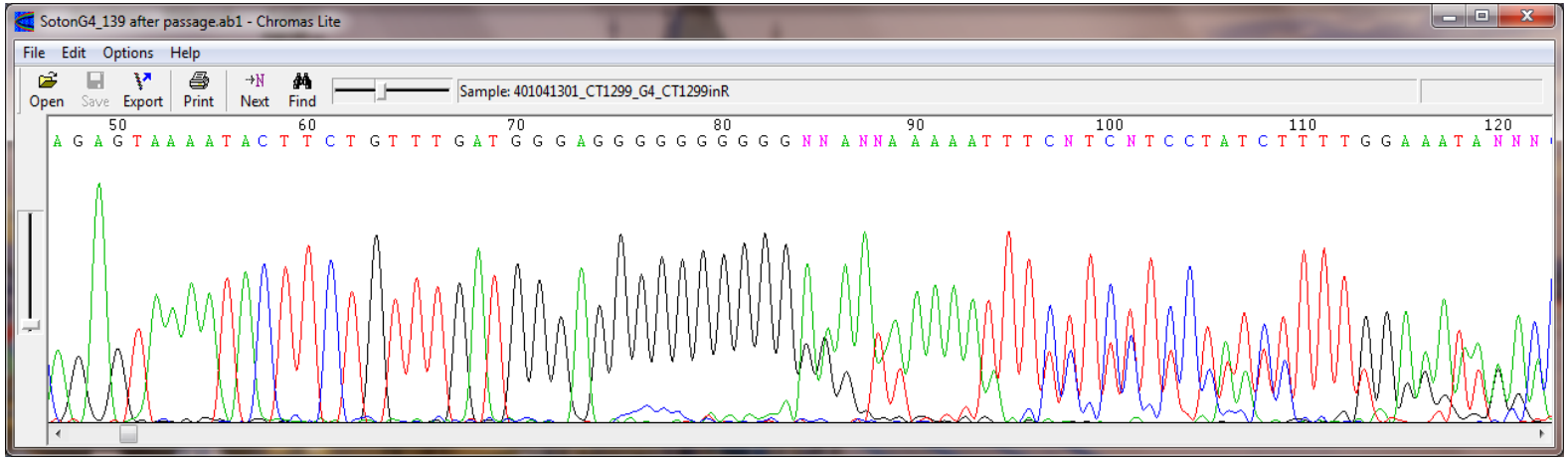


Figure 3.4 Chromatograms showing the VNTR region CT1299 for Southampton isolate G4
The diagram shows chromatograms before (A) and after (B) eight passages in cell culture. The arrows point to a) a dominant peak and b) a minor peak at a nucleotide position.

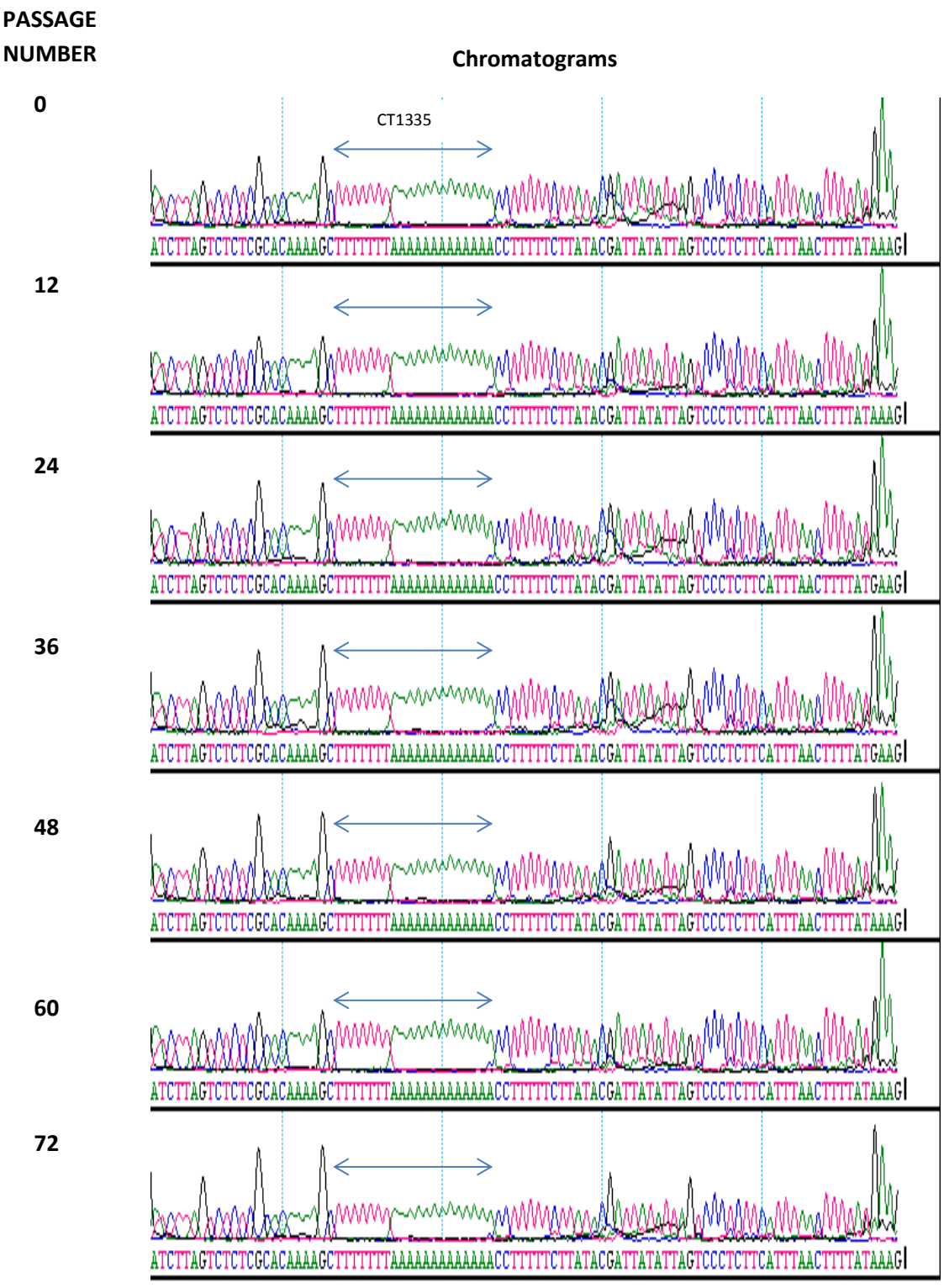


Figure 3.5 Alignment of marker CT1335.
Chromatograms show alignments from before the Swedish new variant was passaged and for every twelfth passage up to and including passage 72. Arrows indicate the VNTR region in each of the samples.

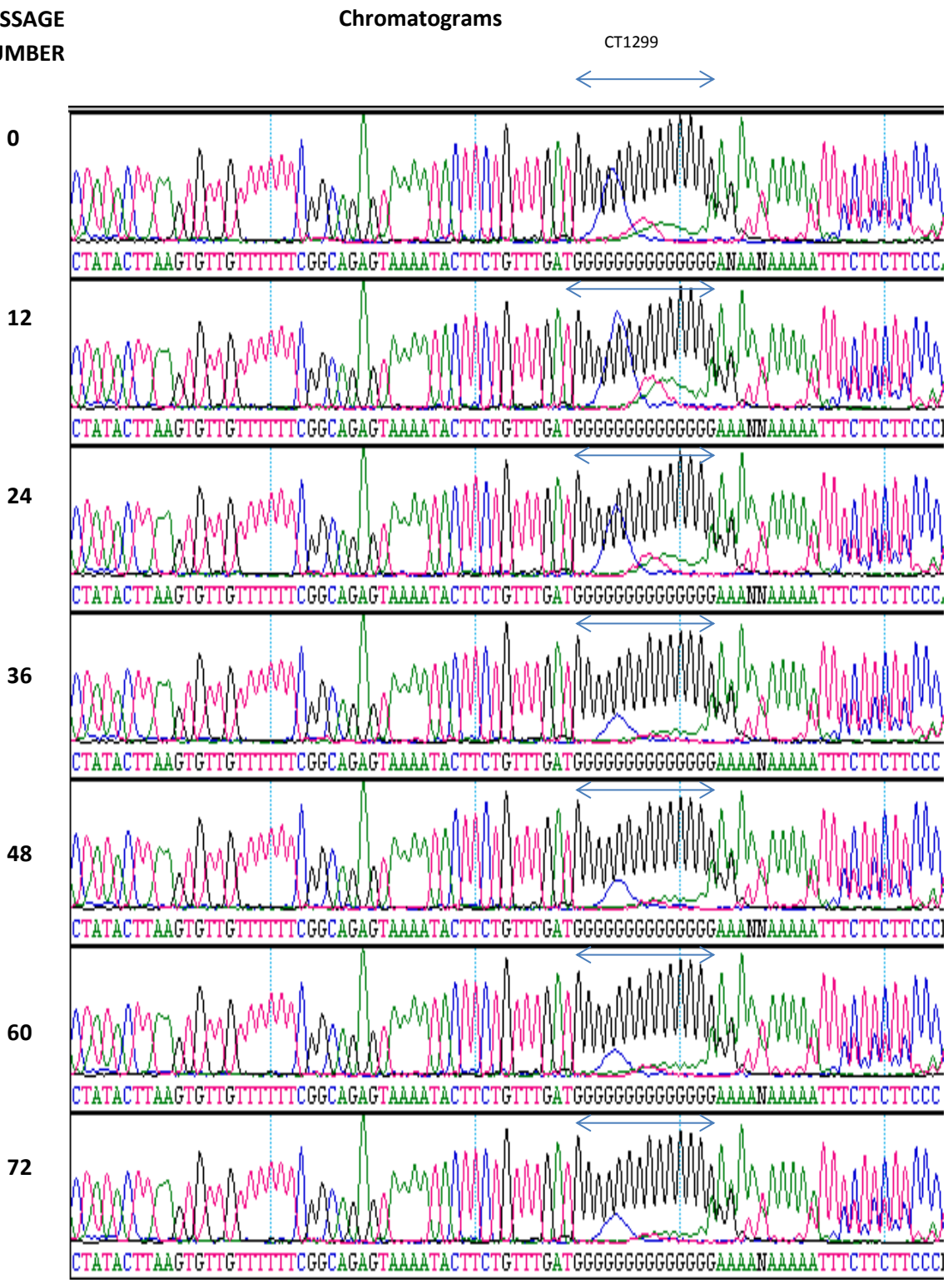


Figure 3.6 Alignment of marker CT1299.
Chromatograms from before the Swedish new variant was passaged and for every twelfth passage up to and including passage 72. Arrows indicate the VNTR region in the samples.

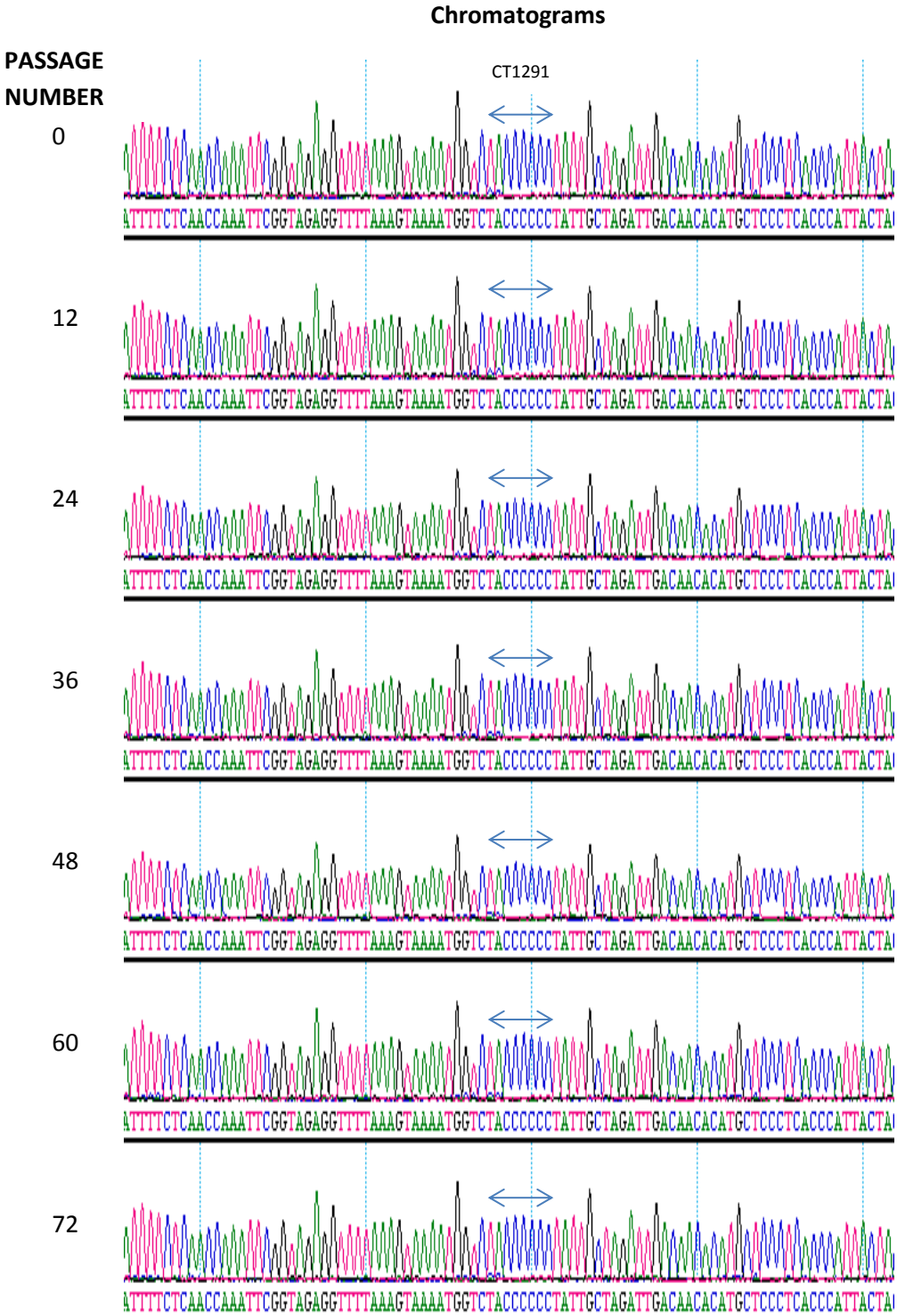


Figure 3.7 Alignment of marker CT1291.
Chromatograms from before the Swedish new variant was passaged and for every twelfth passage up to and including passage 72. Arrows indicate the VNTR region in the samples.

3.4.4 Whole genome sequencing of two isolates at Passage 36 and passage 72 of the Swedish new variant.

Although the data on the marker stability of the MST and the MLVA-*omp*A typing scheme suggests that the *C. trachomatis* Swedish new variant strain remains unchanged after long term passaging it was important to find out if there were any other changes that remained undetected elsewhere in the genome that would not be revealed by using this typing method. This would indicate whether stability of the MLVA-ompA markers could be equated to genome stability. Even though stability of various markers used for MLVA-*omp*A and MST typing had been determined there may be other changes elsewhere in the genome that are undetected by only assessing the stability of these markers. To investigate genome stability the entire genome of the isolate at passage 72 (the last passage carried out in cell culture for this study) was compared to the sequence to that of the original sequence of the new variant (available on the NCBI website). The 36th passage was also sequenced. It was necessary to sequence the isolate from passage 36 so that it could be determined if there were any changes within the genome that may have reverted back to the original sequence by the 72nd passage.

Following the whole genome sequencing process binary alignment/Map files (BAM file) which are compressed forms of sequence alignment/Map (SAM file) files were generated. These files show the aligned reads generated from the whole genome sequencing process. This alignment can then be mapped against the reference genome. The Bam files generated from the whole genome sequencing process of passage 36 and passage 72 were mapped to the reference sequence of the Swedish new variant whole genome that was readily available on the NCBI database.

SNPS between the genomes were identified using the script multiple_mapping_to_bam. py which is a script designed at WTSI by Dr. Simon Harris [personal communication]. The script was run through Putty (a free open source terminal which allows file sharing and the use of scripts available at the WTSI) by entering the following command:

```
~sh16/scripts/multiple_mappings_to_bam.py -r SW2.fasta -g -t *fastq.gz
```

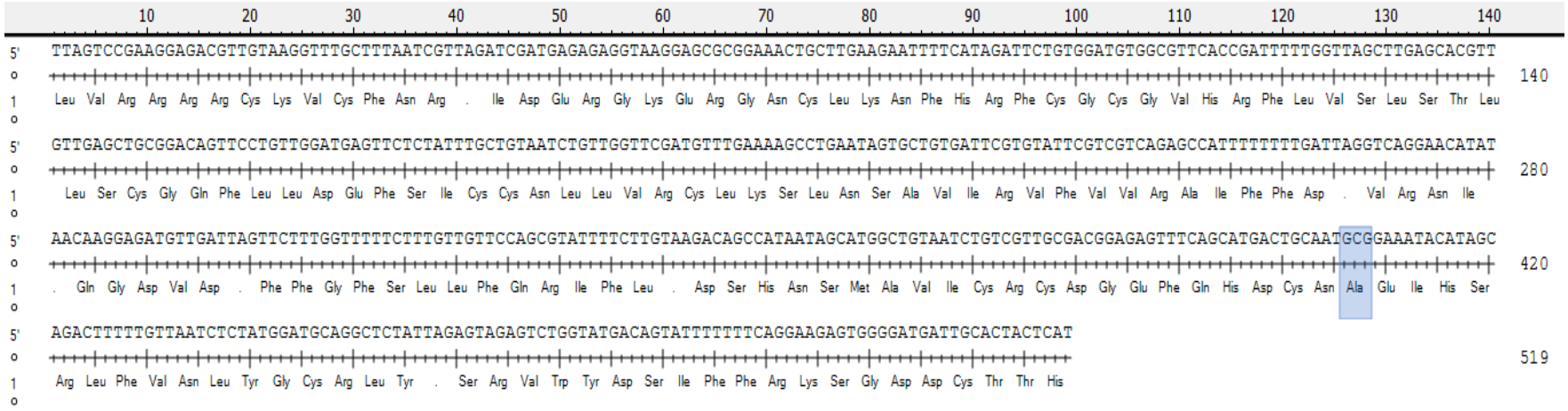
This command instructed the script to map the Swedish new variant sequence that had been passaged 36 times and the sequence of the isolate that was passaged 72 times (these files were in a fastq format which is a text based format of the sequence) to the original Sweden 2 sequence (SW2.fasta).The –g option was included in the command so that the script created mapping plots. The –t option was also added to the command so that a tab file of the SNPs would also be produced.

3.4.4.1 Determination of the stability of the whole genome of the Swedish new variant following long term passaging.

Following the mapping process two SNPs were identified. These SNPS were located at positions 998457 of the 36th passage of the Swedish new variant and in position 634198 in the 72nd passage of the isolate. Interestingly, the SNP found in the isolate at the 36th passage was not found in the 72nd passaged isolate and *vice versa*. The SNP found in passage 36 of the Swedish new variant is in a hypothetical protein which has been given the name SW2_8621 and is 519bp in length from position 998050 to 998568 (figure 3.8). This point mutation is a synonymous mutation as it does not change the translated amino acid. The amino acid that was produced in the original sequence of the Swedish new variant is alanine and as the SNP is an A to G point mutation the amino acid remains as alanine in the 36th passage and therefore it does not have any effect on the downstream protein synthesis. This can be seen in figure 3.8 where the amino acid sequence remains unchanged despite the mutation.

The second SNP identified was found in the 72nd passage of the Swedish new variant and was found in a non-coding region of the genome and therefore did not have any apparent effect on protein expression. The SNP was located between two genes- upstream of SW2_5701 which codes for a hypothetical protein and downstream of SCTL1 which codes for a type III secretion protein. As this SNP was located upstream of the 3' end of SW2_5701, it was necessary to determine that the SNP did not interfere with termination of transcription for this gene as a SNP within this region may affect the transcription of SW2_5701 and cause transcription not to terminate correctly. Likewise, as it was located downstream of the 5' end of SCTL1 it was necessary to check that the SNP was not located in the promoter region for the SCTL1 gene as this is the region that initiates transcription and a SNP within this region may interfere with transcription of SCTL1. To check this, the transcription termination factor for SW2_5701 was identified by finding the sequence for inverted tandem repeats located at the end of the gene followed by a string of adenines. The SNP found in the 72nd passage was not located in the region associated with termination of transcription for this gene and therefore it is unlikely that the SNP would have any effect on the termination of transcription of SW2_5701. Similarly, promoter sequences were identified using an online promoter identifying tool called BPROM. The SNP identified in the 72nd passage was not located in the promoter region. Following this analysis it is not likely that this SNP would have any obvious effect on protein expression, termination of transcription for SW2_5701 or on the promoter region of SCTL1.

A



B

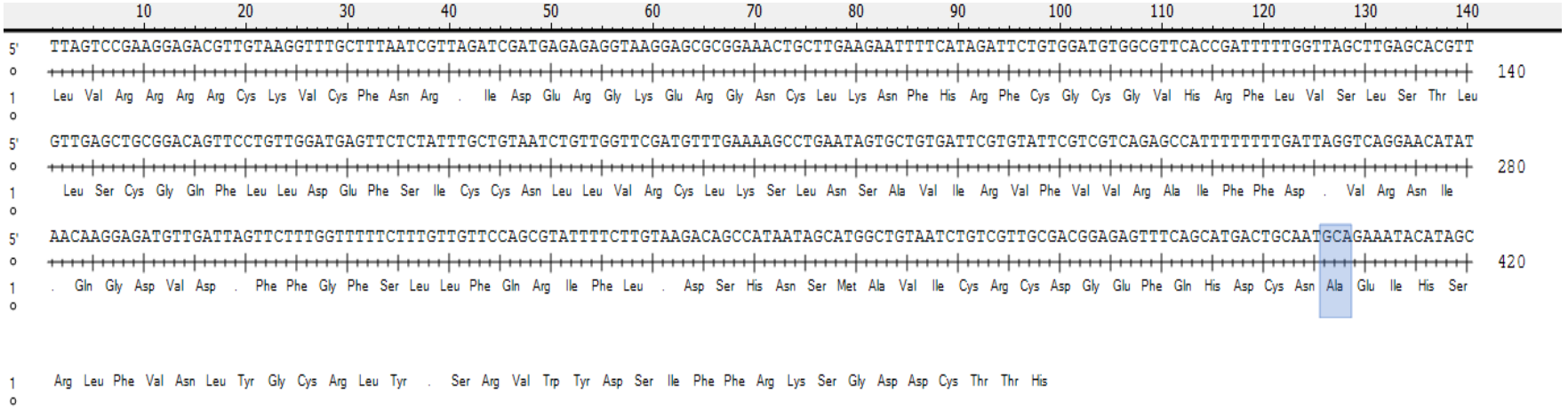


Figure 3.8 Sequence of the hypothetical protein SW2_8621 in Sweden 2.
SW2_8621 is located at position 998050 to 998568 of the Sweden 2 *Chlamydia trachomatis* genome. Below the sequence of SW2_8621 are the corresponding amino acids. A shows the original sequence of SW2_8621. B shows SW2_8621 with the point mutation at position 998457. Highlighted in blue is the amino acid sequence for the region with the point mutation. The SNP change from an A to G can be seen at the third position of the highlighted region.

3.5 DISCUSSION

The guidelines for validation and application of typing methods for use in bacterial epidemiology suggest that in a good typing system the sequences selected should remain stable during the period of time in which the study is conducted, as well as having a high discriminatory power, good typeability and reproducibility (van Belkum *et al.*, 1997). A review by Bjorn-Arne Lindstedt states that it is imperative that the VNTR *loci* selected for MLVA typing are thoroughly checked before being applied to real life situations and this includes long term passaging of isolates as the use of genomic regions which are not stable may not give a true epidemiological picture of the strains circulating within a community (Lindstedt, 2005). *Omp*A stability has previously been determined and was shown to be stable over 20 *in vitro* passages (Stothard *et al.*, 1998). However, two other studies carried out in Sweden that have used the *omp*A gene to assign *omp*A genotypes to samples, in order to study the epidemiology of the Swedish new variant and of L2b, have shown that there are minor point base substitutions in the *omp*A region. Following amplification of the *omp*A gene for several samples of the Swedish new variant and of L2b collected from patients, it was found to differ at a single nucleotide position in both studies for at least one of the samples collected in the study (Herrmann *et al.*, 2008, Christerson *et al.*, 2010). This shows that the *omp*A gene may not be stable *in vivo.* If the gene was stable *in vivo* we would expect all samples of the Swedish new variant and of L2b to be clonal.

The purpose of this study was to compare and evaluate two typing systems – MLVA-*omp*A (Pedersen *et al.*, 2009) and MST typing systems (Klint *et al.*, 2007). The chances of observing genetic alterations were increased in this study, as it was carried out to determine the stability of more *loci* around the genome and not just a single locus such as *omp*A with a lower discriminatory power. As the VNTR *loci* are generated as a consequence of an increased error rate during DNA replication at these regions, they are therefore potentially prone to additional variation (Gomes, 2012). At the time this study was carried out the paper by Peuchant *et al* (2012) had not yet been published and it is important to note that the markers described by this group also offer a high discriminatory power and the stability of these markers in different isolates can be determined as part of future work. However, the stability of the markers in this typing scheme can be determined from the whole genome sequences of the Swedish new variant.

The study by Pedersen *et al.* (2008) showed some evidence that markers in their typing scheme are stable. Their stability study of the markers CT1335, CT1299, CT1291 and *omp*A, is based on results obtained from typing isolates from recurrent or persistent infections in 24 patients and did not use tissue culture to grow defined generations of a single strain against which to measure stability. Such an *in vivo* test can only be conducted by making assumptions such as the 24

individuals were still infected by the same isolate and they had not contracted chlamydia from someone else. However, their study does have the benefit of observing stability in human hosts rather than cell culture and so highlights stability of the markers in the context of the real world. Similarly the Swedish new variant has been shown to be clonal by isolating the variant from various patients (Herrmann *et al.*, 2008, Jurstrand *et al.*, 2010, Klint *et al.*, 2011). The isolates were shown to have similar *omp*A genotypes and also the same MST type (Herrmann *et al.*, 2008, Jurstrand *et al.*, 2010, Klint *et al.*, 2011). These studies show that the Swedish new variant may be stable *in vivo*. However, it was not known as to how many developmental cycles the Swedish new variant has actually been through in these studies. Therefore it was useful to determine the stability of the VNTR and MST markers in cell culture as we could carefully define the number of passages and even though an *in vitro* system lacks the pressure from the immune system, normal genital microflora, the pH of the system and hormones (Pedersen *et al.*, 2008, Gomes, 2012), cell culture is still a suitable model to determine the stability of these markers as it is impossible to perform such analysis *in vivo* under controlled conditions (Gomes, 2012).

The first part of the study was carried out to determine if cell culture conditions had an effect on the stability of the markers chosen for a MLVA-*omp*A typing scheme for *C. trachomatis* (Pedersen *et al.*, 2008). The initial thinking was that the markers would be unstable as the chlamydia was growing in conditions where it was not replicating in its usual host and was possibly subject to different selective pressures. The initial experiment was conducted using swabs obtained from a study in 2009 (Wang *et al.*, 2011b). The samples were taken through eight passages in cell culture. This study has shown that over eight passages the MLVA-*omp*A markers are stable. The stability was determined by aligning the chromatograms and sequence data after eight passages to sequences obtained from the original swabs. Analysis of the alignments of the *loci* showed that the markers were stable through adaptation to cell culture as the sequence was unchanged in comparison to the sequence before the sample had been grown in mouse cells and also that there is no change in the sequence after eight passages in cell culture.

As part of the experimental design all of the samples that were chosen for the first study had different MLVA-*omp*A types. This acted as a way of ensuring that no contamination had taken place during the passage process. Three isolates with different *omp*A types but the same MLVA type were included in this study and this could be evidence that recombination of the *omp*A gene does occur (Harris *et al.*, 2012), as the marker sequence was the same for the three VNTRs in these samples but different for the *omp*A sequence.

Although the first study showed stability of the MLVA-*omp*A markers over eight passages *in vitro*, it was also imperative to evaluate the stability of these markers over a longer period of time to determine if there were any changes that occurred following long term passaging. This was important for future studies where one of the aims of a project may be to define sexual networks. It is useful to know that the markers are stable over longer periods of time and as they are passed from individual to individual so that networks can be better understood. The second part of the study was therefore designed to take the Swedish new variant which is an *omp*A genotype E strain through 72 passages. The Swedish new variant differs from other typical genotype Es as it has a 377bp deletion and a 44bp duplication in the cryptic plasmid (Unemo *et al.*, 2010). The reason for choosing the variant for this study was because it had already undergone changes and could be seen as unstable (Gomes, 2012) and is putatively vulnerable to acquiring changes, therefore if there were any differences to be observed in the markers this would be the best strain available with which to test stability. With this variant both the MLVA-*omp*A and MST markers were tested for stability.

For determination of the stability of the MLVA-*omp*A markers in the *C. trachomatis* Swedish new variant all the sequence data for each of the markers that make up the scheme were aligned against each other and the reference sequence for the Swedish new variant available on the NCBI website. Again there were no changes seen in the MLVA-*omp*A profile at any of passages that had been sequenced (Figures 3.5 to 3.7). The figures show the alignment for the sequences obtained for every twelfth passage including passage 72 for markers CT1335, CT1299 and CT1291, and how the three *loci* are stable through all the seven passages. Following this analysis it was concluded that there were no changes in the sequence even after 72 passages.

The stability of the MST markers in cell culture was also determined for the Swedish new variant by colleagues in Sweden as these markers were used in an MST typing system and offered high discrimination between *C. trachomatis* strains. The sample from passage 72 of the new variant was sent to the Department of Medical Sciences, Malmö, Sweden. For this study, five MST regions were analysed including *hct*B, CT058, CT144, CT172, *pbp*B and were first described in 2007 by Klint *et al.* (Klint et al. 2007). The five regions were sequenced and then aligned against corresponding sequences for each marker that was readily available from the sequence data of the reference genome of the Swedish new variant. The MST profile obtained was 21-19-1-2-1 and this was the same as the MST profile of the Swedish new variant. Their data showed that these five markers were also stable when the nvCT strain was passaged in cell culture up to and including passage 72.

Although a group in Amsterdam has found difficulty with interpretation of the Pedersen *et al* (2008) VNTR regions as the Taq polymerase they use for PCR is unable to resolve ambiguities such as double peaks and shifts within these regions (Bom *et al.*, 2011), this phenomenon has not been experienced in my study and this can probably be attributed to the use of high fidelity DNA polymerase for the PCR in this study compared to the DNA polymerase used in their study. However, what was observed in the chromatograms was the presence of larger peaks (assigned dominant peaks) and smaller peaks (assigned minor peaks) in one of the VNTR loci- CT1299 which made assignment of the sequence a little ambiguous. Figure 3.4 shows the presence of a dominant peak and also a minor peak at a single nucleotide position. The dominant sequence suggests that this base is an A whilst the minor peak suggests that at this nucleotide position the base is a G. For this study the dominant sequences have been recorded and compared. For all of the isolates sequenced, the nucleotide for the dominant and minor sequence observed before and after passage was always the same and therefore stability was confirmed. The phenomenon of dominant and minor bases has previously been observed (Kommedal *et al.*, 2009). Multiple populations occur when an individual is infected with more than one strain. It has been shown that chromatograms which have mixed sequences do offer accurate information of the individual strain within a population eliminating the need to synthesis new primers or having to clone the individual variants (Dueck, 2009).

Following recommendations by one of the reviewers of the paper published for this study (Labiran *et al.*, 2012b) and who later wrote a commentary (Appendix 2) after the publication of our paper suggesting that some of the samples should be whole genome sequenced after long term passage (Gomes, 2012). Therefore two samples were sent to be whole genome sequenced from the 36th and 72nd passage of the Swedish new variant. The samples were sent to the WTSI. The whole genome sequences were subsequently compared to the reference genome to determine if there were any changes elsewhere in the genome at both passages. There were two definite SNPs recognised following the filtering process, which calls a 'true' SNP based on its appearance in the majority of reads. It is important to note that there were other possible mutations found in some of the reads but because they were not found in the majority of the population then they were not classified as 'true' mutations following the filtering process. Additionally even though the SNP found in the genome sequence of passage 36 of the Swedish new variant was not found in passage 72, the SNP was seen in some of the (Passage 72) reads when the Bam files were compared. However, because the SNP was not detected in the majority of the reads the SNP calling program did not call it as a 'true' SNP. Likewise the SNP detected in passage 72 could be detected in some reads of passage 36 but again as it was not seen in the

majority of (Passage 36) reads it was therefore not called a 'true' SNP. Following analysis the SNPs identified did not have any effect on the genome.

The VNTR and MST markers have been shown by other groups who have studied the suitability of the markers in these typing schemes for typing samples, to have high typeability, reproducibility and discriminatory power (Pedersen *et al.*, 2008, Bom *et al.*, 2011) according to the definitions outlined by Van Belkum *et al.* (van Belkum *et al.*, 2007). My study has shown that the markers in these typing schemes are also stable over several passages. These typing schemes can therefore be used to study both short and long term epidemiology even though sequences elsewhere in the genome may vary as the chlamydia is sexually transmitted from person to person. However, the MLVA-*omp*A typing scheme was chosen to be the best typing scheme for future epidemiological studies as it did not require the need for nested PCR, which was avoided due to contamination problems that can occur when this method is used.

Additionally, in a study by Wang *et al* (Wang *et al.*, 2011b) there was a selection of samples that could not originally be assigned types directly from the swab but once the samples were isolated in cell culture they could then be given a MLVA-*omp*A type (Wang *et al.*, 2011b). It is not known why this was the case at the time but what can now be said is that multiple passage in cell culture does not affect the sequence of the MLVA-*omp*A markers and so although types were very likely to be assigned after culture to a proportion of the samples, it can be said with confidence that the types were identical to those of the initial swabs and this strongly suggests that the type would not have changed during culture.

3.6 FUTURE WORK

Future work for this chapter includes assessing the stability of the markers in the MLVA-5 typing system described by Peuchant *et al* (Peuchant *et al.*, 2012) as it also offers high discriminatory power to type *C. trachomatis*. Although the regions used in this typing scheme do not have any mutations following long term passaging in the Swedish new variant, it will be useful to assesss the stability of the markers in this typing scheme using different strains as marker stability may be strain specific.

CHAPTER 4 DIVERSITY OF GENITAL TRACT STRAINS IN HIV POSITIVE AND HIV NEGATIVE MEN WHO HAVE SEX WITH MEN (MSM) IN BRIGHTON.

4.1 INTRODUCTION.

Brighton is a city located on the South East coast of the United Kingdom. The city has a large homosexual population and has previously been known as the 'gay' capital of the UK (Manchester and London also have a large gay population) (Browne and Lim, 2010).

Some men who have sexual intercourse with other men choose not to identify themselves as homosexual or bisexual, therefore the term men who have sex with men (MSM) is often used to identify any male who engages in sexual intercourse with another man. There have been large increases in sexually transmitted infection (STI) diagnoses in MSM due to high risk sexual behaviour and increased testing in MSM. Young people aged between 15 and 24 have also contributed significantly to the increase in STI diagnoses (HPA, 2012b). Among male GUM clinic attendees, 79% of syphilis cases, 58% of gonorrhoea cases, 17% of chlamydia cases, 11% of genital herpes and 9% of genital warts cases were attributed to MSM in England in 2012 (HPA, 2012b).

The diagnosis of STIs reported in MSM continues to rise on a yearly basis. In 2012 gonorrhoea diagnoses increased by 37%, chlamydia diagnoses by 8%, syphilis by 5% and genital warts by 8% compared to diagnoses of these STIs in 2011 (HPA, 2012b). The factors that have contributed to the increase in diagnoses of STIs in MSM include increased screening of extra genital sites (rectal and pharyngeal) and the use of highly sensitive Nucleic Acid Amplification Tests (NAATs) (HPA, 2012b). Furthermore, reporting of sexual orientation has continued to improve, therefore a greater number of diagnoses are now attributed to MSM than in previous years (HPA, 2012b).

A report published by the HPA shows that there were approximately 40,000 MSM in the UK in 2012 who were human immunodeficiency virus positive (HIV+) and a quarter of these men were unaware that they had the infection. One in twenty MSM were living with HIV in the UK, with one in twelve infected MSM residing in London and one in 32 living elsewhere in the UK (HPA, 2012b). The data also indicated that in 2011 approximately 48% of all new HIV diagnosis were in MSM (HPA, 2012b). MSM living with a diagnosed HIV infection in the UK, if diagnosed promptly, can expect to live to a near normal life expectancy as there are now effective antiretrovirals (ARVs) for individuals who are HIV+. One of the reasons for the continued increase in the rate of STIs amongst the MSM population is due to serosorting (Heymer and Wilson, 2010, Wilson *et al.*, 2010). Serosorting is the practice of choosing sexual partners based on their HIV status (Kennedy *et al.*, 2013) and leads to HIV positive MSM often having unprotected sex with other HIV positive

men. It also encourages HIV negative MSM to have unprotected sex with other HIV negative MSM as they believe there is no risk of passing on HIV but are unaware of the risks posed by other STIs. In theory, practising serosorting should lead to a decrease in the rate of new diagnoses of HIV but as individuals who practice serosorting rarely use condoms, data has shown that serosorting leads to an increase in other STIs including chlamydia (Heymer and Wilson, 2010, Wilson *et al.*, 2010, Mayer, 2011). In 2012, a study conducted in Sydney investigating the impact of having HIV and contracting other STIs found that serosorting was associated with an increase in both urethral and rectal chlamydia (Jin *et al.*, 2012).

The recommendations given to MSM by the HPA are as follows: they should be screened for HIV and STIs at least annually and more frequently if they are engaging in unprotected anal intercourse with casual and new sexual partners. Furthermore, the use of condoms should be encouraged. Despite these recommendations the number of HIV+ patients who are co-infected with other STIs, especially chlamydia, continues to rise yearly (HPA, 2012b). In 2011, approximately one in eight people newly diagnosed with HIV were simultaneously diagnosed with an acute STI (HPA, 2012b). One in four (21%) MSM diagnosed with HIV were also diagnosed with an acute STI compared to one in 25 heterosexual men and one in 30 heterosexual women. According to HPA data, chlamydia diagnoses in the UK MSM population increased from 1,300 in 2001 to 5,000 in 2010. Additionally, the number of individuals diagnosed with LGV (which is an invasive infection, affecting the lymphatic system and lymph nodes) more than doubled between 2009 and 2010, with more than 99% of LGV cases diagnosed in MSM, 80% of which were HIV positive. Figures also indicate that diagnoses of LGV rose from 190 infected people in the UK in 2009 to 530 in 2010. The rise in LGV cases is mainly due to the increased number of MSM who have unprotected anal intercourse (HPA, 2011).

Amongst MSM, transmission of STIs is often by anal sex, although anal sex has also become increasingly popular in the heterosexual community (Christerson *et al.*, 2012). It has been shown that the *omp*A genotype distribution of *C. trachomatis* is different between MSM and the heterosexual population. The prevalence of genotypes D, G and J is higher in the MSM population, whilst genotypes D, E, and F are more prevalent in the heterosexual population (Twin *et al.*, 2010, Li *et al.*, 2011, Quint *et al.*, 2011, Christerson *et al.*, 2012, Bom *et al.*, 2013b). A more recent study using the low resolution *omp*A typing and conducted on a Spanish MSM population, has shown that in this population, genotypes D, G and J are not the most prevalent genotypes and in fact the most prevalent genotypes in their MSM population were genotypes E, G and D (Mejuto, 2013). As only a small number of studies (Christerson *et al.*, 2012, Bom *et al.*, 2013a, Bom *et al.*, 2013b) had been conducted investigating the types of *C. trachomatis* found in the rectum and urethra of MSM further studies were required to investigate why this difference in distribution occurs

between heterosexuals and MSM. It was also important to use high resolution typing methods to identify if there were any differences between the two groups. Previous studies have mainly used low resolution *omp*A typing when assigning types to clinical samples (Barnes *et al.*, 1987, Boisvert *et al.*, 1999, Geisler *et al.*, 2002, Geisler *et al.*, 2008). During the past six years, the use of high resolution typing methods has meant that other differences within *omp*A genotypes can be determined (Klint *et al.*, 2007, Pedersen *et al.*, 2008, Bom *et al.*, 2011, Wang *et al.*, 2011b, Christerson *et al.*, 2012). It is also important to use typing methods other than low resolution *omp*A typing as recent studies have shown that the *omp*A gene is mobile between strains (Dean *et al.*, 2009, Harris *et al.*, 2012, Joseph *et al.*, 2012). Therefore assigning types using only the *omp*A gene may not be reliable when carrying out prevalence studies or when trying to track strains in the community to identify networks. Studies have been published where the *omp*A genotype E accounts for almost half of the samples typed (Bom *et al.*, 2011, Wang *et al.*, 2011b). However, high resolution genotyping methods have now been designed to study epidemiology in communities (Klint *et al.*, 2007, Pedersen *et al.*, 2008, Peuchant *et al.*, 2012).

The hypothesis for this study was that the tropism of *C. trachomatis* has a strong influence on the distribution of genotypes, i.e. that there are strains of *C. trachomatis* that have a selective advantage for growing in rectal tissue. This pilot study was designed to determine the distribution of chlamydia MLVA-*omp*A genotypes in MSM in Brighton and to determine whether the most prevalent genotypes found in MSM are different to the genotypes found in a heterosexual population as reported by other groups (Twin *et al.*, 2010, Li *et al.*, 2011, Quint *et al.*, 2011, Wang *et al.*, 2011b, Bom *et al.*, 2013b). These data can also be compared to the heterosexual population in Southampton (Chapter 5). The study was also designed to allow the comparison of chlamydia genotypes found in patients who were HIV+ and those who are HIV-, to investigate if there were differences between the two cohorts. As the HIV+ patients are immunocompromised, it may be the case that certain types of chlamydia grow better in immunocompromised people compared to those who do not have HIV and are therefore unlikely to be immunocompromised. In addition, HIV+ patients may engage in higher risk sexual behaviour than HIV- MSM and therefore different distributions of chlamydia types in these two groups may be observed.

4.2 AIMS OF PROJECT.

4.2.1 Primary aim.

To determine the distribution and prevalence of *C. trachomatis* genotypes in an MSM population using samples from patients attending the Claude Nicol Clinic, Royal Sussex County Hospital GUM clinic in Brighton.

4.2.2 Secondary aims.

- To determine the distribution of *C. trachomatis* genotypes found in urethral and rectal infection in HIV-positive and HIV-negative MSM. This is based on the hypothesis strains have different tissue tropisms and this may determine infection outcome.
- To determine the distribution of *C. trachomatis* genotypes found in urethral and rectal infection in HIV-positive and HIV-negative MSM, as these groups may represent different networks.
- To measure concordance of *C. trachomatis* types between the urethral and rectal sites in MSM infected at both sites.
- To observe microscopically if there are any phenotypic differences between chlamydia genotypes in cell culture.
- To examine correlations between ethnicity, clinical symptoms and other sexually transmitted infections with *C. trachomatis* genotypes.

4.3 STUDY DESIGN.

This study was set up as a feasibility study to determine if it was possible to conduct studies between the institutes in Brighton (Claude Nicol Clinic, Royal Sussex County Hospital GUM clinic) and Southampton (University of Southampton) and to investigate how to set up a larger study in the future. Additionally, the study was set up to determine if MSM networks could be identified from the typing data. The study design was to collect 100 samples in each group for comparison purposes (100 chlamydia positive samples from HIV positive and 100 chlamydia positive samples from HIV negative men). As this was a pilot study, 100 samples in each group was considered suitable to obtain an idea of the distribution of chlamydia types in each group in preparation for a full scale study. A total of two hundred samples were estimated as a practical number of samples collectable in the time frame of the study based on current diagnosis in the Brighton GUM clinic.

Samples collected at the GUM clinic were obtained by clinicians or nurses. When the swabs were taken during rectal examination there was no additional time cost or invasive sampling to the patient. In situations where swabs were self-collected, patients were only asked to provide a second sample taken by a clinician if they were asked to return to the clinic for treatment following diagnosis with *C. trachomatis*. Before this study began a single swab was collected to test for chlamydia, and if gonorohoea was suspected, an additional swab was collected from the patient to test for its presence. Initially, the swabs used to test for chlamydia were manufactured by Sigma as this was the make of swab that was used in the Brighton GUM clinic at the time for diagnostic purposes. However, the swabs in use were later changed to the BD ProbeTec™ swab and transport media. Following collection, the swabs were sent to the diagnostic laboratory in Brighton (which is in a different location in Brighton to the GUM clinic) for routine diagnostic tests

before being sent to the HPA sexually transmitted infection bacteriology reference laboratory (STBRL) in Colindale for re-testing (swabs were sent to Colindale due to the rise in asymptomatic LGV cases to confirm that they were LGV strains). This was a hindrance to the study because to maximise the chances of extracting DNA and for assigning genotypes to each sample, the cold chain had to be maintained at all times. Additionally, the BD ProbeTec™ swabs were unsuitable for the isolation of *C. trachomatis* in cell culture as they contained a lysis buffer which would inactivate the chlamydia. Therefore, to achieve all of the aims of this study additional swabs were collected. Ethical permission was sought and granted (REC reference 11/LO/0745) for the collection of an extra two swabs. The swab and transport media system manufactured by BD ProbeTec™ was used as the source of sample DNA for MLVA-*omp*A genotyping. The second swab and transport media system was manufactured by Sigma for the isolation of viruses, chlamydia and mycoplasma. This swab was collected for the isolation of chlamydia from samples in cell culture at the Molecular Microbiology Department, University of Southampton. Following collection, the swabs were stored at -70°C prior to genotype analysis and for the isolation of *C. trachomatis*. This equated to a total of three swabs collected at the GUM clinic from each individual who had given consent for their samples to be included in this study. Additionally, participants were also asked to provide a urine sample to test for urethral *C. trachomatis* and the residual sample was stored for analysis in Southampton if the sample was found to be positive. All swabs and urine samples were anonymised before transportation to Southampton. A retrospective case note review of symptoms, signs and sexual history over the preceding three months was conducted by Dr. Suneeta Soni and Dr. Judith Zhou in Brighton to examine the correlation between these factors and *C. trachomatis* genotypes.

4.4 SWABS CHOSEN FOR THE ISOLATION OF *CHLAMYDIA TRACHOMATIS* IN CELL CULTURE AND FOR GENOTYPING.

To assign genotypes to samples, the residual material from the BD ProbeTec™ swabs were used to extract DNA for typing purposes. Following changes to the agreed procedure due to an increase in cases of LGV, which meant that samples were sent to the STBRL in Colindale before they were sent to Southampton, an extra BD Probetech swab was collected for typing.

The virus, chlamydia, mycoplasma and ureaplasma (VCM) swabs (Medical wire, Wiltshire) were chosen for the isolation of *C. trachomatis* in cell culture as live chlamydia could be recovered from these swabs (according to the manufacturer). The staff at the GUM clinic were already accustomed to using the swabs by this manufacturer and therefore it was considered reasonable to obtain these swabs for isolation of chlamydia from samples. Additionally, these swabs do not contain chloramphenicol (a bacteriostatic antimicrobial effective on Gram-positive bacteria and

also on Gram-negative bacteria such as chlamydia) as some swabs and transport media systems do, therefore, these swabs were suitable for the study requirements.

4.4.1 BD ProbeTec™ swabs and transport media chosen for the storage of *C. trachomatis* samples before DNA extraction and typing.

The BD ProbeTec™ ET *C. trachomatis* and *N. gonorrhoeae* amplified DNA assay is used with the BD Viper™ system which uses Strand Displacement Amplification (SDA) technology for the direct, qualitative detection of *C. trachomatis* and *N. gonorrhoeae* DNA in endocervical swabs, urethral swabs and urine specimens. The information sheet (Appendix 3) provided with the swabs states that after collection, the endocervical swabs and male urethral swabs must be stored and transported to the laboratory at 2-27°C within 4-6 days. Storage up to 4 days had been validated with clinical specimens and storage up to six days had also been demonstrated with seeded specimens by the manufacturer.

4.4.2 Virus, chlamydia and mycoplasma (VCM) swabs and transport media chosen for the isolation of *C. trachomatis*.

Σ-VCM™ is a universal transport medium used to recover virus, chlamydia, mycoplasma and ureaplasma (VCM) from swabs. There is currently no validated, approved cell culture system for rectal isolation of *C. trachomatis,* however, according to the manufacturers, the VCM swabs offer optimum absorption and release of material. Validation studies (Appendix 4) have been conducted by the manufacturer to show the correlation between NAATs tests and culture using these swabs (Appendix 5). Even though the manufacturers had recommended these swabs for the isolation of chlamydia using their equipment, to assess the suitability of these swabs for this particular study, using the equipment available at the Molecular Microbiology Department, University of Southampton the VCM swabs were spiked with $2.5x10^5$ifu/ml of chlamydia genotype E/Bour and successfully recovered. Therefore, these swabs were considered suitable for routine isolation of *C. trachomatis* from samples using the resources available.

4.5 DETERMINATION OF THE OPTIMAL STORAGE CONDITIONS FOR RECOVERY AND ISOLATION OF *C. TRACHOMATIS* FROM THE VCM TRANSPORT MEDIUM.

4.5.1 4°C and -20°C stability study.

After deciding that the Σ-VCM™ transport media and swab were suitable for isolation of *C. trachomatis*, the optimal storage conditions for recovering chlamydia were determined. This was necessary because there was a time lag of up to a month before the swabs were collected from the GUM clinic in Brighton and delivered to the Molecular Microbiology Department, University of Southampton and also between delivery at University of Southampton and isolation of chlamydia in cell culture. A set of experiments were conducted at 4°C and -20°C to ensure that the chances of recovering live chlamydia would be maximised. The reason for initially choosing these two temperatures to perform these stability studies was because there was already a 4°C fridge and -

20°C freezer available at the GUM clinic. As the GUM clinic was regularly very busy with only limited space for sample storage, it was considered a reasonable approach to firstly determine if chlamydia could be recovered from the swabs following storage at these temperatures. Additionally, it would have been ideal to use a storage facility already available at the GUM clinic to reduce confusing the staff during busy clinic periods when trying to decide which samples go in to which storage facility and also to reduce costs associated with buying a new storage facility. Furthermore, biological samples are easily stored at 4°C and -20°C but the duration in which chlamydia samples could be stored in the Σ-VCM™ transport media at these temperatures was unknown.

To analyse the effect of storing these samples for an extended period of time at these temperatures the transport medium was spiked with *C. trachomatis omp*A genotype E/Bour and stored at the various temperatures. The proposed experimental plan is shown in figure 4.1.The sample was diluted in DMEM (+10% (v/v) FCS) for the dilution series as this medium is suitable for the growth of *C. trachomatis* in cell culture and would thus preserve viability. The concentration of infectious elementary bodies present was determined by performing a dilution series and staining the inclusions using the X-gal staining method (Section 2.10). Seven, ten-fold dilutions were set up for each of the 'time points' to be tested (0, 7, 14 and 28 days post spiking) and for both of the temperatures selected (4°C and -20°C). On the day the dilution series was carried out, McCoy cells were infected with a set of dilutions to evaluate whether *C. trachomatis* could be cultured from this transport medium and also to quantify the EBs in each dilution at time zero. This was used as a comparison for the other time points to determine if infectivity had reduced following storage.

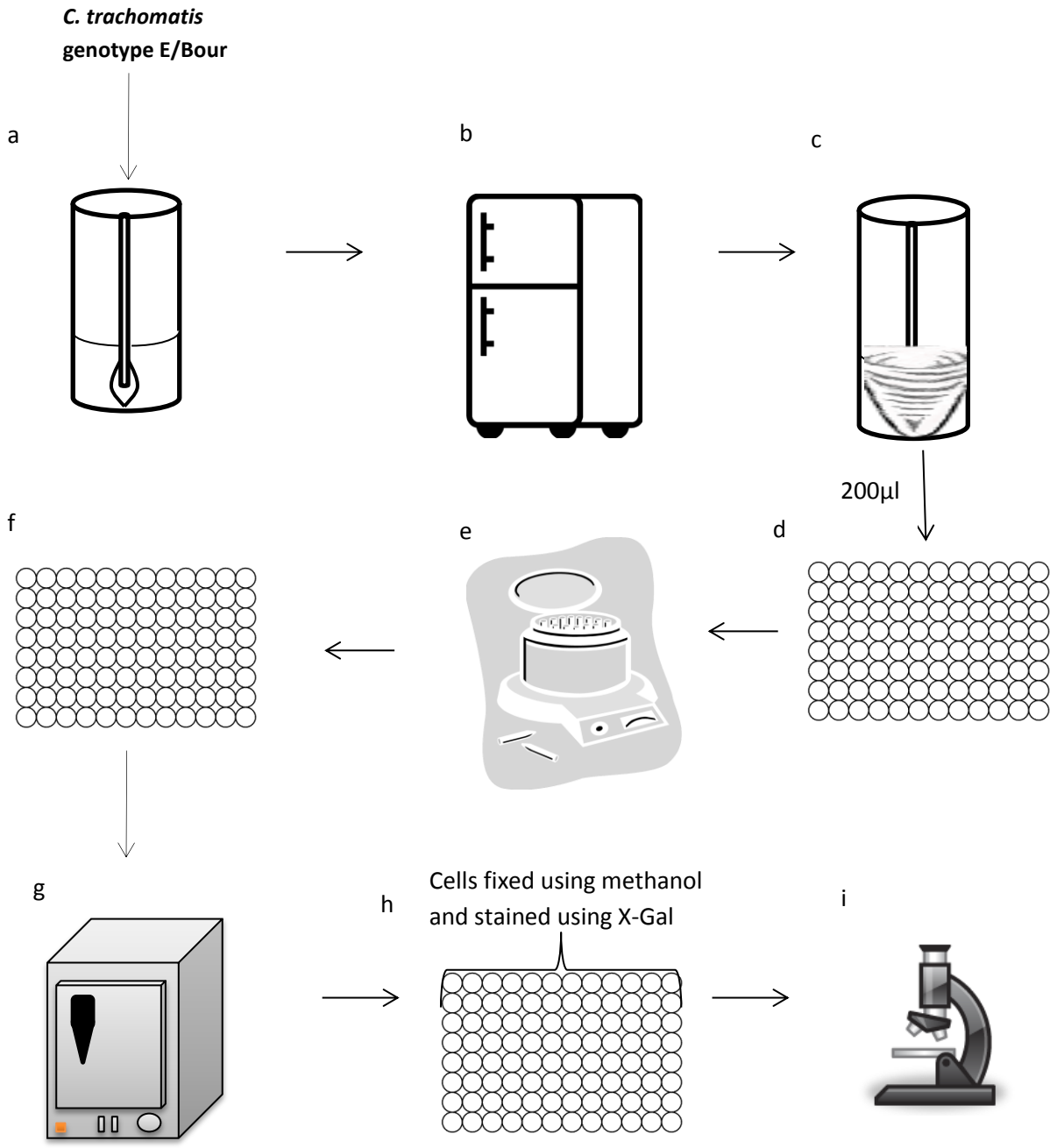


Figure 4.1. Experimental plan to analyse the effect of storing samples for different periods of time at 4°C and -20°C.

Swabs were spiked with *C. trachomatis* E/Bour, using the following concentrations: neat, 2.5 x 10^5 ifu/ml), 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6}. These dilutions were then stored in either a fridge (4°C) or a freezer (-20°C) for different durations of time (b). On days 7, 14 and 28, the plan was to take the swabs out of the fridge and the freezer and leave them to equilibrate at room temperature, then vortex the samples for one minute (c). 200µl of each sample would then be added to McCoy cells in a 96 well plate (d). Subsequently, the plates would be centrifuged for an hour at 754 x g (e). The media in the wells would be replaced with media containing gentamicin (1mg/ml) to kill other bacteria (but not chlamydia) and cycloheximide (10mg/ml) which would be added to reduce the growth rate of the cells (f) and incubated at 37°C/ 5%(v/v) CO_2 for 72 hours (g). Two days post infection the cells were to be fixed using cold methanol and the inclusions stained using X-gal (h) so that they could be clearly visualised under a microscope (i) and counted to determine the quantity of inclusions in comparison to before the samples were stored at the various temperatures.

4.5.1.1 Results of the 4°C and -20°C stability study.

C. trachomatis was successfully recovered up to the 10^{-3} ifu/ml dilution at time point zero as expected. At this time point the swabs had not been stored at either 4°C or at -20°C. Seven days post spiking of the swabs, the first set of dilutions were taken out of the fridge and freezer and used to infect McCoy cells using the method outlined in section 4.5.1 for day zero. No inclusions were seen when the cells were visualised under the light microscope at day seven in any of the wells infected with the various dilutions or at any of the two temperatures tested. These results indicated that there was no need to continue with the other time points as no inclusions could be seen in these dilutions even after just seven days.

This data showed that, firstly it was possible immediately to recover *C. trachomatis* that had been placed in the virus chlamydia and mycoplasma transport medium which was shown by recovering *C. trachomatis* at time point zero. However, storage at 4°C and -20°C did not maintain the viability of *C. trachomatis.* Therefore, it was concluded that chlamydia would not survive up to seven days in the transport medium at these temperatures.

Subsequently, this experiment was repeated, but this time 4SP was added to the transport media before the samples were stored at -20°C. 4SP is a cryopreservative which works by lowering the melting point of water so that ice crystals do not form within the bacteria as the majority of the bacteria is composed of water. Once ice crystals begin to form within the bacteria this causes the bacteria to expand which could result in damage, therefore the experiment was repeated to determine if the addition of 4SP would prevent ice crystals from forming and thereby improve recovery of chlamydia. Unfortunately, the addition of 4SP did not improve the recovery of chlamydia. No chlamydia could be recovered seven days post spiking of the transport media with EBs and after storage of the swabs in transport media at the two temperatures. At this stage it was concluded that there may be an additive in the transport media which lysed the chlamydia during long term storage. However, this seemed unlikely. It was also a possibility that the various temperatures at which the transport media was stored were inadequate storage conditions as it has been shown that ice crystals grow the fastest by diffusion at -20°C and therefore this temperature may not be ideal for storage of chlamydia (aufm Kampe *et al.*, 1951). Additionally 4°C is inadequate for long term storage and preservation.

4.5.2 -70°C swab stability study to determine the viability of *C. trachomatis* in VCM transport media.

Following the experimental determination that storage of chlamydia at 4°C and -20°C (although convenient for the GUM clinic staff) were not optimal storage conditions for the samples in this

study, a comprehensive literature review on storage conditions used previously was conducted to determine the optimal storage facility that had to be employed for the GUM clinic in Brighton.

All the studies investigating the best storage conditions for *C. trachomatis* were over 18 years old and showed that to increase the recovery rate of chlamydia following long term storage, freezing at -70°C or liquid nitrogen is required (Reeve *et al.*, 1975, Ngeow *et al.*, 1981, Mahony and Chernesky, 1985, Maass and Dalhoff, 1995). A study by Mahony *et al* (1985) investigating the storage of specimens at various temperatures (4°C, -23°C, -70°C and snap freezing in a dry ice acetone bath) supported the findings of my 4°C and the -20°C stability study. Their results showed that recovery of *Chlamydiae* from specimens stored at 4°C quickly decreased with only one of the specimens yielding chlamydia after seven days and none of the specimens yielding any *Chlamydiae* after eight days. Storage at -23°C did not improve the isolation rates of *C. trachomatis* and isolation rates from specimens were at 0% after 2 days. Good recovery rates were only achieved after the samples were slowly frozen to -70°C or snap frozen at -70°C. With even better results recorded when samples were stored in liquid nitrogen, here no decrease in isolation rates were observed (Mahony and Chernesky, 1985).

Other groups have experienced difficulties in isolating chlamydia using cell culture from samples stored at 4°C and -20°C but were successful in isolating *C. trachomatis* from samples stored at -70°C. The next aim was to attempt to isolate chlamydia from transport media stored at -70°C, to prove that chlamydia can be cultured from the VCM transport media and that chlamydia is stable in this transport media after various periods of time at -70°C. The experiment was conducted as for the previous study (Section 4.5.1). However, this time the spiked samples were stored at -70°C. There was also a set of samples stored at -70°C which contained an equal amount of 4SP as the transport media (1ml) to give a total amount of 2mls of liquid in each tube. This second set of swabs containing 4SP were included to determine if the addition of 4SP would improve the recovery rate of chlamydia. Dilutions were not conducted before storing the swabs at -70°C (to reduce wastage of swabs if the experiment did not work, as was the case for the 4°C and -20°C stability study), instead 2.5 x 10^5 ifu/ml of an *omp*A genotype E/Bour isolate was added to the transport media containing 1ml of media as it had already been predetermined that this concentration of EBs was enough to infect at least 90% of the cells in a 96 well plate. One tube of transport media for each time point containing EBs, with and without 4SP was stored at both temperatures. There were three time points chosen post spiking of the swabs, these were 7, 14 and 28 days. These time points were chosen on the basis that a full working week was required to infect cells and for inclusions to reach maturation. Cells were seeded on a Monday and were ready to be infected on a Tuesday. The inclusions took at least 48 hours to reach maturation and therefore would have reached this point on the Thursday of the week the cells were infected. The

upper limit was chosen as 28 days as the plan was to not exceed the period of 28 days between storing the samples, collection of the samples and infecting McCoy cells with each sample. Additionally, storage conditions were limited in Brighton and only a small -70°C freezer could be placed in the clinic in Brighton due to space restrictions. Therefore swabs had to be collected frequently.

4.5.2.1 Results of the -70°C stability study.

Figure 4.2 shows the $\log_{10}$ graph of the inclusions per ml after storage. As can be seen from the graph there are differences in recovery when 4SP is added to the transport media. However, stability is maintained over 28 days and perhaps longer. The results indicated that even after 28 days at -70°C chlamydia can still be isolated from the transport media and there is no decrease in infectivity. However, with the addition of 4SP there is about a 10-fold decrease in infectivity. Therefore, the addition of 4SP to the media does not have beneficial effects on the recovery rate. These findings had significant implications on the study as a -70°C freezer was purchased for storage of the samples and placed in the GUM clinic for the duration of the study.

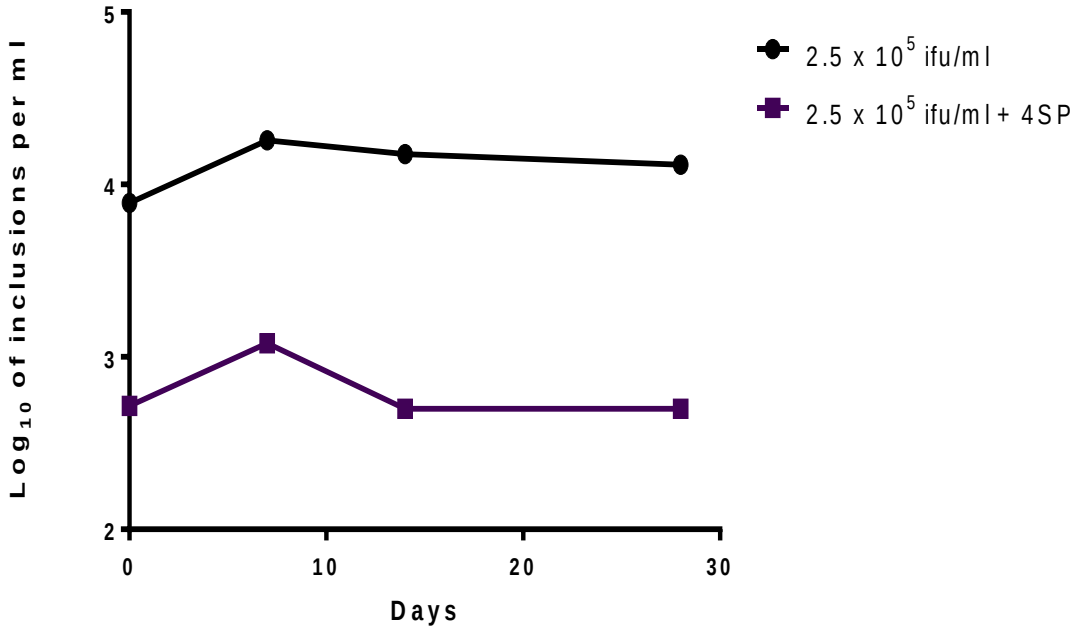


Figure 4.2 Survival of *C. trachomatis* over a 28 day period.
The Σ-VCM transport medium and swab system was spiked with an isolate of chlamydia genotype E/Bour and stored at -70°C over the 28 day period.

4.6 TYPING OF SAMPLES FROM MSM IN BRIGHTON.

Before the commencement of this study a set of storage conditions for culturing and typing of samples from the VCM and the BD ProbeTec™ transport media respectively were determined (Section 4.5). As the samples were transported from Brighton to Southampton for typing on a monthly basis, it was not always possible to follow the guidelines outlined by the manufacturers for the BD ProbeTec™swabs (Section 4.4.1). The samples were initially stored at 4°C following routine diagnostic tests as the laboratory was in a different building to the clinic and there was no access to a -70°C freezer in the laboratory. Subsequently, the swabs were collected by a consultant who then stored the swabs at -70°C, before they were transported to Southampton. It was not until ethical permission was granted for the collection of an extra BD ProbeTec™ swab (due to the rise in asymptomatic LGV which meant that swabs were sent elsewhere before they were received in Southampton), that these swabs were stored immediately at -70°C as they were collected solely for the purpose of this study.

4.6.1 Methods used to assign genotypes to the samples.

Following DNA extraction using a DNA extraction kit (Section 2.11), samples were assigned MLVA-*omp*A genotypes. The PCRs were conducted using a high fidelity Taq DNA polymerase called Phusion Taq DNA polymerase (New England biolabs UK, Hitchin, UK) which ensured that the sequences were accurate due to the low error rate of this enzyme. Samples were typed using the MLVA-*omp*A typing system (Pedersen *et al.*, 2008). O*mp*A sequences were aligned with known *omp*A reference sequences from the NCBI website using BLAST searches and types were assigned to the three VNTR loci using the rules outlined in Pedersen *et al* (2008). Types for each VNTR are shown in table 4.1.

Amplification of target regions in the swab and urine samples were initially conducted with the primers described by Pedersen *et al* (2008) and Wang *et al* (2011b). However, to gain a higher percentage of samples with types for all four MLVA-*omp*A markers, new primers were designed. The primer sequences for the new primers are outlined in section 2.12.6–Table 2.3. Primers were designed for the three MLVA markers- CT1335, CT1299 and CT1291 using Seqbuilder (Section 2.15.2.2). The new set of primers were designed on either side of the initial forward and reverse primers, to hybridise outside the original amplicon sequence of each locus so that it could be determined if there were any mutations in the original primer region which would prevent the primers from binding and therefore amplifying the regions for typing purposes. The new primers were between 18-30 bases in length with a T_M of 55°C, so that the primers were long enough to increase specificity but not too long to allow the primers to bind easily at the annealing temperature. The G/C concentration of the primers were between 40-60% for each primer as a

higher G/C concentration allows for stronger bonds between the primer and the template. The PCR conditions remained unchanged to what they were previously (Section 2.12.6) with a T_M of 55°C for all the primers. PCR of locus CT1291 produced two different sized amplicons (~225bp and 510bp) as seen in previous work where this typing method has been used (Pedersen *et al.*, 2008, Wang *et al.*, 2011b).

For a sub-set of samples a full MLVA-*omp*A profile was only possible once the chlamydia had been isolated in cell culture (Section 4.11), and had been passaged several times to increase the amount of DNA that could be extracted for PCR. This approach has previously proven to be successful (Wang *et al.*, 2011b) and chapter three shows the loci are stable through passaging, therefore, this approach is viable. New VNTR variants were also found during this study and were added to the existing panel of variants (Table 4.1), already published (Wang et al, 2011, Pedersen et al, 2008). Further analyses were conducted on samples where all four loci were successfully amplified, sequenced and given a full MLVA-*omp*A type. Samples with an incomplete MLVA-*omp*A profile were excluded from further analyses as minimum spanning trees could not be generated using incomplete data and comparisons could not be made.

A

VNTR	Variant number of VNTR type and corresponding repeat sequence
CT1335	1 9T8A
	2 10T7A
	3 10T8A
	4 11T7A
	5 11T8A
	6 12T7A
	7 12T8A
	8 13T7A
	9 13T6A
	11 7T9A
	12 8T9A
	13 9T9A
	14 14T7A
	15 11T9A

B

VNTR	Variant number of VNTR type and corresponding repeat sequence
CT1299	1 7C
	2 8C
	3 9C
	4 10C
	5 11C
	6 12C
	7 13C
	8 14C
	9 6C
	10 15C
	11 16C
	12 18C
	13 21C

C

VNTR	Variant number of VNTR type and corresponding repeat sequence
CT1291	1 6C
	2 8C
	3 9C
	4 10C
	5 11C

Table 4.1 Variant numbers assigned to the three VNTR loci (CT1335, CT1299, CT1291).
The table includes variant numbers from Pedersen *et al* (2008), Wang *et al* (2011b) and new variants found in this project.
A Shows the variant numbers corresponding to the VNTR loci CT1335, B Shows the variant numbers corresponding to the VNTR loci CT1299, C Shows the variant numbers corresponding to the VNTR loci CT1291.

97 | Page

4.6.2 Genotyping results.

4.6.2.1 *omp*A distribution.

In total there were 112 positive samples, 85 (76%) of which were assigned full MLVA-*omp*A profiles. All of the other samples had types assigned to at least two markers. Amongst the 85 fully genotyped samples there were 66 different MLVA-*omp*A sequence types.

All of the genital tract *omp*A genotypes apart from *omp*A genotypes H and K were detected in this sample set. The most prevalent *omp*A genotypes were G, E and D comprising of 25.7%, 22.77% and 21.78% respectively. There were two D subgroups: D/UW-3Cx and D/IC-CAL8 which comprised of 5.94% and 15.84% respectively. The distribution of *omp*A genotypes is shown in figure 4.3.

4.6.2.2 MLVA-*omp*A distribution.

In most cases the MLVA types were unique to the *omp*A types. However, eleven MLVA signatures were associated with more than one *omp*A type. These MLVA types were; 3.2.4, 3.5.2, 3.5.3, 3.6.4, 8.2.4, 8.5.5, 8.6.1, 8.6.3, 8.7.2 (Table 4.2). The MLVA-*omp*A profile most commonly detected was 8.5.1-E-Bour which was typed in five patients. MLVA type 8.5.1 was also the MLVA type found in the most *omp* types as it was typed in four *omp*A types. The distribution of MLVA-*omp*A genotypes is shown in figure 4.4.

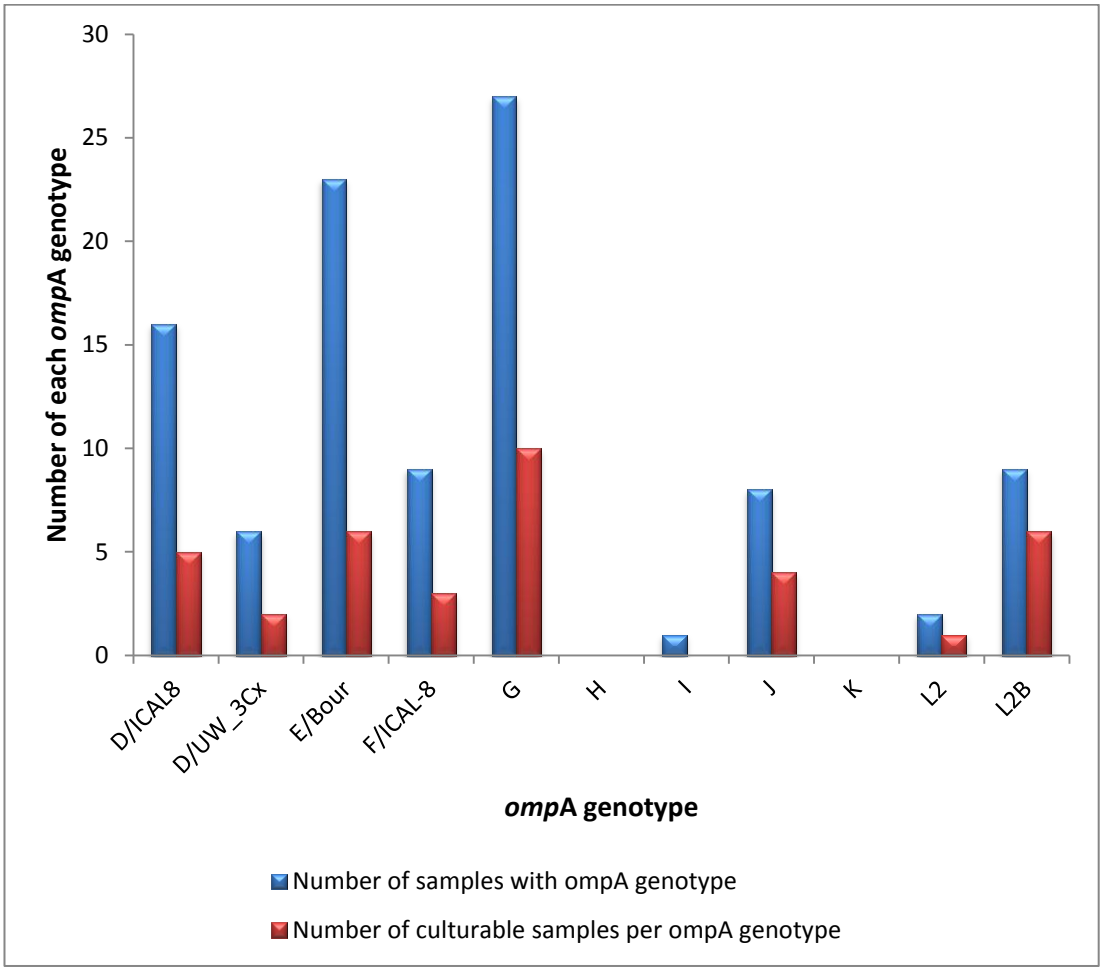


Figure 4.3 The distribution of each *ompA* genotype from MSM in Brighton (Blue bar) and the number of samples isolated in cell culture (Red bar) for each *ompA* genotype.

	omp A genotype												
	D/IC-CAL8	D/UW-3	E/Bour	F/IC-CAL8	G	H	I	J	K	L2	L2b	No. of samples with this MLVA type	No.of ompA genotypes with this MLVA type
					1.3.3 (1)							1	1
											1.5.2 (1)	1	1
											1.5.3 (1)	1	1
								1.6.3 (1)				1	1
											1.9.1 (1)	1	1
											1.9.2 (3)	3	1
										1.9.4 (1)		1	1
					2.5.3 (1)							1	1
					2.5.4 (1)							1	1
								2.7.3 (1)				1	1
					2.11.4 (1)							1	1
					3.2.4 (1)			**3.2.4 (1)**				2	2
					3.3.3 (1)							1	1
					3.4.3 (2)							2	1
								3.5.2 (1)			**3.5.2 (1)**	2	2
					3.5.3 (1)			**3.5.3 (1)**				2	2
					3.6.3 (2)							2	1
	3.6.4 (1)	**3.6.4 (1)**			3.6.4 (1)							3	3
			3.7.2 (1)									1	1
					3.7.3 (3)							3	1
	3.8.1 (1)											1	1
MLVA type	**6.13.3 (1)**											1	1
			8.2.1 (1)									1	1
		8.2.4 (1)		8.2.4 (1)								2	2
			8.3.1 (1)									1	1
			8.3.2 (1)									1	1
								8.3.4 (1)				1	1
		8.4.1 (1)										1	1
				8.4.5 (1)								1	1
	8.5.1 (3)		**8.5.1 (5)***	**8.5.1 (3)#**	**8.5.1 (1)**							12	4
					8.5.2 (1)							1	1
			8.5.3 (2)									2	1
					8.5.4 (1)						8.5.4 (1)	2	2
					8.5.5 (2)					8.5.5 (1)	8.5.5 (1)	4	3
		8.6.1 (1)	**8.6.1 (1)**									2	2
					8.6.2 (1)							1	1
		8.6.3 (1)			8.6.3 (1)							2	2
	8.6.4 (1)											1	1
	8.6.5 (1)											1	1
			8.7.1 (3)									3	1
	8.7.2 (1)		8.7.2 (1)									2	2
				8.7.3 (1)								1	1
			8.7.4 (2)									2	1
							8.7.5 (1)					1	1
	8.8.1 (1)											1	1
				8.8.2 (1)								1	1
	8.8.3 (1)											1	1
			8.9.1 (1)									1	1
								8.9.2 (1)				1	1
			8.9.3 (1)									1	1
	14.7.3 (1)											1	1
No. of samples with this MLVA-ompA type	12	5	20	7	22	0	1	7	0	2	9	85	
No. of MLVA types per ompA genotype	10	5	12	5	17	0	1	7	0	2	7		

Table 4.2 Distribution of different MLVA types according to *omp*A genotype (total of each type).
Coloured in grey are the MLVA-*omp*A genotypes of samples isolated in cell culture.
*One isolate was cultured with this MLVA-*omp*A type.
Two isolates were cultured with this MLVA-*omp*A type.

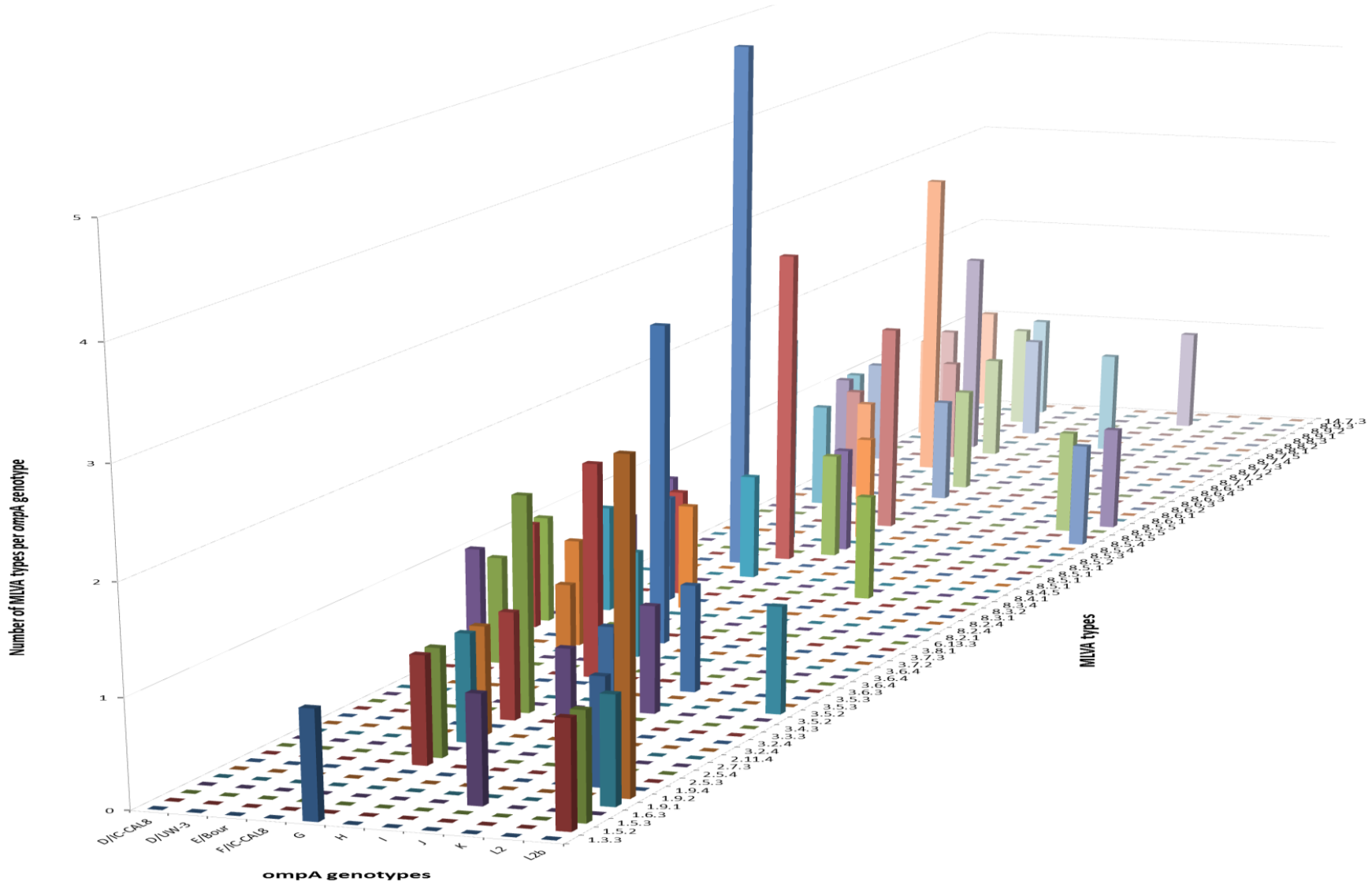


Figure 4.4 Graph showing the distribution of MLVA types according to *ompA* genotype.

4.7 SIMPSONS INDEX OF DIVERSITY.

The Simpson index of diversity (Hunter and Gaston, 1988, Stephens *et al.*, 2001) was calculated using the formula outlined in section 2.16.1. The data obtained from this sample set was used to determine the level of discrimination offered by the MLVA-*omp*A typing system and to see if the addition of three markers increased the discriminatory power compared to using just *omp*A as a typing system. Firstly, the diversity of each of the markers individually was assessed, this was then followed by assessing the diversity of the three MLVA markers individually and then calculating the diversity of the three MLVA markers when used in a typing system together. Lastly, the diversity of the MLVA markers in conjunction with *omp*A was calculated. The following values were obtained: The marker CT1335 had a diversity of 0.61, CT1299 had an index of diversity of 0.84, CT1291 had an index of diversity of 0.78 and *omp*A had an index of diversity of 0.84. The three MLVA markers together had an index of diversity of 0.97 when used together as a typing system for this samples set. The MLVA-*omp*A typing system had a Simpsons index of diversity of 0.99 for this sample set. This shows that the highest discriminatory power is offered when all four markers are used in this typing system.

4.8 TYPEABILITY OF MARKERS.

The typeability of the markers, which is the ability of the marker to assign a type to all samples was calculated by expressing as a percentage the number of typed samples over the total number of samples. The typeability for each one of the markers with respect to this samples set were as follows: CT1335= 99%, CT1299 = 87%, CT1291 = 84% and the typeability for *omp*A was 86%. This shows that there were more samples where the marker CT1335 could be amplified than the marker CT1291 which offered the lowest typeability for these samples.

4.9 MINIMUM SPANNING TREE.

A minimum spanning tree (MSpT) (Figure 4.5) was generated for the samples with full MLVA-*omp*A profiles using Bionumerics 6.5 (section 2.17). The circles on the MSpT were coloured coded according to the *omp*A genotype. Samples that only differed by one locus were linked by black lines on the MSpT. Samples where the MLVA-*omp*A genotype differed by more than one locus from any of the samples in MSpT can be seen around the MSpT as singletons and are not linked to the rest of the MLVA types. The MLVA-*omp*A types in this study formed a single network where two segments could be identified. These segments were labelled i and ii on figure 4.5 and the two segments are divided using a black line. Segment i consisted of isolates which predominately had genotypes D, E and F (which are the *omp*A genotypes found mainly in the heterosexual population), whilst the samples with *omp*A genotype D, G, J, L2 and L2b formed segment ii (which are the *omp*A genotypes found mainly in the MSM population). Pearson's Chi squared analysis

showed that this difference in *omp*A genotypes between segments i and ii was significant with a P value of $P<0.001$ (a P value of $P\leq0.05$ was considered significant).

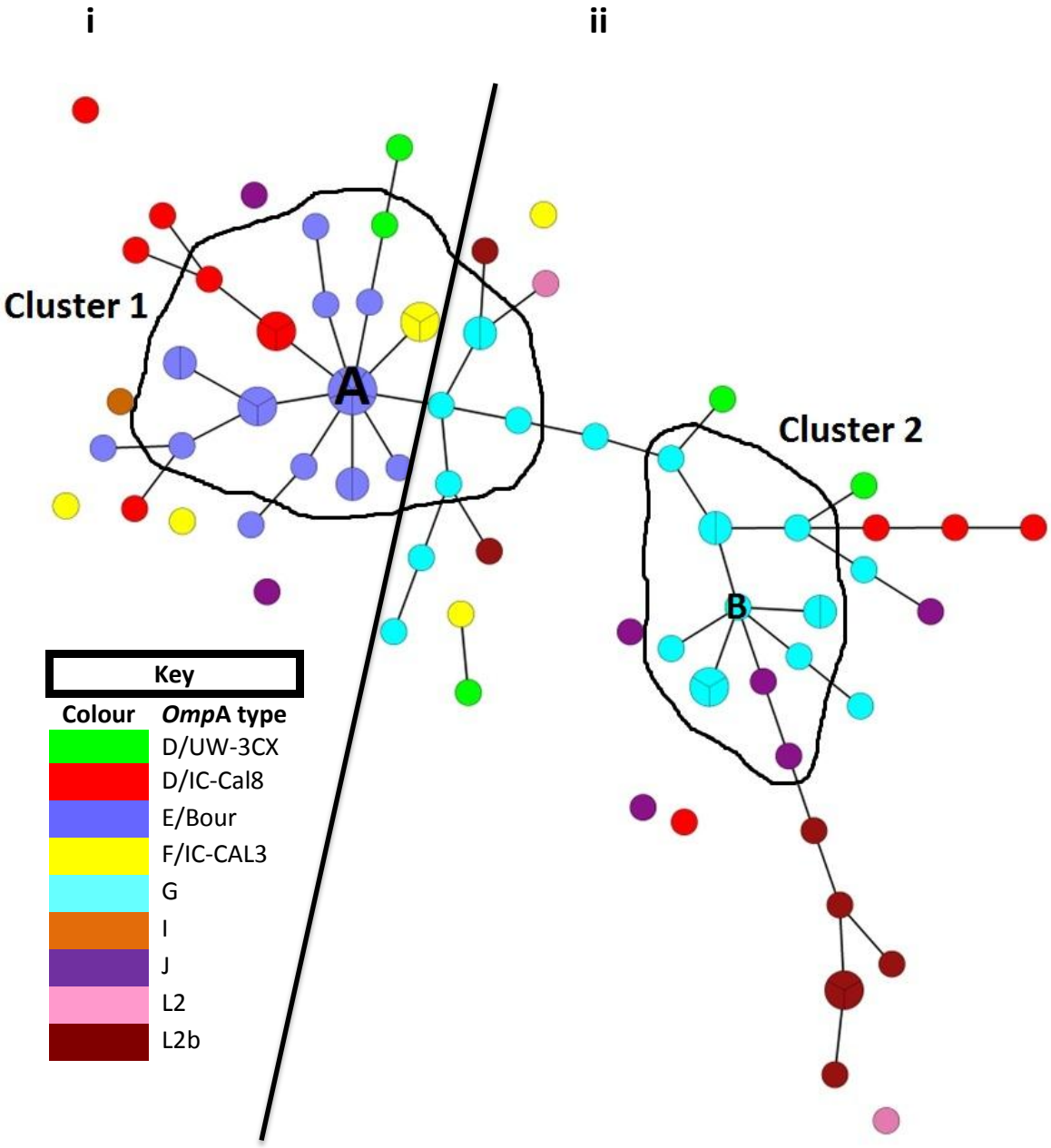


Figure 4.5 Minimum spanning tree using the MLVA data from samples from MSM.
The size of the spheres are indicative of how many samples are in each sphere. All samples within a sphere have the same MLVA-*omp*A type. The branches show single locus variants. The black line splitting the figure in two shows the two segments – I and ii described in section 4.9. Two clusters where identified (A and B) using user defined rules.

4.10 COMPARISON OF *OMP*A AND MLVA-*OMP*A GENOTYPES WITH HIV STATUS, SYMPTOMOLOGY, CO-INFECTION, PREVIOUS *C. TRACHOMATIS* AND OTHER STI INFECTIONS.

Comparison of *omp*A and MLVA-*omp*A genotypes with various factors (HIV status, symptomology, previous infection and coinfections) was conducted to determine if there were certain types associated with these factors. The initial analysis was conducted by determining if there were *omp*A genotypes associated with each factor and then analysis was conducted to determine if there were MLVA-*omp*A types associated with the factors. To determine if the differences were significant the Pearson's Chi squared test or the Fishers exact test were used where appropriate (Statistical analysis was conducted by a medical statistician, Stephanie Goubert in Brighton).

As there were 66 different sequence types for this study, to compare the MLVA-*omp*A types for each of the demographic factors two possible subgroups were identified (labelled A and B on figure 4.5) on the basis that they had five or more SLVS associated with them. Only single locus variants (SLVS) and double locus variants (DLVs) of the subgroup founders A and B were included in further analysis. The two clusters were numbered 1 and 2 (figure 4.5). These two clusters were interesting for comparison purposes to determine if there were any differences between the samples in both clusters of the MSpT. For instance samples from patients in cluster 1 may all be from asymptomatic patients, whilst in cluster 2 the patients may all be symptomatic and this may indicate groups of chlamydia MLVA-*omp*A types which are associated with presentation of symptoms.

4.10.1 HIV status.

There were 66 HIV negative and 46 HIV positive MSM in this study. Figure 4.6 is an MSpT showing samples which are HIV+ (green) and those that were HIV- (red). The *omp*A genotypes for individuals who were HIV negative and those who were HIV positive were compared to test the hypothesis that there are certain *omp*A genotypes associated with HIV status. The most common *omp*A genotypes were G, D and E for both men with and without HIV. Chi squared analysis showed that there was no significant difference between HIV status and *omp*A genotype (P=0.672).

Statistical analysis conducted using the Fisher exact test to determine if there was any significant difference between HIV status of samples from those in cluster 1 compared to those in cluster 2 showed that there was a significant difference (P=0.025). Cluster 1 had predominately HIV negative men and cluster 2 consisted predominately of HIV positive men.

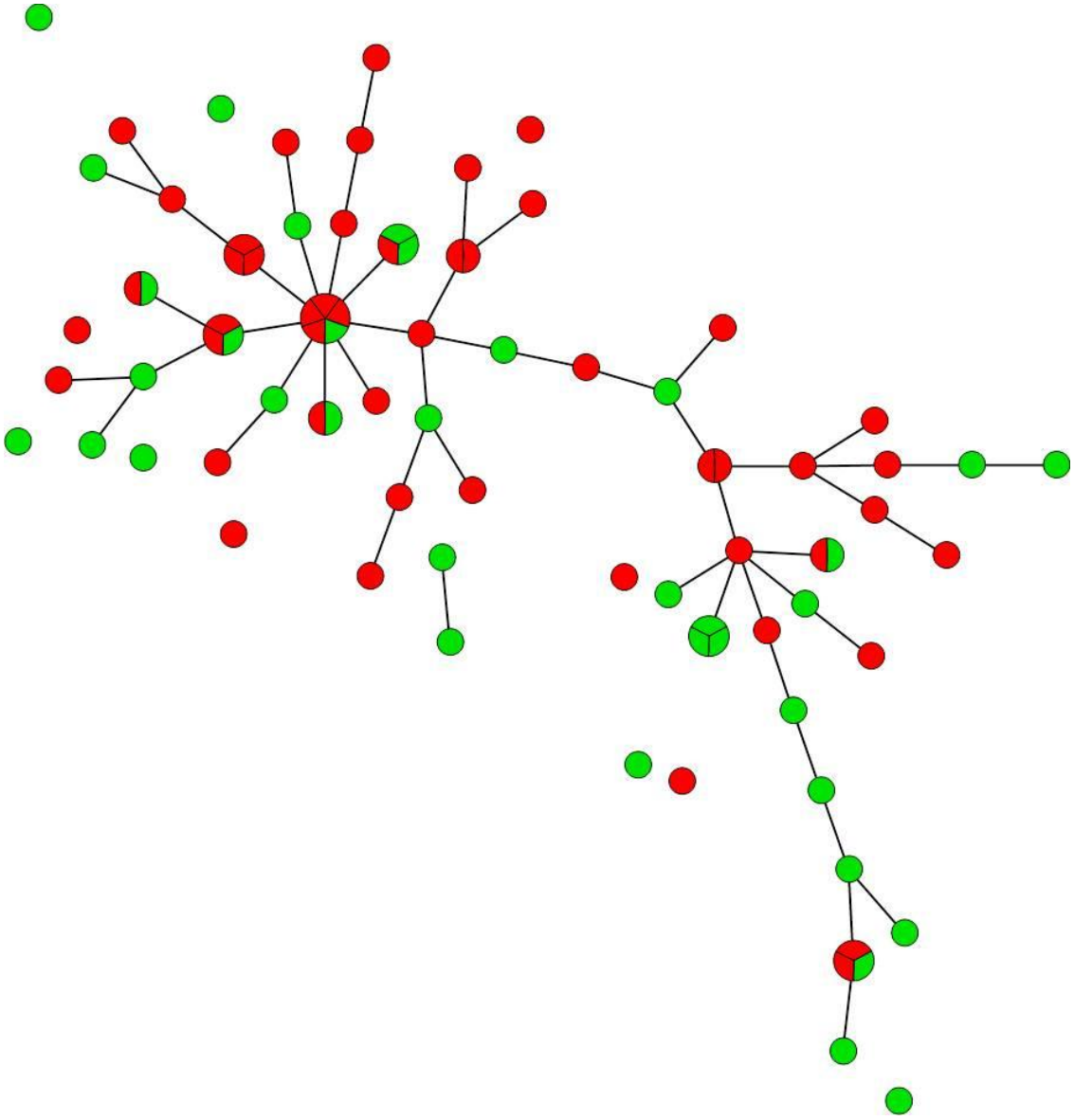


Figure 4.6 Minimum spanning tree showing samples from individuals who are HIV+ and those who are HIV-.
Samples from HIV+ MSM are shown in green. Samples fron HIV- MSM are shown in red.

4.10.2 Symptoms and signs.

There was no significant difference between *omp*A genotypes and urethral symptoms but there were significant differences between rectal symptoms and *omp*A type (P=0.004). Additionally, there were significant differences between some of the *omp*A genotypes and individuals having both rectal and urethral symptoms (P=0.002). Individuals with *omp*A genotypes E, G , I L2/L2b were more likely to have both rectal and urethral symptoms. Overall 75% of individuals were asymptomatic with regards to rectal symptoms. Individuals with *omp*A genotype L2/L2b were more likely to be symptomatic (64% of individuals with this genotype were symptomatic) than any of the other *omp*A genotypes indicating possible trends.

Further analysis conducted using the Fishers exact test showed that there was no statistical significant difference between having rectal symptoms (P=0.725) or signs (P=1.000) and the MLVA-*omp*A types in cluster 1 and 2.

4.10.3 Previous CT and previous STIs.

Interestingly, just over half of the participants (51.4%) had never had chlamydia before but the majority, 72.6% had previously been diagnosed with other STIs. However, there was no significant difference between *omp*A genotypes and previous chlamydia infection (P=0.386) or previous STIs (P=0.127).

Further analysis conducted using the Fishers exact test showed that there was no statistical significant difference between previous infection with chlamydia (P=0.199) or any other STI (P=1.000) and the MLVA-*omp*A types in cluster 1 and 2.

4.10.4 Co-infection.

At the time of visit to the clinic 21.20% of the individuals were co-infected with other STIs. The most common STI was *N. gonorrohoeae* followed by the herpes simplex virus. There was no significant difference between *omp*A genotype and whether the individual was co-infected with another STI (P=0.179). Neither was there any significant difference between *omp*A and the type of STI the individuals were co-infected with (P=0.350).

4.11 PARTICIPANTS WITH BOTH URETHRAL AND RECTAL INFECTIONS.

Rectal swabs and urine samples were collected at the time of examination from all patients. Urine samples were only sent to the lab in Southampton if they were positive and if the patients also had a rectal swab which was positive by NAAT testing. The purpose of collecting both the rectal swabs and the urine sample, if available, was to measure the concordance of types in men infected at both sites. In this study there were seven participants who were infected at two anatomical sites – the rectum and the urethra (Table 4.3). Interestingly, none of the patients in this study had concordant genotypes for both sample sites. This strongly suggests that these

patients have been infected by more than one partner. For four out of seven of these participants infected at both sites the typing data is incomplete and therefore it was difficult to draw any conclusions as to if these samples have the same types. For the other three participants whose urine and rectal samples had full MLVA-*omp*A types one sample differed in one locus and the other two differed in two loci. Four out of seven urine samples could not be fully genotyped compared to only two out of the seven swab samples with a paired urine sample. Table 4.3 shows the anonymised sample number and the differences between both samples for each of the participants infected at both sites. Also included in the table is the HIV status of the participants who had both urethral and rectal chlamydia. Two of the participants were HIV+ whilst the other five participants were HIV negative.

Sample number	Urine or swab	CT1335	CT1299	CT1291	*Omp*A	Complete/Incomplete MLVA-*omp*A type	HIV Status
006	**Swab**	13T7A	11C	11C	L2/434/BU	**Complete**	**Neg**
	Urine	13T7A	11C	6C	E/Bour	**Complete**	
046	**Swab**	13T7A	12C	10C	D/IC-CAL8	**Complete**	**Pos**
	Urine	13T7A	**-**	**-**	**-**	**Incomplete**	
050	**Swab**	10T8A	11C	9C	G/9301	**Complete**	**Neg**
	Urine	10T8A	12CT3C	9C	G/9301	**Complete**	
051	**Swab**	12T7A	-	-	D/IC-CAL8	**Incomplete**	**Neg**
	Urine	13T7A	-	-	D/IC-CAL8	**Incomplete**	
053	**Swab**	13T7A	-	9C	K/UW3/CX	**Incomplete**	**Neg**
	Urine	13T7A	-	9C	-	**Incomplete**	
067	**Swab**	13T7A	13C	6C	E/Bour	**Complete**	**Pos**
	Urine	13T7A	12C		**-**	**Incomplete**	
079	**Swab**	10T8A	12C	10C	D/IC-CAL8	**Complete**	**Neg**
	Urine	10T8A	12CT3C	10C	G/9301	**Complete**	

Table 4.3 Typing data from participants who gave a swab and a urine sample.
Included in the table are the sequences for each of the VNTR and *omp*A loci for both the urine and swab samples of participants who had both urethral and rectal chlamydia.

4.12 ISOLATION OF *CHLAMYDIA TRACHOMATIS* FROM SAMPLES COLLECTED FROM MSM IN BRIGHTON.

4.12.1 Methods used for isolation of *C. trachomatis* from the VCM swab and transport media system.

Isolation of rectal strains of *C. trachomatis* using cell culture was carried out to subject the different types of chlamydia in this study to the most detailed phenotypic analysis possible and also to allow the possibility of whole genome sequencing which could reveal whether the rectal strains had any specific differences related to their tropism.

Samples used for isolation of *C. trachomatis* were collected from patients using swabs and these swabs were immediately inserted into the relevant transport medium. Due to the fact that the -70°C freezer was not in the same room as the consultation room due to lack of space, the noise and heat created by the freezer, the swabs were initially placed in a 4°C fridge which was available in the consultation room until the end of the GUM clinic day which was no longer than an eight hour period. Subsequently, the swabs were stored at -70°C which had been shown to be the optimum storage temperature for sample isolation and viability remained stable up to 28 days at this temperature (Section 4.5.2). The samples were collected from Brighton and transported to the Molecular Microbiology laboratory in Southampton for the isolation of *C. trachomatis* before the 28 day period had elapsed. To ensure the cold chain was maintained the samples were transported on dry ice, so that they remained frozen. This was essential to ensure the cells remained intact as freeze-thawing of the samples leads to ice crystals forming within cells which could lead to damage of the cell membrane (Dumont *et al.*, 2006).

Initially the experimental design for isolating chlamydia from the clinical samples was to allow the samples to thaw, vortex the sample and then add 1ml of the transport media to cells in one well of a 24 well tray. This approach was initially taken as it had proved successful in another study carried out in Molecular Microbiology Department, Southampton (Wang *et al.*, 2011b). However, Wang *et al* were isolating chlamydia from genital tract samples and not from samples that had been collected from the rectum. For this study this experimental design failed to work effectively. The samples proved to be toxic to the tissue culture cells as they frequently contained faecal matter which completely destroyed the monolayer of cells and turned the media yellow which showed the sample was acidic.

The experimental design was refined following consultation with colleagues (Prof. Bertille De Barbeyrac and Dr. Kenneth Persson, personal communication), who also had difficulty with culturing *C. trachomatis* from rectal swabs. The refinements to culture following their advice were

as follows; after the samples were vortexed they were then transferred into a 20ml tube and centrifuged at 110xg for five minutes. This step was added to concentrate the large particles in the sample and to ensure that the elementary bodies remained in the supernatant. Each sample was then diluted 1/10 using DMEM media and placed on a monolayer of cells, initially in a single well of a 24 well tray. Cells were infected with the sample in diluted transport media and incubated for 2-3 days after which the samples were harvested (Section 2.7). The cryopreservative, 4SP was subsequently added to the harvested sample which was then stored at -70°C.

4.12.2 Results of the isolation of *C. trachomatis* from positive rectal samples collected from MSM in Brighton.

In total 112 samples were collected for this study and all 112 samples were taken through at least three passages in cell culture to determine if the chlamydia in the sample could be isolated in McCoy cells. This procedure was used because it had previously been shown that blind passaging increases recovery rates of *C. trachomatis* from urogenital specimens (Jones *et al.*, 1986) and therefore this may also be true for rectal chlamydia specimens. If, after three passages, no inclusions were seen when the cells were visualised under a light microscope it was concluded that these samples were culture negative. If inclusions were seen then passaging was continued until the isolate was able to infect 100% of the cells in a T25 flask and therefore it would proliferate successfully if required for further research purposes. The main problem associated with isolation of *C. trachomatis* from samples was that they were mostly highly contaminated with faecal matter and/or sometimes blood.

33% of the samples (37 culture positive samples out of a total of 112 samples) were observed to produce inclusions when McCoy cells were infected with the sample. Of the 37 viable isolates the distribution of *omp*A genotypes were as follows; G (n=10), followed by E (n=6), D (n=8), and L2B (n=6) (table 4.4). Furthermore, four genotype J and three genotype F/ICAL-8 isolates were obtained. There were no genotype H or K isolates as none of the samples were typed as having the *omp*A genotype H. Additionally, there were no isolates with an *omp*A genotype I or K isolated from any of the samples. Figure 4.7 shows an MSpT of the MLVA-*omp*A types of culturable and non-culturable isolates. There were some MLVA-*omp*A types with multiple samples which could not be cultured. Additionally, there were some MLVA-*omp*A types with multiple samples where none of the samples were isolated in cell culture. There were no large clusters of MLVA-*omp*A types of samples which could or could not be isolated in cell culture.

Table 4.4 shows the MLVA-*omp*A types of the 37 isolates. Most of them had a unique MLVA-*omp*A type. However, the MLVA-*omp*A types 3.4.3-G, 8.5.1-F, 8.7.1-E and 1.9.2-L2b had multiple culture positive samples with these types. Two culture positive samples had the MLVA-*omp*A type 3.4.3-

G, two samples had the MLVA-*omp*A type 8.5.1-F, three samples had the MLVA-*omp*A type 8.7.1-E and three samples had the MLVA-*omp*A type 1.9.2-L2b. Two of the samples with MLVA-*omp*A type 1.9.2-L2b were collected on the same day which could mean that these samples were collected from partners, or possibly indicating part of a network.

The morphology of the inclusions formed by 36 of the 37 samples that were isolated all seemed to be typical when observed under a light microscope as there was only one inclusion per cell. However, the inclusions produced by one sample, which was called Bri 008, were different to what is usually seen in *C. trachomatis* infected cells. The morphology of the inclusions produced by this isolate were unusual as there appeared to be more than one inclusion forming per cell (Figure 4.8a). This phenomenon of non-fusogenic inclusions was first described by Suchland *et al* (1999) who attributed this formation to mutations within the *inc*A gene. Figure 4.8 shows the image of the inclusions formed from infecting cells with sample 008 after five passages in cell culture, there are multiple inclusions within a single cell. Figure 4.8b shows the phase contrast image of the usual infection seen when *C. trachomatis* infects McCoy cells, where one inclusion is formed per cell. Figure 4.8c shows a control well where none of the cells have been infected with chlamydia.

Further investigations were required to determine whether the inclusions were non-fusogenic inclusions or whether this phenomenon was due to inclusion splitting, therefore a dilution series was set up. The idea behind this experiment was that if this phenomenon was actually due to non-fusogenic inclusions then at a dilution where there was only one chlamydia per cell, i.e., where the multiplicity of infection (MOI) was one there should only be one inclusion per cell. However, if there is more than one inclusion per cell at an MOI of one then this phenomenon is not due to inclusion non-fusogenicity but due to inclusion splitting. The results obtained from conducting this experiment showed that at an MOI of one there was only one inclusion per cell and therefore it was concluded that the inclusions seen were indeed non-fusogenic chlamydial inclusions.

	MLVA-*ompA* type			
Sample No	**CT1335**	**CT1299**	**CT1291**	**ompA**
083	2	5	4	G/9301
078	2	7	3	J/UW36/Cx
069	2	11	4	G/9301
113	3	4	3	G/9301
100	3	4	3	G/9301
098	3	5	2	L2b
050	3	5	3	G/9301
030	3	5	3	J/UW-3X
102	3	6	4	D/UW-3/CX
028	3	7	3	G/9301
108	3	8	1	D/IC-CAL8
054	3	2	4	J/UW36/Cx
048	3	2	4	G/9301
105	6	13	3	D/IC-CAL8
023	8	5	5	G/9301
008*	8	5	1	E/Bour
026	8	5	1	F/IC-CAL3
055	8	5	1	F/IC-CAL3
084	8	5	1	G/9301
046	8	6	4	D/IC-CAL8
071	8	6	1	D/UW-3/CX
060	8	6	1	E/Bour
021	8	7	1	E/Bour
067	8	7	1	E/Bour
090	8	7	1	E/Bour
107	8	7	2	D/IC-CAL8
103	8	8	2	F/IC-CAL3
106	8	8	3	D/IC-CAL8
095	8	9	2	J/UW36/Cx
037	8	2	1	E/Bour
052	8		2	G/9301
085	1	5	2	L2b
087	1	5	3	L2b
104	1	9	4	L2
038	1	9	2	L2b
086	1	9	2	L2b
088	1	9	2	L2b

Table 4.4 MLVA-*ompA* types of the 37 isolates from rectal swabs.
The samples are ordered according to MLVA-*ompA* type. Highlighted in red, yellow, green and turquoise are four sets of samples with identical MLVA-*ompA* types.
*Non-fusogenic isolate

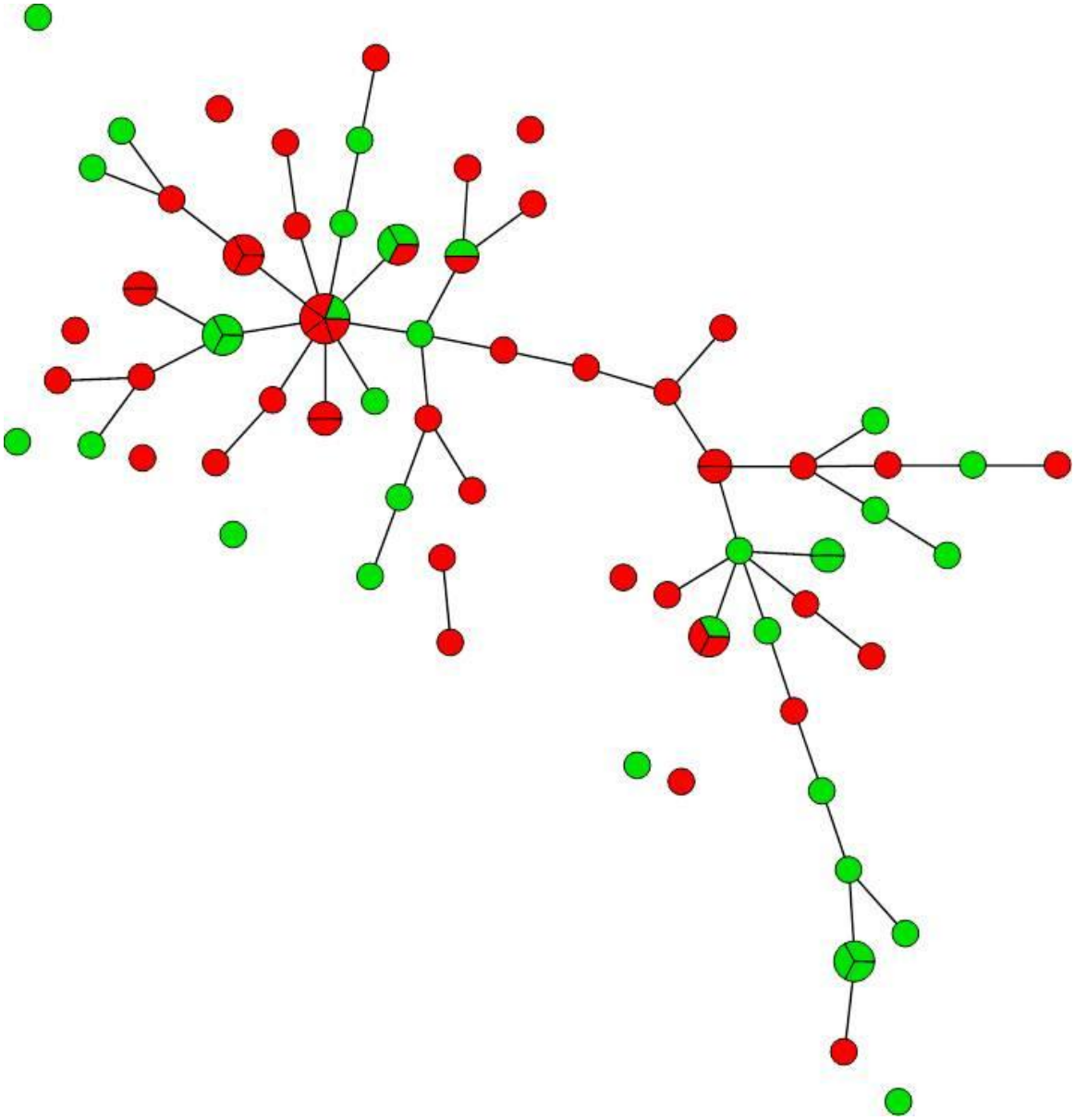
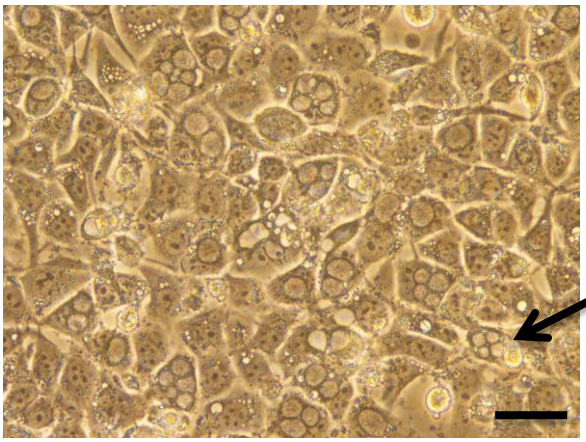


Figure 4.7 Miminimum spanning tree showing culturable and non culturable isolates. Samples which were isolated in cell culture are shown in green. Samples which could not be isolated in cell culture are shown in red.

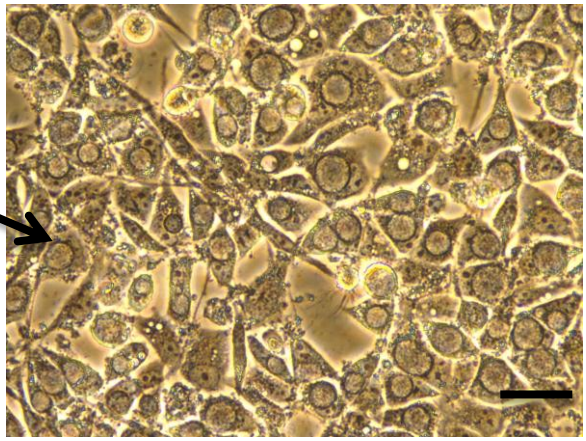


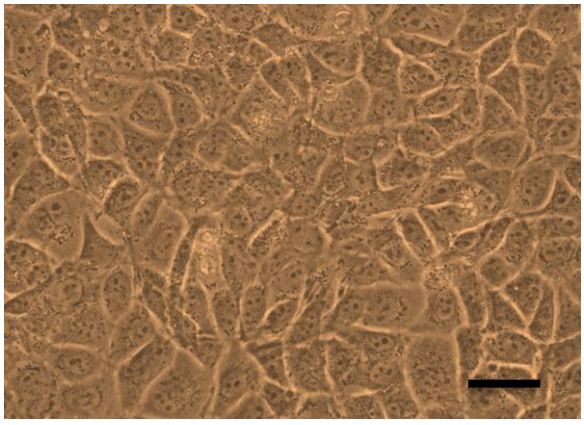
a)

Four inclusions within one cell

b)

One inclusion within one cell





c)

Figure 4.8 Phase contrast microscope image of cells infected with *C. trachomatis* in a single well of a six well plate from sample 008 (a) and sample 078 (b).
The control well (uninfected cells) is shown in (c). Images taken using x40 magnification**.** The scale bar represents 10µm.

4.13 DISCUSSION

To date there are only a small number of studies where high resolution typing systems have been implemented to study the epidemiology of MSM patients (Christerson *et al.*, 2012, Bom *et al.*, 2013a). The identification of genotypes within demographic groups is important when trying to gain a better understanding of the spread of chlamydia as there may be certain genotypes which are more prevalent within specific groups of people. The study of genotypes within various groups of people can explain the differences seen within the community especially if there are certain genotypes which have specific tissue tropism.

Once a suitable transport medium was sourced, which did not contain any antibiotics capable of inhibiting the growth of chlamydia in cell culture, a series of stability studies were conducted, firstly to determine if chlamydia could be isolated in cell culture from these swabs as a study by Mahony and Chernesky (Mahony & Chernesky 1985b), showed that the type of swab used is important as live chlamydia cannot be recovered from certain types of swabs. In their study the best recovery rates were experienced with cotton on aluminium swabs, whilst the calcium alginate swabs were toxic to cell culture. The inclusion recovery from certain types of swab were as low as zero depending on the materials the swabs were made from so it was important to determine if the Σ-VCM swabs which were made from cellular foam were suitable for recovering live chlamydia. Secondly the experiments were conducted to determine the temperature at which the transport medium should be stored at over an extended period of time. Long term storage in the Molecular Microbiology laboratory in Southampton is traditionally in -70°C freezer or in liquid nitrogen. However, these facilities were not available at the GUM clinic in Brighton before the project, therefore it was necessary to conduct a series of experiments to determine the optimal temperature to store the swabs, to increase the chances of assigning genotypes to each sample and to recover live chlamydia. The -70°C stability study (Section 4.5.2) showed that chlamydia can be stored over a long period of time at -70°C and that live elementary bodies could be recovered from the transport media. It also showed that the staff at the clinic do not need to add 4SP to the transport medium when they collect the swabs. In fact, 4SP seemed detrimental (figure 4.2). Therefore, a small under bench -70°C freezer was considered appropriate for this study and was purchased for storage of swabs collected in Brighton (Only a small freezer could fit under one of the benches in the GUM clinic due to the cramped working conditions and the lack of space in the clinic).

Initially, there were difficulties with isolating chlamydia from the rectal samples. One of the problems with obtaining *C. trachomatis* isolates from these swabs was that rectal swabs are likely to be highly contaminated with other bacteria, faecal matter and blood, which destroy the cell monolayer. Under these conditions the media turned yellow due to the acidic conditions which

made it impossible to view the samples under the light microscope. Although, isolation of chlamydia proved to be problematic at first, the issues associated with isolating chlamydia from the samples were overcome by initially centrifuging the sample to pellet any large particles and also by diluting the samples before cells were infected, this increased the chances of recovering live isolates. A method used in another study to increase the chances of recovering live chlamydia from rectal specimens was by sonicating the samples, this reduced contamination by faecal microorganisms and improved sensitivity (Barnes *et al.*, 1987). This approach can be used in future studies to determine if sonication increases the rate of isolation when attempting to isolate chlamydia from rectal samples. In addition, specimens were inoculated within 24 hours after the samples were taken and only 100µl of the specimen was used for inoculation (Barnes *et al.*, 1987). Sonication was not attempted in my study because there was no safe method of sonicating chlamydia at the time of the study. As the EBs were sonicated in a tube which did not have a lid, the apparatus for sonication had to be set up in a class one hood as aerosols may be produced. There was also a chance of spillage with this apparatus, therefore this system was unsafe and it was decided that sonication should only be used if it was absolutely necessary and if no other methods worked well.

Extremely poor recovery rates from rectal samples were expected. The first isolation of *C. trachomatis* from the rectum of a homosexual man occurred in 1977 (Goldmeier and Darougar, 1977). This study showed that isolation of *C. trachomatis,* although difficult, was actually possible. In this paper no details on how isolation of *C. trachomatis* was achieved was described. In a more recent study carried out in 2012, it was reported that analysis of 14 NAATs positive *C. trachomatis* samples from the rectum of MSM were all negative when they tried to isolate chlamydia in cell culture from the samples (Neu *et al.*, 2012). In most studies where the researchers have attempted to isolate chlamydia from rectal samples, the success rate has been between 4 to 8% (McMillan *et al.*, 1981, Munday *et al.*, 1981, Quinn *et al.*, 1981, Munday *et al.*, 1985, Sulaiman *et al.*, 1987). Compared to previous studies, of which, there have only been a small number of culturable isolates, the innovations made to improve recovery of culturable isolates proved extremely advantageous. Out of a total of 112 samples, 37 isolates were successfully cultured, which is an unprecedented rate of recovery from rectal swabs. There has also not been any recent publications on the isolation of *C. trachomatis* from rectal specimens, possibly due to the experimental difficulties associated with isolation of chlamydia from rectal material. This makes this study, not only the most successful study when it comes to isolation of rectal *C. trachomatis* but also the only study at the time this thesis was written where the researchers have managed to isolate so many recent samples from a population of MSM.

Figure 4.3 shows the distribution of each *omp*A type isolated in cell culture in this study and table 4.4 shows the MLVA-*omp*A types of the 37 isolates. Unlike genital tract strains LGV strains do not require centrifugation on to cells. This shows how invasive LGV strains are as they have the capability to grow at ease within cells without the need of having to spin the inoculum on to the cells by centrifugation and this also reflects the ease in which these strains can be obtained. However, it was not possible to culture all of the samples with *omp*A genotype L2 and L2b. Only one chlamydia sample with genotype L2 and six out of nine possible isolates with genotype L2b were isolated in cell culture. This may be due to the level of contamination which was present in the sample which hindered the ability of the sample to infect the cells and produce inclusions. There were some MLVA-*omp*A types with multiple samples which were cultured. This could be indicative of MLVA-*omp*A types of chlamydia which grow well in cell culture. There were also some MLVA-*omp*A types with multiple samples where none of the samples were isolated in cell culture this may be indicative that there may be some MLVA-*omp*A types of chlamydia which are difficult to isolate in cell culture or do not grow at all in cell culture.

One of the chlamydia isolates was rather unusual. When McCoy cells were infected with this isolate, the inclusions proceeded through the developmental cycle but did not fuse, instead multiple inclusions were seen within a cell. This phenomenon is rarely seen and only a small percentage of samples have been reported to possess this mutation. In a 'normal' infection inclusions fuse to produce a single inclusion (Hodinka *et al.*, 1988, Matsumoto *et al.*, 1991). incA proteins are hypothesised to be important in chlamydial pathogenesis (Fields *et al.*, 2002) and are found in the inclusion membrane of *C. trachomatis* (Bannantine *et al.*, 1998). Every non fusogenic isolate that was 'whole genome sequenced' by Suchland *et al* (1999), did not localise incA to the inclusion membrane and therefore this gene could play a role in inclusion fusion. The incA gene of some of the isolates that possess this characteristic were sequenced and compared to wild type *C. trachomatis* and their data showed that there are two amino acid changes within the genome which could be attributed to incA not localising to the inclusion membrane (Suchland *et al.*, 1999, Pannekoek *et al.*, 2001).

A study by Christerson *et al* (2012) showed that full profiles of types could only be obtained in 55% of the samples (out of a total of 203 samples) obtained from the rectum of MSM which they tried to type using a high resolution genotyping method – MLST (Christerson *et al.*, 2012). This shows that it may be difficult to obtain full profiles from rectal swabs, possibly due to inhibitors within the sample that interfere with the PCR reaction. The quality of the samples received for typing purposes for my study was initially an issue as the samples did not produce full profiles when subjected to the MLVA-*omp*A typing scheme. The usual protocol for sample collection at the GUM clinic in Brighton was for samples that are collected from LGV patients and those who are positive

for gonorrhoea to be sent to STBRL, Colindale. The samples were then sent back in the post to the clinic in Brighton. This meant that samples were not immediately stored at -70°C and this may be the reason why initially it was difficult to assign types to all four loci in every sample. Due to an increase in asymptomatic LGV cases, all of the samples collected from GUM clinic attendees and not just samples from individuals who were positive for LGV and gonorrohoea had to be sent to the STBRL. This delayed the time in which swabs were received at the laboratory in Southampton and extended the period of time that the samples were stored at room temperature and decreased the chances that viable DNA would be extracted from the sample to perform the PCR. There were also issues with collecting samples from the diagnostic laboratory in Brighton and some of the samples were stored in the fridge for up to four weeks before being stored at -70°C which rendered them useless for assigning types to the samples. The length of time it took the staff working in the laboratory in Brighton to put the samples in the fridge is unknown which also may have had implications for the study. There were several meetings to address these issues but it was difficult to rectify these issues as there were many different members of staff working in the clinic and the laboratory. Therefore to ensure the samples received were adequate for the study requirements extra swabs were collected for this study.

As problems with the primers could also be the reason for the production of poor quality PCR amplicons or no amplicons, a new batch of the same primers were ordered. However, this did not improve the typing results obtained. The primers and polymerase were also tested on samples in the freezer from previous studies and these samples produced a full MLVA-*omp*A profile. Due to the poor sequencing data initially received using the primers described by Pedersen *et al* (2008), new primers were designed as it was only possible to fully type 30% of samples with the initial set of primers. This was considered unacceptable to achieve the aims of this study and to determine genotypes of chlamydia in MSM in Brighton. Therefore, new primers were designed on either side of the initial forward and reverse primers to increase the number of fully typed samples and this approach was successful as 76% of the samples were assigned full MLVA-*omp*A genotypes. This suggested that there were possibly mutations in the primer binding sites of the samples which initially did not produce a full MLVA-*omp*A profile.

Initially, the intention was not to include samples from patients with LGV in this study but due to the increase in asymptomatic LGV cases it was inevitable that these samples were included. It has previously been shown by another group that men infected with LGV strains were more likely to be co-infected with HIV and other STIs as they had high risk sexual behaviour (Bom *et al.*, 2013b). This particular study also demonstrated that different *C. trachomatis* genotypes were not associated with patient demographics such as age, ethnicity, lifestyle, sexual partners and sexual

behaviour. Interestingly, the paper also mentions that there are no distinct genotypes that are associated with solely MSM or solely heterosexual groups but that all genotypes can be found in both groups of people. Perhaps this is evidence of a bridging population which may identify themselves as bisexual and therefore have sexual relations with both men and women and are the reason why these types are found in both the heterosexual and MSM cohorts. The study carried out in Brighton does not include a heterosexual population for comparison purposes but can be compared to a heterosexual population from Southampton (chapter 5) to understand better the types circulating in the population (chapter 7).

The most prevalent genotypes were (in descending order) G, E and D which differs slightly from the distribution seen in most studies where the most prevalent *omp*A genotypes in MSM were genotypes G, D and J (Christerson *et al.*, 2012, Bom *et al.*, 2013a, Bom *et al.*, 2013b). The findings in the Brighton study are similar to those found in Northern Spain, where the most prevalent genotypes were genotypes E, G and D (Mejuto, 2013). This could be due to an increase in bisexual individuals which act as a bridging population between the heterosexual population and the MSM population. Over time, the most prevalent *omp*A genotypes found in MSM and the heterosexual population may continue to change. Additionally, this may also indicate the UK MSM networks are more closely associated with those of Southern Europe than those of the Northern Europe studies.

Seven participants (6.3%) had *C. trachomatis* positive urine and swab samples at the same visit to the clinic. This percentage is double the percentage of a recent study (Bom *et al.*, 2013a). In the study described in this chapter none of the samples were concordant, due to either incomplete data or the loci having different types between the urine and the rectal samples. This was also an observation made by Bom *et al* (2013a). They found that the paired swab and urine samples had different sequences. Three of the nine participants, who had given both urine and swab samples were in different clusters of the minimum spanning tree and for the remaining five participants the paired samples belonged to the same cluster, but for one of the participants the sequence type was different due to a difference in one of the loci. This is evidence of high risk sexual behaviour as it suggests that there is a percentage of men who have more than one sexual partner. The majority of participants (71%) in my study also claimed to have had unprotected anal intercourse in the past three months before their visit to the GUM clinic which is also indicative of high risk sexual behaviour amongst MSM. In the Bom *et al* (2013a) study they also experienced difficulty in obtaining full MST genotypes for the urine samples and this could be due to how the samples were stored. Urine samples also contain inhibitory material which interferes with the PCR and although a DNA extraction method was used which can remove some of the

inhibitors it does not always work, therefore, the target loci in these samples are not always amplified.

The Simpsons index of diversity which is defined as a method's ability to assign different types to two unrelated strains randomly (van Belkum *et al.*, 2007) was calculated for this sample set and there was a diversity index of 0.99. As a diversity index of one means complete diversity, a value of 0.99 suggests that the genotypes in this study were almost completely diverse (i.e different).

Table 4.2 shows the MLVA types found in more than one *omp*A genotype. The majority of MLVA types were only found in one sample which shows that this typing system does offer high resolution. The most common MLVA-*omp*A type was 8.5.1-E which was typed in five samples. The relationship amongst closely related genotypes are shown using an MSpT (Figure 4.4). There were two large segments identified in the MSpT. One segment consisted of genotypes D, E and F – cluster i, whilst the other cluster consisted mainly of genotypes D, G, J and L2b – cluster ii. The *omp*A genotypes in cluster i were genotypes which are the most prevalent genotypes found in heterosexuals (*omp*A genotype D, E, F) whilst the *omp*A genotypes in cluster ii were the genotypes which are most prevalent in the MSM population (*omp*A genotype D, G, J, L2 L2b). The differences in *omp*A genotypes in both segments was considered significant following Pearson's Chi squared analysis. This *omp*A genotype distribution amongst heterosexual and MSM related samples has been seen in other studies where they analysed samples from heterosexual and MSM populations using the MLST typing scheme (Bom *et al.*, 2011, Christerson *et al.*, 2012, Bom *et al.*, 2013b). In these studies it was apparent that there were clear divisions between the genotypes found in the MSM population and those found in the heterosexual population which formed different clusters in the population.

One of the aims of the study was to determine if there were MLVA-*omp*A genotypes associated with HIV status. Comparison of HIV status between cluster 1 and 2 (figure 4.6) showed that there was a significant difference between HIV status in cluster 1 and cluster 2. Cluster 1 consisted predominately of individuals who were HIV negative, this is also the cluster which contains mainly *omp*A types mainly associated with heterosexuals. On the other hand, cluster 2 predominately contained samples from HIV+ men and these samples were mainly samples with *omp*A types mainly associated with MSM. This data shows that there are closely related MLVA-*omp*A types associated with HIV-status in MSM. This may be because those who have HIV tend to have sexual intercourse with other MSM who have HIV and therefore they pass on certain MLVA-*omp*A types of chlamydia. Likewise, MSM who do not have HIV tend to have sexual intercourse with each other and therefore pass on chlamydia with certain MLVA-*omp*A types in their sexual networks. These data could also be proof that serosorting is occurring in Brighton. These data also support our hypothesis that certain types of chlamydia grow better in HIV+ MSM compared to HIV- MSM,

possibly due to the fact that HIV+ MSM are immunocompromised and HIV- MSM are unlikely to be immunocompromised.

There was no significant difference between previous chlamydia infections, previous infection with other STIs or no significant difference with co-infection with other STIs, this was probably due to the low numbers of samples. However, there were significant differences between some of the *omp*A genotypes and rectal symptoms. Participants with *omp*A genotype L2 and L2b were more likely to present with symptoms than those with other *omp*A genotypes.

4.14 FUTURE WORK.

It will also be interesting to further analyse the sample that has produced non-fusogenic inclusions by firstly performing an incA specific PCR and sequencing the gene to identify any mutations and secondly by whole genome sequencing the isolate and comparing the genome to other genomes and indeed comparing the incA genes between wild type *C. trachomatis* and the non-fusogenic isolate. Comparison of genotypes of chlamydia found in the rectum of men compared to the genotypes found in rectal chlamydia in women can also be conducted using high resolution typing to determine if there are any differences between genders. Sonication of sample could also be attempted on rectal samples to determine if this increases isolation rate from rectal samples.

4.15 CONCLUSIONS.

The GUM clinic staff in Brighton had reported a rectal chlamydia prevalence of about 10% amongst their MSM population who had been tested for rectal chlamydia, a figure which is slightly higher in HIV+ than HIV- MSM (Dr. Suneeta Soni, personal communication). Therefore, based on the frequency of positive *C. trachomatis* patients who attended the clinic it was estimated that it would take nine months to recruit the 200 patients considered necessary for this pilot study. However, as the study progressed it became apparent that it would take longer than nine months to recruit 200 participants for this study. Recruitment for the study began in October 2011 and finished in March 2013. However, despite extending the finishing date it was not possible to obtain 200 samples. In total, 112 samples were collected and not all of them were in pairs, for instance some patient samples only had the Σ-VCM swab and not the BD ProbeTec™ swab and *vice versa*. There are several reasons for the low recruitment numbers. These include participants not consenting to having their samples taken for the purpose of this study, the GUM clinic staff sometimes forgetting to inform the patients of the study and therefore samples collected from these patients could not be included in the study as the patients had not given their consent. Samples were also sometimes misplaced during transportation following diagnostic tests from the laboratory back to the GUM clinic and samples were sometimes misplaced once

they had been returned to the GUM clinic. The numbers of patients with chlamydia at both sites were also low due to medical staff forgetting to save the urine samples of individuals who were positive at both sites. Another issue was that some patients were only approached when they had come back for treatment as they had given a self- taken swab for testing. These individuals were only asked to participate in the study if their tests were positive for chlamydia at which point they were unlikely to want to give an extra swab and would probably prefer to collect their medication and leave the clinic immediately.

This is the first study where the MLV-*omp*A typing scheme has been used to type rectal samples from an MSM population (in Brighton). To summarise the findings in this study, there were 37 isolates successfully isolated in cell culture using McCoy cells. In addition, 76% of the samples were assigned full MLVA-*omp*A genotype and these samples were further analysed by generating a minimum spanning tree using the typing data and also determining if there were any significant differences between demographic factors and the MLVA-*omp*A types of each sample.

CHAPTER 5 EPIDEMIOLOGY OF CHLAMYDIA IN SOUTHAMPTON.

5.1 INTRODUCTION

Genital *Chlamydia trachomatis* is the most common bacterial sexually transmitted infection (STI) diagnosed in the UK. If chlamydia is diagnosed early, it can be easily treated with antibiotics and will not cause any long-term complications. However, as chlamydia infections are frequently asymptomatic (in up to 70% of women and 50% of men) chlamydia infections tend to be undiagnosed and subsequently go untreated leading to further complications such as pelvic inflammatory disease (PID) and ectopic pregnancy.

Since 1999, the number of reported annual cases of chlamydia in England has more than doubled. In 2012 an estimated 1.7 million tests for chlamydia were taken in the UK; with approximately 136,000 patients producing a positive result (PHE, 2013). There were also 189,612 new diagnoses of chlamydia in England during this year. According to the Health Protection Agency (HPA) Southampton has one of the highest rates of STIs in the UK. The city has previously been highlighted by the HPA as having one of the worst infection rates in 2010 and saw a 38% increase in *C. trachomatis* between 2000 and 2009 (HPA, 2011). Additionally, between 2009 and 2011 there was a 10% increase of new diagnoses from 1,108 new cases to 1,225 new cases respectively. The increases seen in the recorded rates of chlamydia are potentially due to the introduction of the National Chlamydia Screening Programme (NCSP) in 2003 (NHS, 2012) resulting in a significant increase in data.

As well as the increase in individuals diagnosed with *C. trachomatis,* there have been recorded mutations in the genetic makeup of *C. trachomatis*. The most notable variant was discovered in Sweden where *C. trachomatis* evaded detection in the regular Nucleic Acid Amplification Technology (NAATs) (Ripa and Nilsson, 2007). To understand better the detailed epidemiology in defined local populations a high resolution typing method for chlamydia is required to define the distribution of types of chlamydia and to identify the spread of the disease through sexual networks. A technique known as MLVA-*omp*A analysis has been evaluated for *C. trachomatis* (Pedersen *et al.*, 2008, Wang *et al.*, 2011b). This method is highly discriminatory (Pedersen *et al.*, 2008, Wang *et al.*, 2011b) and a recent study has shown that the markers used for this typing scheme are stable over several passages in cell culture (Chapter 3).

A study was conducted over a six month period in 2009 using the MLVA-*omp*A typing system to test the system on samples from a local population in Southampton (Wang *et al.*, 2011b). In the study chlamydia samples from 157 women were genotyped. DNA extracted from endocervical swabs from the women was analysed by the COBAS Taqman® CT test, V. 2.0 (Roche diagnostics, UK). DNA from positive samples were further analysed using the MLVA-*omp*A typing system. The

primary swab sample was also retained so that the chlamydia could be isolated in cell culture to obtain full genotypes for samples that could not be genotyped directly from the extracted DNA. In the Wang *et al* (2011b) study a total of 59% of the samples were successfully assigned a full MLVA-*omp*A genotype. The study only included samples from women living in Southampton and no demographic data could be recorded for the patients as ethical permission was not obtained for the collection of demographic data. The study described in this chapter includes samples collected from both women and men together with the demographic data from consenting patients so that we could attempt to define sexual networks in Southampton and to gain a more comprehensive understanding of the epidemiology of chlamydia in the city. Samples from MSM were also accepted for my study.

The purpose of my study was to collect anonymised patient samples with their accompanying demographic data and to genotype these samples using the MLVA-*omp*A typing system (Pedersen *et al.*, 2008). An important component of this investigation was to determine if it would provide suitable information to define sexual networks in Southampton. Southampton with its high infection rate and as a major commercial port city was an ideal location for this study. As well as being the entry and exit routes for many individuals who embark on cruises and also for merchant ships, Southampton also has an airport, therefore the city attracts many visitors on a daily basis.

There have been similar epidemiological studies carried out using clinical samples where demographic data were also analysed. However, these studies have used multi sequence typing (MST) (Bom *et al.*, 2011, Christerson *et al.*, 2012, Bom *et al.*, 2013b). No large scale studies have been carried out where both demographic and MLVA-*omp*A typing data have been collected and analysed.

5.2 AIMS:

5.2.1 Primary aim

To define sexual networks in which bridge populations can be identified (these are people who have sexual relations with people from different networks and therefore link the networks together) by assigning genotypes to chlamydia positive samples using the MLVA-*omp*A typing system.

5.2.2 Secondary aims

- To determine if there are genotypes associated with certain demographic profiles such as the age of the participant.
- To compare the demographic profiles of those who are *C. trachomatis* positive and those who are negative using anonymised patient information.

- To determine if there are genotypes associated with asymptomatic and symptomatic disease presentation.
- To determine if co-infection could be linked to symptomatology.
- To determine if sexual networks can be identified using information from those who are co-infected with *C. trachomatis* and other STIs.

5.3 STUDY HYPOTHESIS.

It was hypothesised that there are specific *C. trachomatis* genotypes associated with specific demographic groups and from the identification of the genotypes associated with each demographic group sexual networks within the community can be identified and analysed.

5.4 STUDY OBJECTIVES AND STUDY DESIGN.

This study was designed as a prospective study in which genotypes of chlamydia in the community were analysed and was conducted so that the spread of *C. trachomatis* in the community could be followed and could possibly be used in the future as a tool to tackle the spread of chlamydia and gain a more in depth understanding of the sexual networks in Southampton.

To carry out this study ethical permission was obtained from the University (school ethics ref - 2130- ERGO) and from the National Research Ethics Committee (NREC) (Reference number - 12/LO/1021) as patient samples and demographic information from the patients had to be obtained (Figure 5.1). The University of Southampton also provided insurance and sponsorship for the study (research governance office reference -RGO REF- 8598). Details and letters with regards to the ethics obtained can be found in the appendix section of this thesis. This study was also included in the National Institute for Health Research portfolio (NIHR) under the reference 12854. Inclusion on the portfolio ensured that the clinic could obtain funding and support from the UK Clinical Research Network (UKCRN) (Section 2.18). Once the study is complete (following publication of the results and findings) the samples will be stored in the UK Human Tissue Bank so that the samples can be accessible to other researchers.

The main technical objective of the study was to collect residual chlamydia positive samples from patients visiting the GUM clinic at the Royal South Hants (RSH) hospital, Southampton and to assign MLVA-*omp*A types to each sample based on the sequencing data. To determine the sample size that was required to make the study statistically relevant previous data on the number of positive chlamydia samples previously received per month at the HPA was obtained from the HPA, Southampton (Dave Browning, Personal communication). These figures showed that following routine tests there were on average 100 *C. trachomatis* positive samples per month. Following discussions with Mr Scott Harris, the Statistician for the Faculty of Medicine at Southampton General Hospital (SGH) the number of samples we chose to collect was based on the length of

time in which the study was conducted. As the study was part of a PhD project there was a constraint of a maximum of eight months for the collection of samples. Looking at data from the past year (May 2011- April 2012) there were a total of 1262 positive samples. Therefore, we hoped to achieve somewhere between 600 and 800 positive samples. As *omp*A is one of the genes in the MLVA-*omp*A typing scheme we chose to look at the prevalence of each *omp*A genotype in a previous study conducted in Southampton (Wang *et al*, 2011), so that we could determine the confidence intervals by which we would be able to estimate the prevalence of each *omp*A type in the population. We chose to group the samples by the *omp*A genotypes and not by the MLVA-*omp*A types because there were 47 MLVA-*omp*A types identified in the study by Wang *et al* (2011b), most of which were unique to the *omp*A types, therefore, we reasoned that there would be several new MLVA signatures in this study as we aimed to collect between three to five times the number of samples collected in the Wang *et al* (2011b) study. Therefore the best way to conduct this initial analysis was to analyse the *omp*A types. The reason for doing this was because if we had enough samples which represent the various *omp*A genotypes in Southampton, then this ensured that the data would be representative of a wide range of networks in the city.

Table 5.1 shows the precision with which we would be able to estimate the prevalence (using 95% confidence intervals) of the various chlamydia *omp*A genotypes provided that we achieve the lower and upper sample size of 600 and 800 positive samples, respectively. As can be seen when the sample size is increased from 600 to 800 this decreases the confidence interval because with more samples there is an increased chance of the genotype being a good match for the whole population. To decrease the confidence intervals further we would need to collect more samples, but as this was not possible due to time constraint we reasoned that these confidence intervals would be adequate to capture enough of the population to represent the individuals with each type in our study. Therefore, on advice from Mr Scott Harris, we aimed to collect 600 to 800 samples during the eight months that the study was going to take place.

Collection of samples and demographic information was co-ordinated by Dr. David Rowen (a consultant at the RSH hospital in Southampton). Samples were collected from individuals over the age of 16 and no individuals below this age were asked to participate as we did not have ethical permission to include them. Additionally, according to the Human Tissue Act (HTA) to include children into the study consent would need to be obtained from someone who has parental responsibility and we realised that it may be almost impossible to gain consent from parents as they very rarely attend the GUM clinic with their children. Every sample obtained was analysed in accordance with the 2004 HTA. Each patient was given an information sheet (Appendix 7) and the project was explained to them by a nurse so they understood what they were consenting to before they were given a consent form to sign and date (Appendix 8). After each patient sample

had been tested and the individual had been given their results the residual material from the positive samples (which were otherwise discarded) were collected. Before the PCR was performed to amplify the regions of interest, DNA was extracted from the samples at the Molecular Microbiology Department, SGH using the QIAamp viral RNA extraction kit (as this kit was recommended by the manufacturer – Qiagen, for the extraction of DNA from urine samples) (Section 2.11.3). MLVA-*omp*A genotypes were assigned to each sample according to sequences obtained (Table 4.1).

As well as assigning genotypes to the samples, anonymised demographic patient information was collected to conduct a survey of patient cohorts and for use in network analysis by identifying specific bridge populations. The patient samples and demographic information were anonymised using a unique study number. The demographic data collected from the patients included their age, gender, occupation, ethnicity and the postal district (first four digits of postcode) of their home address. Information on symptomatology, of their previous history of chlamydia and other STIs they were infected with at the time of examination were also obtained.

Even though a large proportion (approximately 90%) of patients whose samples were tested were negative for *C. trachomatis*, the demographic information they provided was still useful for this study. The information provided by the individuals was used to compare the demographic information provided by those who tested negative to those who tested positive to determine if *C. trachomatis* is more likely to be detected in individuals in certain demographic groups. Comparisons between MLVA-*omp*A types and the demographic information collected were conducted to determine if there were MLVA-*omp*A types associated with specific demographic groups. For instance there may be certain types which are found in certain age groups or in individuals who reside in certain areas.

Type	Prevalence	Width of Confidence interval	
		600 samples	800 samples
D	20.4%	3.2%	2.8%
D/UW-3/CX	*11.8%*	*2.6%*	*2.2%*
D/IC-CAL3	*8.6%*	*2.2%*	*1.9%*
E	40.9%	3.9%	3.4%
F	16.1%	2.9%	2.6%
G	6.5%	2.0%	1.7%
H	0%	0.3%	
Ia	4.3%	1.6%	1.4%
J	4.3%	1.6%	1.4%
K	7.5%	2.1%	1.8%

Table 5.1 Prevalence estimates and expected precision for different *C. trachomatis ompA* types.

5.5 COLLECTION AND STORAGE OF SPECIMENS.

Samples were collected at the GUM clinic by trained GUM clinicians and then sent to the Southampton HPA chlamydia diagnostics service in the Molecular Diagnostics Unit (MDU) for routine analysis. There was no risk to the participants during collection of the samples as the samples they provided were part of the routine GUM clinic consultation and had been subject to independent risk assessment by the HPA. There was no additional invasive sampling other than what was required for routine testing. Once the samples had been collected from the patient they were labelled with the patient number. Patients who were willing to participate in the study provided informed consent at their consultation which meant that their sample and the demographic information they provided could be used in this study. Once the samples had been tested and stored for the necessary period required by the HPA for repeat testing, the patient numbers of the individuals who had consented was sent to the team at SGH and only these samples were collected and transported to the Molecular Microbiology Department, SGH for further analysis. Upon receipt of the samples, the samples of those who had consented were given a new study number and were then stored at -20°C to reduce denaturation of the DNA, until it was extracted from the samples for the MLVA-*omp*A PCR. Samples belonging to patients who did not wish to participate in the study were discarded. The results of diagnostic tests were given by staff at the HPA to the GUM clinic staff who advised the patients of their results in the normal manner. The demographic information for those who had given consent for their sample to be used in the study were given to the research team in an anonymised form by Dr. David Rowen. The process by which samples were collected and analysed is shown in Figure 5.2.

5.6 SEXUAL NETWORKS.

Sexual networks are defined as groups of people who are behaviourally connected to one another sexually (Ghani *et al.*, 1996). How quickly a STI is transmitted within the network is dependent on where high risk individuals are within the network, the number of people within that particular network, and if individuals are in monogamous relationships. The identification of those with high numbers of partners can also be used as a tool to reduce the spread of STIs (Ghani *et al.*, 1996). These individuals can be educated about carrying out safe sex and also be given treatment so that they do not pass the STI to other individuals and effectively make the network larger.

To define sexual networks in a community the patients have to be able and willing to identify who their sexual partners have been. However, many are not always so willing to do this. Incomplete partner links may lead to gaps in sexual network analysis and lead to an under estimation of how big networks actually are. Molecular typing and epidemiological data are therefore very important when looking at sexual networks for STIs, such as *C. trachomatis* as this information can

give insight to transmission patterns, which can help to improve strategies which can be used to drive down the number of people infected with *C. trachomatis*.

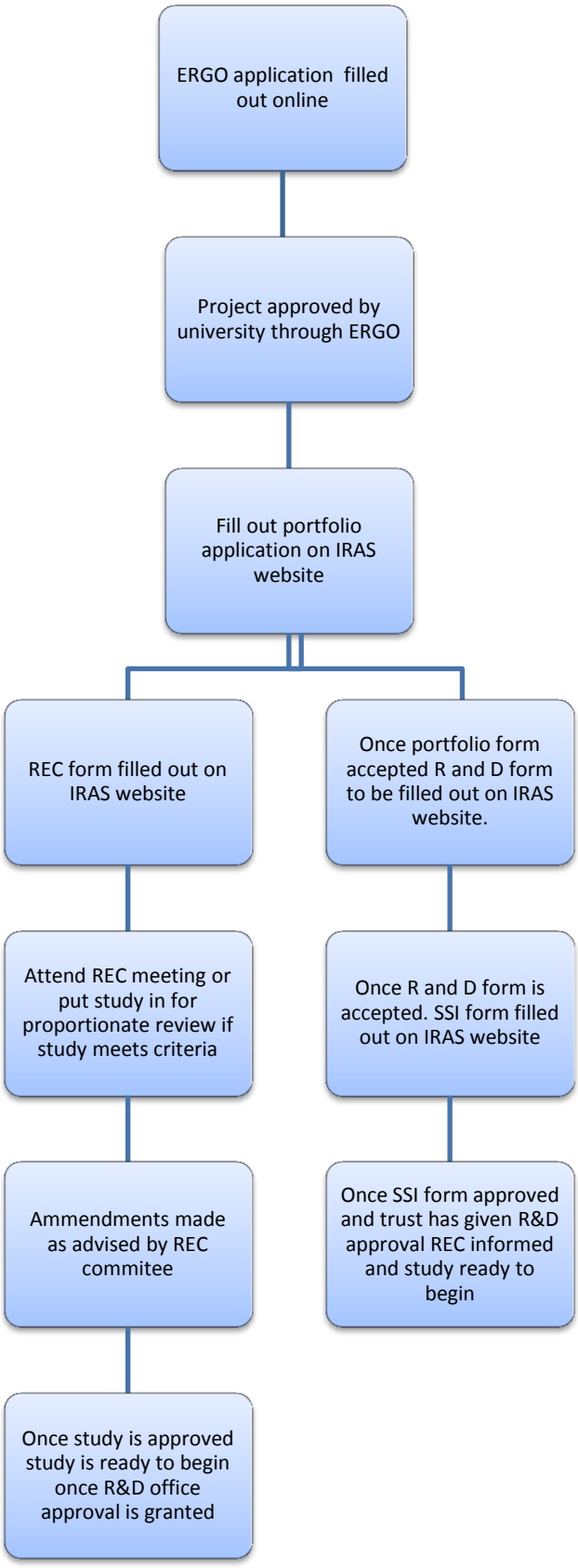


Figure 5.1 Flow diagram of process to gain ethical approval before a study requiring ethical approval can begin.

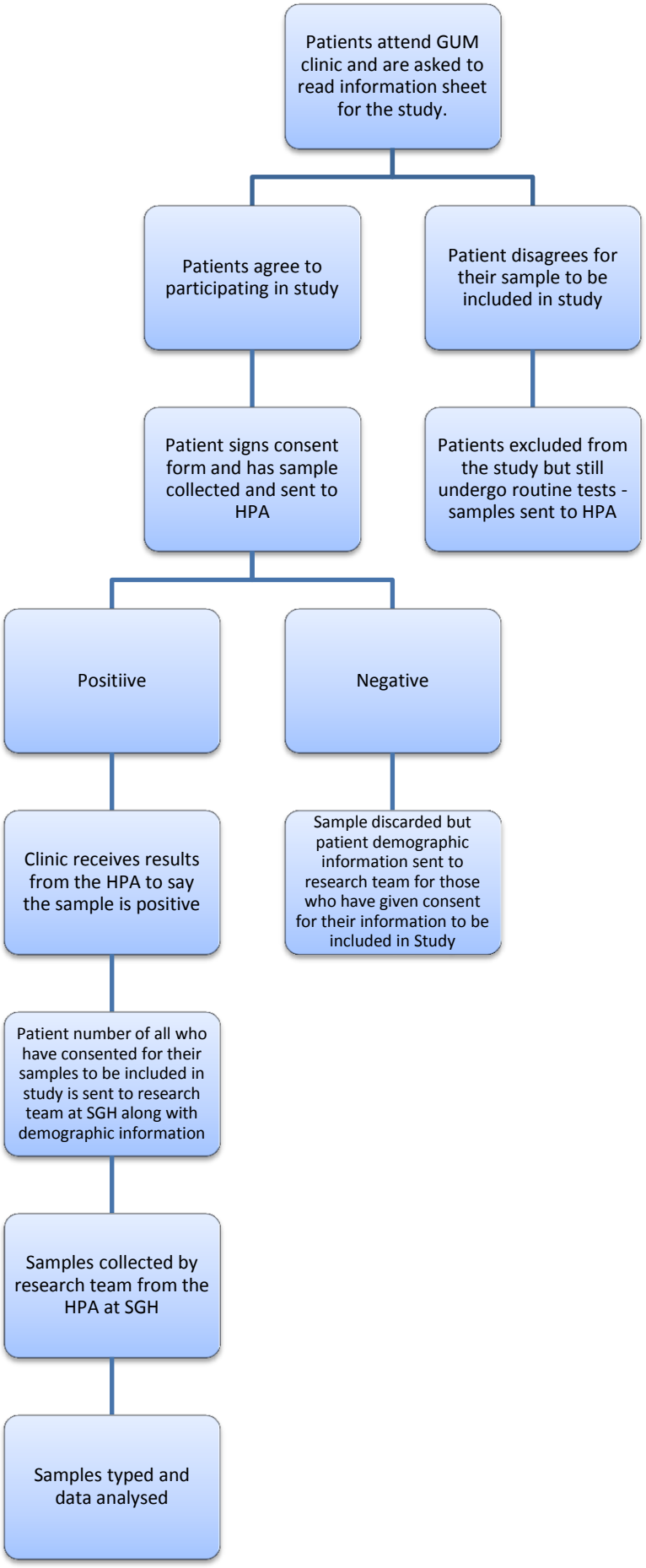


Figure 5.2 Flow diagram of the process by which samples were collected and subsequently analysed.

5.7 STUDY POPULATION AND TYPING OF SAMPLES.

During the study period, 3118 individuals consented to participate in this study (1465 males and 1653 females respectively). Samples were taken by the GUM clinic staff, most of the men in the study gave urine samples whilst the women gave swabs. The swabs collected from women were taken from either the higher vagina, the vagina or from the cervix and for men who gave swabs the swabs were taken from the urethra. In total 199 swabs, 179 urine samples and two pharyngeal swabs were provided by male participants in this study. The total number of individuals who visited the GUM clinic during the study period is not known as this information was not recorded. Additionally, not all of the samples from patients who had positive results and gave consent for their sample to be used could be included because on some occasions there was not enough transport media in the tube for DNA extraction and sometimes the samples had already been discarded before collection. Nevertheless, 12% of individuals who had given consent for their sample to be included in this study had positive *C. trachomatis* diagnoses.

In total there were 380 positive samples collected from 375 patients who had given consent over the period of eight months (September 2012 to April 2013). Five patients gave multiple samples. The types of samples collected from these patients are shown in table 5.2. Surprisingly out of the five patients who gave multiple samples, the only two samples which showed concordance and had the same genotype when typed, were the two urine samples which were collected from one of the male participants.

Number	Gender	
	Female	**Male**
1	2 urine samples	2 urine samples
2	2 urine samples	1 pharyngeal, 1 urine
3		1 urine, 1 urethral

Table 5.2 Showing the type of samples collected from two female and three male patients who gave multiple samples at one visit to the clinic.

Following extraction of the DNA from the swabs and urine samples, samples were assigned types using the MLVA-*omp*A typing system described by Pedersen *et al* (2008). Sequencing was conducted at SourceBioscience, Nottingham (Section 2.13.3).

In total samples from 191 females and 184 males were typed. Of the male participants ten of them identified themselves as being MSM whilst only one female identified herself as being bisexual. 302 of the 380 samples that tested positive (79.5%), all four loci were successfully

amplified, sequenced and assigned types and therefore as these samples were assigned full MLVA-*omp*A profiles they were included in further analysis.

5.8 *OMP*A DISTRIBUTION.

As *omp*A is part of the MLVA-*omp*A typing scheme the *omp*A sequences were compared to known sequences on the NCBI database. Amongst the typeable samples in this study there were twelve *omp*A genotypes: B3 (n= 3), D -two *omp*A genotype Ds were identified; D/IC-CAL8 (n= 34) and D/UW-3CX (n= 51), E (n= 164), F (n= 35), G (n= 28), H (n = 3), I (n=8), J (n=10), K (n=27) and L2b (n=2) (Figure 5.3).The most prevalent *omp*A type were samples which were identical to the reference strain on the NCBI website E/Bour (55%), followed by those which were identical to D/UW-3CX (14%), D/ICAL8 (9%) and F/ICAL-8 (10%).

5.8.1 *omp*A genotype B3/IU-FQ279.

Surprisingly, there were three samples which were assigned the *omp*A genotype B3. Genotype B is generally associated with patients who have trachoma and not genital tract infections. Genotype B3/IU-FQ279 is a very rare genotype found in trachoma patients (Batteiger *et al.*, 2010), therefore to investigate the difference between the genotype B3/IU-FQ279 and a more common genotype which causes trachoma B/JAli-20-OT, the two *omp*A sequences were aligned against one another using Seqman (Section 2.15.2.1) and were found to possess 99% similarity to each other. When the *omp*A sequence of B3/IU-FQ279 was compared to *omp*A sequences of genital tract strains with *omp*A genotypes D-K there was between 83% to 94% similarity. This showed that the *omp*A genotype B3/IU-FQ279 has a higher percentage similarity to the trachoma strain than it does with the genital tract strains. As the *omp*A genotype B is usually found in countries where trachoma is still a major problem such as Africa and Asia the ethnicities of the females whose samples were genotyped as having an *omp*A genotype B were checked. For one of the females there was no information recorded relating to her ethnicity, whilst one lady was of white British descent and the third was of Indian descent.

5.8.2 *omp*A genotype Ia

In this study, eight samples were typed as *omp*A genotype Ia. Genotype Ia has previously been associated with black individuals (Geisler *et al.*, 2006). Four samples were collected from female participants whilst the other four were collected from male participants. All of the eight participants with this genotype were between the ages of 18-22 years old and were all white British individuals.

5.8.3 *omp*A genotype L2b

It was also surprising to find two samples which were typed as *omp*A genotype L2b. L2b is usually found in rectal samples of HIV+ MSM (Christerson *et al.*, 2012, Bom *et al.*, 2013b). However, two of the male participants in this study who had given urine samples to be tested for chlamydia had

samples which when typed had an *omp*A type L2b. When the demographic data recorded for the individuals with L2b were checked, the males were of white British descent and there was no record of their HIV status. The sexual orientation of both males was recorded heterosexual and both males presented with symptoms on the day.

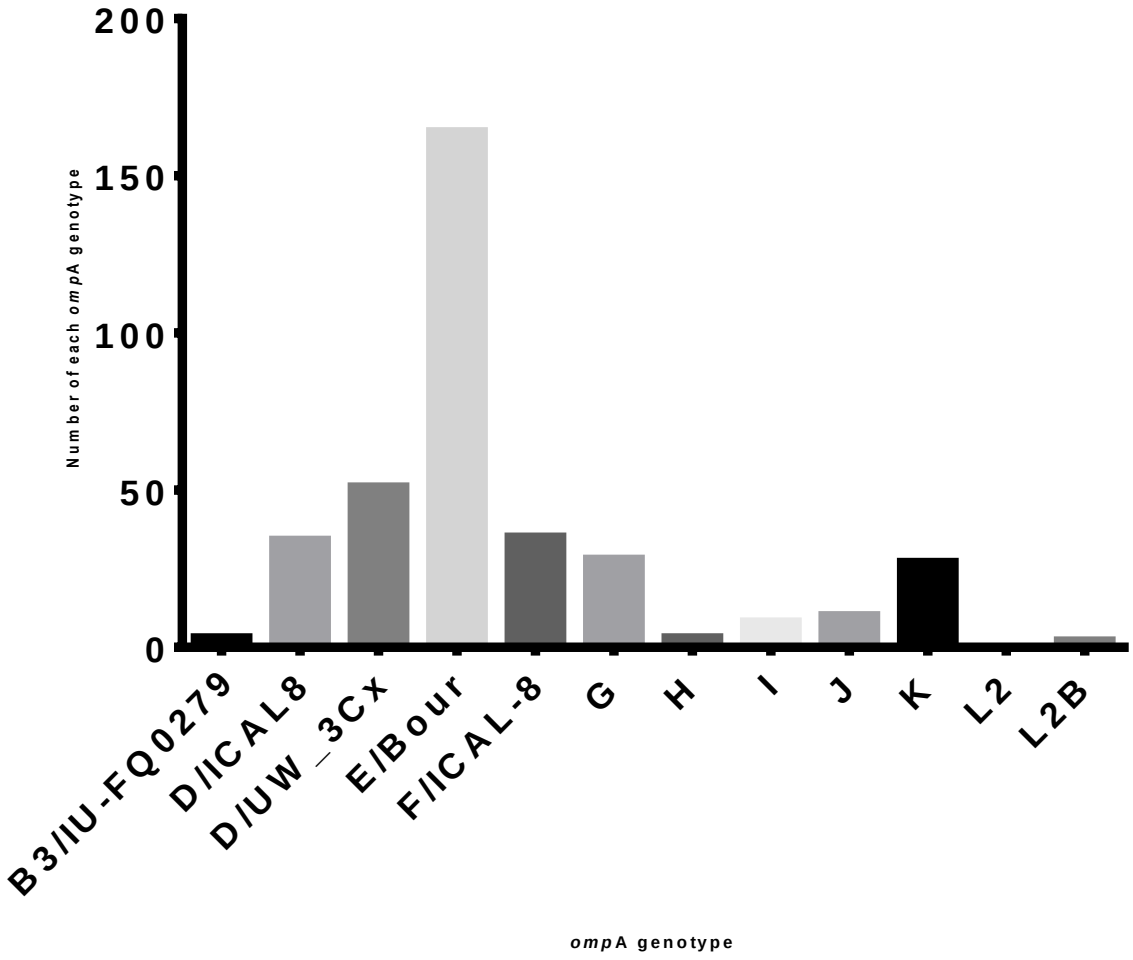


Figure 5.3 Distribution of *omp*A genotypes.

5.9 GENOTYPING USING THE MLVA-*OMP*A TYPING SCHEME.

A detailed table showing all the types and variants of the variable number tandem repeats can be found in chapter 4 – Table 4.1. A new type was assigned to each new sequence of repeated mononucleotides for each VNTR locus. Continuous numbers, starting with 1, were selected, resulting in a three digit MLVA type for each VNTR locus, CT1335, CT1299 and CT1291, e.g. MLVA type 8.5.1 refers to a VNTR type 8 for VNTR locus CT1335, VNTR type 5 for VNTR locus CT1299 and VNTR type 1 for CT1291. There were some new MLVA *omp*A types found in this study. For each of the new repeat sequences that were found following sequencing, new variant numbers were assigned to the repeat sequence. In total, four new VNTR variants were found for CT1299 and for the VNTR type CT1335, two new variants were found (Table 5.3).

VNTR type	**Variant number of VNTR type and corresponding repeat sequence**							
CT1335	14	14T7A	15	11T9A				
CT1299	10	15C	11	16C	12	18C	13	21C

Table 5.3 New types and variants of the variable number tandem repeats (VNTRs) found in this study.

Amongst the 302 fully genotyped samples, 151 unique MLVA-*omp*A types were found. There were some MLVA types which were common in a range of different *omp*A types. However, the majority of MLVA types were unique to each *omp*A genotype (Table 5.4). MLVA type 3.3.4 was the most common MLVA type found in different *omp*A genotypes, it was typed in seven *omp*A genotypes including samples with identical *omp*A genotypes to:- D/UW-3Cx, D/IC-CAL8, E/Bour, G/392, J/UW-36, K/UW-31. In total there were 91 different MLVA types across the seven *omp*A genotypes, 56 of these MLVA types were only found in one *omp*A genotype – *omp*A genotype E. Even though MLVA type 3.3.4 was linked to the most *omp*A types, it was actually MLVA type 8.5.1 (n=25) which had the most samples associated with it, followed by MLVA type 3.3.4 (n=16) and 8.5.5 (n=15). The *omp*A type with the most MLVA variants was genotype E followed by D, F and G (Genotypes F and G both had eighteen MLVA variants within the *omp*A type).

	Omp A genotype												
MLVA type	B3/IU-	D/UW-3	D/IC-CAL8	E/Bour	F/IC-CAL3	G/392	H	Ia/870	J/UW-36	K/Uw-31	L2b	No. of isolates with this MLVA type	No.of ompA genotypes with this MLVA type
											1.3.1 (1)	1	1
	1.4.3 (1)											1	1
								1.5.7(1)				1	1
		1.7.3 (1)										1	1
				1.8.1 (2)								2	1
							3.1.3 (1)		3.1.3 (1)			2	2
			3.2.1 (1)	3.2.1 (8)	3.2.1 (1)	3.2.1 (1)						10	4
		3.2.3 (1)								3.2.3 (4)		5	2
		3.2.4 (1)		3.2.4 (1)				3.2.4 (1)		3.2.4 (1)		4	4
										3.3.1 (1)		1	1
										3.3.2 (1)		1	1
										3.3.3 (3)		3	1
		3.3.4 (5)	3.3.4 (2)	3.3.4 (3)	3.3.4 (1)	3.3.4 (3)			3.3.4 (1)	3.3.4 (1)		16	7
										3.3.7 (1)		1	1
						3.4.2 (1)						1	1
		3.4.3 (1)							3.4.3 (2)	3.4.3 (1)		4	3
		3.4.4 (3)				3.4.4 (2)						5	2
				3.4.5 (1)								1	1
				3.5.1 (2)				3.5.1 (1)				3	2
			3.5.2 (2)	3.5.2 (9)						3.5.2 (1)		12	3
	3.5.3 (1)			3.5.3 (1)		3.5.3 (1)			3.5.3 (2)			5	4
		3.5.4 (1)		3.5.4 (1)		3.5.4 (1)				3.5.4 (1)		4	4
								3.5.5 (2)				2	1
				3.6.1 (4)		3.6.1 (1)						5	2
						3.6.2 (1)						1	1
		3.6.3 (3)	3.6.3 (1)			3.6.3 (2)						6	3
		3.6.4 (3)				3.6.4 (1)						4	2
		3.6.7 (2)										2	1
		3.7.3 (1)								3.7.3 (2)		3	2
		3.7.4 (1)		3.7.4 (1)								2	2
									3.8.2 (1)			1	1
		3.8.3 (1)										1	1
								3.8.4 (1)				1	1
						3.10.3 (1)				3.10.3 (3)		4	2
		3.11.3 (1)										1	1
				4.2.1 (1)					4.2.1 (1)			2	2
				4.5.1 (1)								1	1
					4.5.2 (1)							1	1
					4.6.2 (1)							1	1
						4.6.4 (1)						1	1
					4.7.2 (1)							1	1
				5.4.1 (1)								1	1
				6.4.1 (1)								1	1
				6.5.1 (9)								9	1
				6.5.2 (1)								1	1
				6.6.1 (1)								1	1
				6.6.6 (1)								1	1
		6.7.1 (1)										1	1
				6.8.1 (1)								1	1
			6.8.4 (1)									1	1
					6.10.1 (1)							1	1
										7.3.4 (1)		1	1
			7.6.1 (1)									1	1
				8.1.1 (1)								1	1
							8.1.3 (2)					2	1
		8.2.1 (1)		8.2.1 (2)								3	2
					8.2.2 (1)				8.2.2 (1)			2	2
				8.2.4 (1)								1	1
		8.3.1 (1)	8.3.1 (1)	8.3.1 (1)	8.3.1 (2)							5	4
		8.3.4 (4)										4	1
				8.4.1 (6)						8.4.1 (1)		7	2
				8.4.2 (1)								1	1
				8.4.3 (2)								2	1
			8.4.4 (1)	8.4.4 (2)								3	2
			8.5.1 (2)	8.5.1 (21)	8.5.1 (1)			8.5.1 (1)				25	4
			8.5.2 (5)	8.5.2 (2)	8.5.2 (7)					8.5.2 (1)		15	4
				8.5.3 (3)						8.5.3 (1)		4	2
			8.5.4 (2)	8.5.4 (2)		8.5.4 (1)						5	3
				8.5.5 (2)								2	1
			8.6.1 (2)	8.6.1 (19)	8.6.1 (1)							22	3
					8.6.2 (5)							5	1
				8.6.3 (2)	8.6.3 (2)							4	2
			8.6.4 (1)	8.6.4 (4)	8.6.4 (2)							7	3
	8.6.5 (1)											2	1
			8.7.1 (3)	8.7.1 (8)								11	2
			8.7.2 (1)		8.7.2 (1)							2	2
			8.8.1 (1)	8.8.1 (3)								4	2
				8.8.2 (1)	8.8.2 (1)							2	2
				8.8.3 (1)								1	1
			8.8.4 (1)									1	1
					8.11.1 (1)							1	1
									8.12.2 (1)			1	1
		11.3.4 (1)										1	1
				11.4.4 (1)		11.4.4 (4)						5	2
						11.6.3 (1)						1	1
						11.8.3 (1)						1	1
						11.8.4 (1)						1	1
						12.3.3 (1)						1	1
								13.6.5 (1)				1	1
				13.7.1 (1)								1	1
					15.7.3 (1)							1	1
No. of samples with this MLVA-ompA type	3	33	28	136	31	25	3	8	10	24	1	302	
No. of MLVA types per ompA genotype	3	19	17	42	18	18	2	7	8	16	1		

Table 5.4 Distribution of different VNTR types according to *omp*A genotype (total of each type).

5.10 SIMPSONS INDEX OF DIVERSITY FOR THE MLVA-*OMP*A TYPING SYSTEM WITH THIS SAMPLE SET.

The Simpsons index of diversity (Hunter and Gaston, 1988) was calculated for the MLVA-*omp*A typing system to determine the discriminatory power of the MLVA-*omp*A typing system for these samples. The discriminatory power of the following combinations of markers were calculated to determine if there were any differences in discriminatory power when different combinations of the markers were used to type samples. Firstly the discriminatory power of the four loci (CT1335, CT1299, CT1291, *omp*A) individually were determined using the equation in chapter two - Equation 2-1. This was calculated to identify if a single MLVA marker could give a higher discriminatory power than using all of the markers together, if this were the case then this will reduce the amount of PCRs that need to be conducted and reduce the need to determine the genotype of several PCR products. The diversity values were as follows: CT1335 = 0.63, CT1299=0.82, CT1291= 0.73 and *omp*A= 0.75.The discriminatory power of all the MLVA loci- CT1335, CT1299 and CT1291 were then determined and this gave a discriminatory value of 0.97. Finally, the discriminatory power of the entire MLVA-o*mp*A typing system was calculated and this had a discriminatory power of 0.98.

5.11 TYPEABILITY OF THE FOUR MLVA-*OMP*A MARKERS.

The typeability of each of the markers in the MLVA-*omp*A typing system, which is defined as a typing system's ability to assign a type to all samples was also calculated. The typeability was calculated by expressing (as a percentage) the number of typeable samples over the total number of samples. The typeability for the three VNTRs and *omp*A were as follows: CT1335= 93%, CT1299=91%, CT1291 = 86% and *omp*A= 96%. This data show that *omp*A offers the highest typeability and 10% more samples were assigned an *omp*A genotype compared to the VNTR CT1291.

5.12 MINIMUM SPANNING TREE.

5.12.1 Generation of the minimum spanning tree.

A Minimum spanning tree (MSpT) was generated using Bionumerics 6.5 (Section 2.17). This program generated groups of closely related strains based on the MLVA-*omp*A data. Amongst the 302 fully genotyped samples, 151 unique MLVA-*omp*A sequence types (STs) were found. Only the sequence data from the samples with a full MLVA-*omp*A type were used to generate the MSpT (Figure 5.4). By allowing single locus variants (SLVs), 291 samples were clustered based on the four loci in the typing scheme and formed one large network representative of the population of people that attend the GUM clinic in Southampton. The remaining eleven samples had more than

a single locus difference to the samples in the network and are shown as singletons which are represented as circles which do not connect to any of the other circles on the outside of the network. The primary founder of the group was defined as the MLVA-*omp*A type which differs from the most STs at only a single locus, in other words the sequence type with the largest number of SLVs. The primary founder of the complex (the largest circle in the complex) was MLVA-*omp*A type 8.5.1 –E which had 21 specimens with this type.

5.12.2 Location of *omp*A genotypes on the Minimum spanning tree.

The MST was also colour coded to show the location of *omp*A types amongst the various MLVA-types (Figure 5.5). This gave a visual representation of *omp*A genotype clustering relative to MLVA-*omp*A genotypes. From the diagram it can be seen that the MLVA types for genotype E cluster together and that there are more samples with an *omp*A genotype E, likewise loci that were identical to *omp*A genotypes D/UW-3CX, G and K clustered together in one region of the MSpT. On the contrary, *omp*A genotypes identical to D/IC-CAL8 and F/IC-CAL3 were dispersed throughout the network. There were not enough samples with the *omp*A genotypes B3, I, J, L2 and L2b to identify any clonal complex clusters formed from these samples.

5.12.3 Partioning of the minimum spanning tree to create groups for which demographic data was analysed.

As all the samples obtained from the study formed one large network consisting of 151 different MLVA-*omp*A types, due to the large number of STs obtained from genotyping the samples, the network was partitioned using different rules. In total three rules were defined and this gave rise to three different sets of analysis- the two segment (mentioned above), three segment and the ten segment analysis. These segments were named as such due to the number of segments the MSpT was split into once the rules were applied.

To compare the demographic variables between types, they were initially divided into two subgroups or segments (two segment analysis) – A and B (Figure 5.6B). Segment A samples contained mainly *omp*A genotypes which are least frequently detected in heterosexual populations - D, G, I and K – (There was also a small number of samples with genotypes E, J and I in this section). On the other hand, section B consisted mainly of samples with *omp*A genotypes D, E, and F which are the most prevalent genotypes in the heterosexual population.

Two other partitioning of the MSpT were conducted and these partitioning were based on the fact that there were other possible subgroup founders and that these can be identified from within larger eburst groups such as the one produced in figure 5.4, where there are several STs besides the predicted primary founder that have a number of SLVs of their own (Wang and Grayston, 1963). When a ST was linked to both primary and subgroup founders, the ST was assigned to the primary group founder. When a ST was an SLV of two or more subgroup founders

the ST was assigned on the basis of the distance from the primary founder i.e. the ST was assigned to the subgroup closest to the primary group founder. Although it has been suggested that if this scenario where to arise it may also be acceptable to assign the ST to the largest subgroup (Wang and Grayston, 1963). Here the primary founder was the MLVA-*omp*A type 8.5.1-E coloured in blue in figure 5.6c, which is also the MLVA-*omp*A type with the most SLVs and which happened to be the MLVA-*omp*A type with the most isolates. A possible sub group founder, in this instance, was defined as any MLVA-*omp*A type which has evolved ≥ six times in at least one of the SLVs associated with the subgroup founder. Subgroup founders are coloured in yellow in figure 5.6c. Applying this rule to the MST split the network into three segments (three segment analysis). The previously defined section A (figure 5.6B) was now split into two groups, numbered one and two, whilst segment B remained as one segment.

Finally, a third rule was defined to divide the network into smaller groups. This was considered necessary as differences may be identified between smaller groups of MLVA-*omp*A genotypes. In this instance an ST was considered to be a subgroup founder if it had at least six SLVs. This rule partitioned the network into ten sections which were all analysed statistically (ten segment analysis) (Figure 5.6d).

When the network was split into two segments there were clear divisions between the *omp*A genotypes in relation to MLVA-*omp*A genotype clusters. Segment A mainly contained genotypes similar to D/UW-3CX, G, H, J and K whilst segment B contained mainly *omp*A genotypes D/IC-CAL3, E and F. For the three segment analysis, section one consisted of all of the samples which were typed as having *omp*A genotype H. Segment two comprised of genotypes D/UW-3CX, G and K, whilst segment three consisted mainly of D/IC-CAL3, E and F. *omp*A genotype J was found in all three segments, with three samples which had *omp*A genotypes J in these sections. For the analysis where the network was split into ten segments most of the segments consisted of a high prevalence of one *omp*A genotype (Table 5.6). The difference in *omp*A genotypes between the segments were all significant with a p value of <0.001. This shows that there are specific *omp*A genotypes which are more prevalent in groups of similar MLVA types.

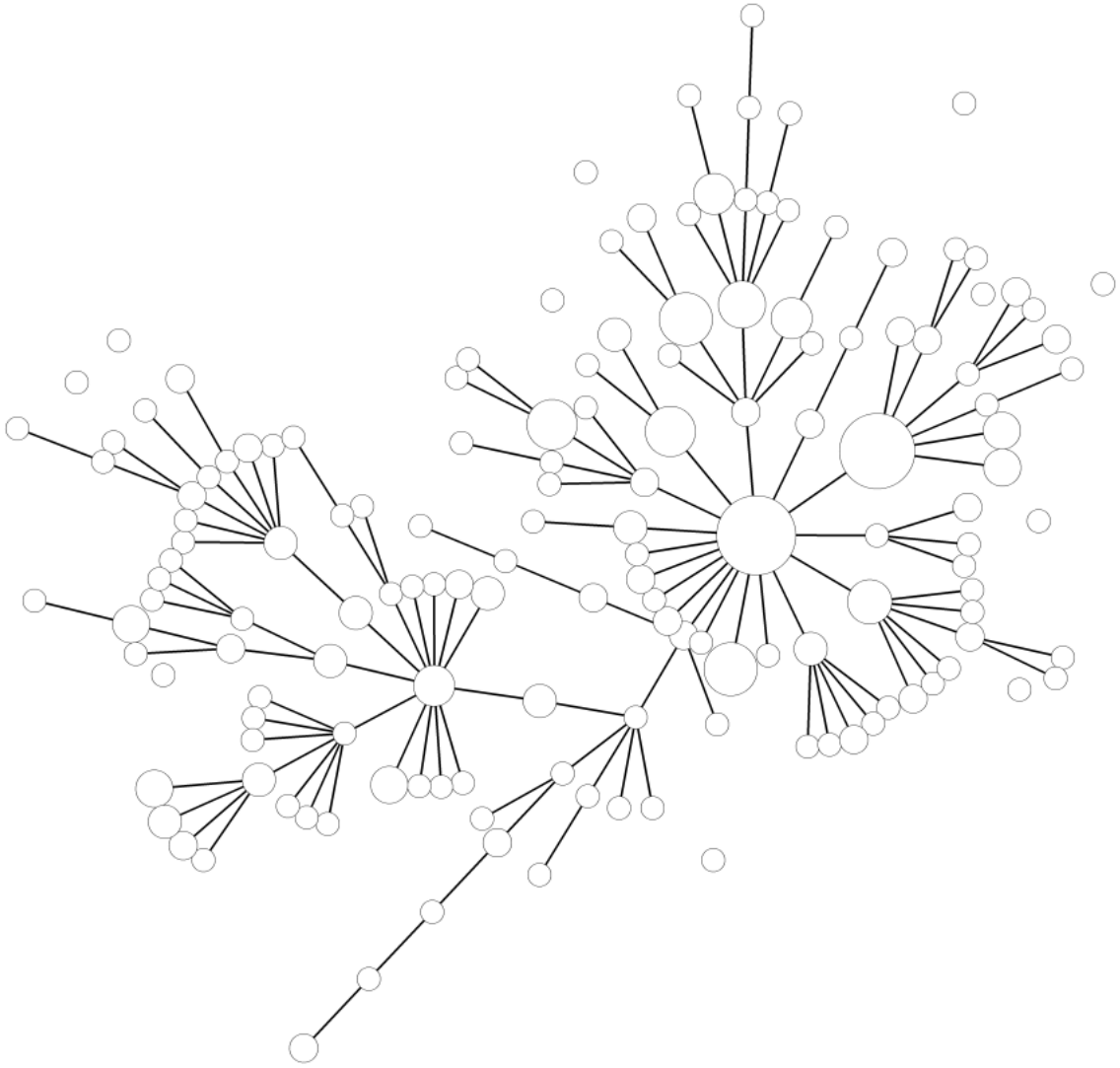


Figure 5.4 Minimum spanning tree generated using MLVA-*omp*A typing data of samples collected in Southampton.
The size of the spheres is indicative of how many samples make up each sphere. All samples within a sphere have the same MLVA-*omp*A type. The branches show the relationship between single locus variants.

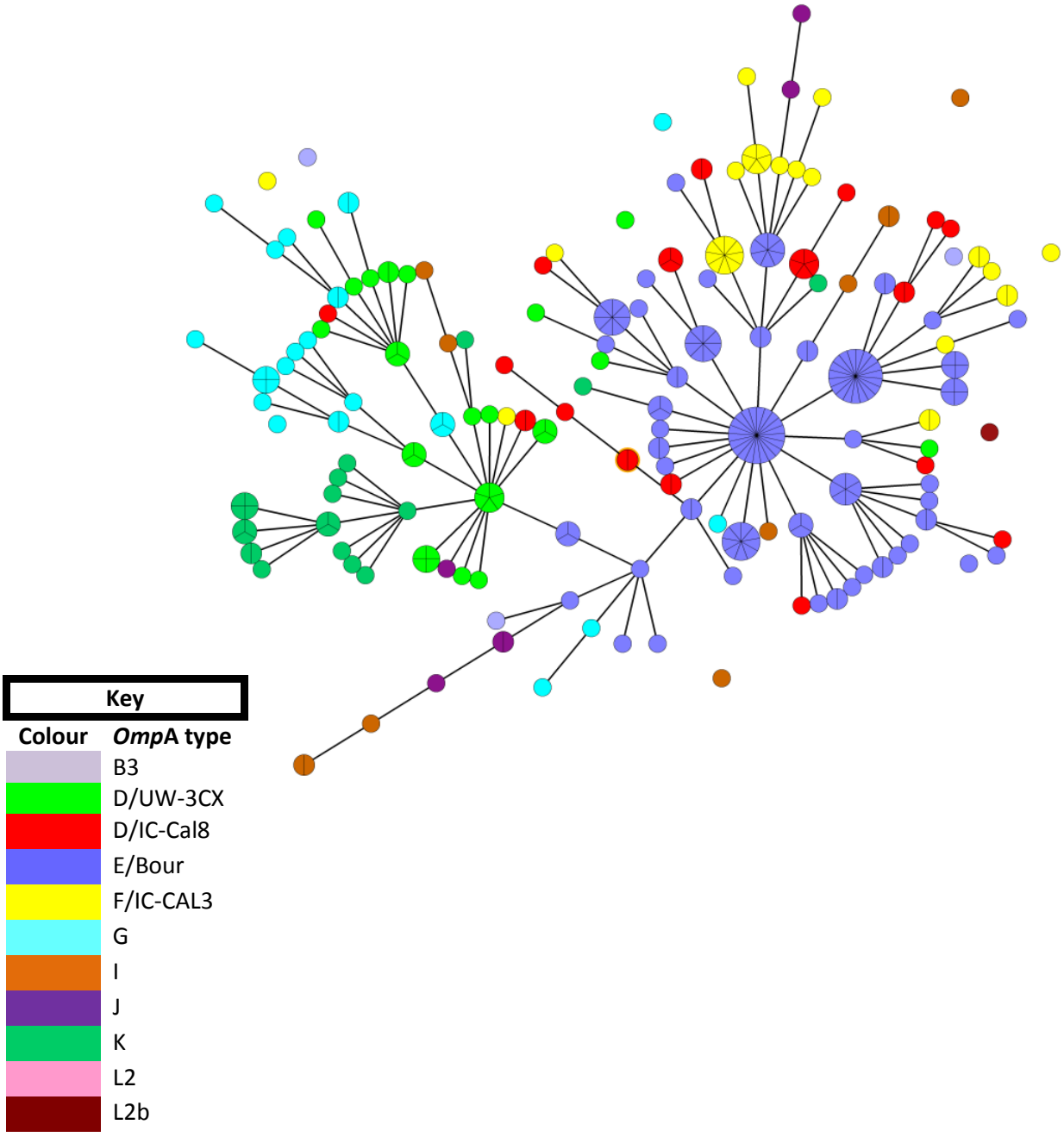


Figure 5.5 Minimum spanning tree showing the location of the *omp*A genotypes within the clonal complex.
Some spheres are divided into sections which are indicative of how many samples are in each sphere, i.e. one section within a sphere is equal to one sample.

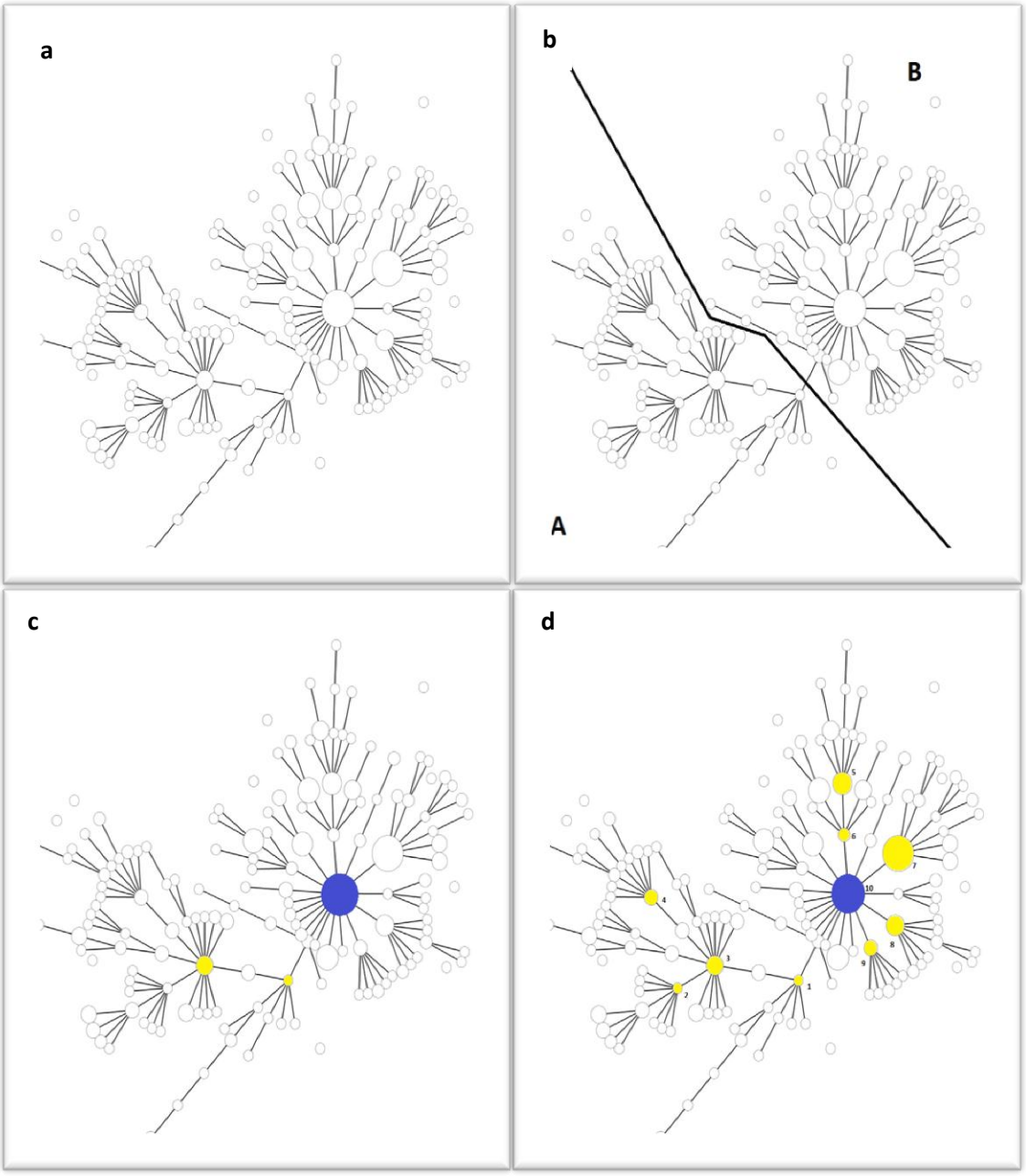


Figure 5.6 Minimum spanning tree of 302 samples from the Royal South Hants GUM clinic, Southampton (a).
The MSpT was then partitioned into two segments based on the most frequent and least frequent *omp*A genotypes found in epidemiological studies of *C. trachomatis* (b). A second and third partitioning were identified based on possible subgroup founders. (c) shows the identification of a primary group founder (blue) and two subgroup founders (yellow), whilst (d) shows the identification of the same primary group founder (blue) but now with nine possible subgroup (yellow) founders based on two different rules used to identify subgroup founders.

		Two segment analysis		Three segment analysis			Ten segment analysis											P values		
		Segment A	Segment B	Segment 1	Segment 2	Segment 3	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8	Segment 9	Segment 10	Residual group (n=11)	Two segment	Three Segment	Ten segment
MLVA-*omp* A type of group founder				3.5.4-E	3.3.4-E	8.5.1-E	3.5.4 -E	3.3.4 -D	3.3.4-D	3.6.3-D	3.5.2-E	8.5.2-E	8.6.1-E	8.4.1-E	8.8.1-E	8.5.1-E				
Corresponding figure		Figure 5.6b		Figure 5.6c			Figure 5.6d													
ompA gene	B	1 (1.1)	0	1 (6.2)	0	0	1 (6.2)	0	0	0	0	0	0	0	0	0	2			
	D1	30 (31.6)	2 (1.0)	0	30 (38.0)	2 (1.0)	0	0	18 (46.2)	9 (52.9)	0	0	0	0	0	2 (2.3)	1			
	D2	3 (3.2)	25 (12.8)	0	3 (3.8)	25 (12.8)	0	0	2 (5.1)	1 (5.9)	0	8 (40.0)	4 (9.8)	1 (6.2)	1 (11.1)	11 (12.6)	0			
	E	7 (7.4)	128 (65.6)	7 (43.8)	0	128 (65.6)	7 (43.8)	0	0	0	9 (40.9)	3 (15.0)	31 (75.6)	14 (87.5)	8 (88.9)	63 (72.4)	1			
	F	1 (1.1)	28 (14.4)	0	1 (1.3)	28 (14.4)	0	0	1 (4.8)	0	11 (50.0)	7 (35.0)	6 (14.6)	0	0	4 (4.6)	1			
	G	22 (23.2)	1 (0.5)	2 (12.5)	20 (25.3)	1 (0.5)	2 (12.5)	1 (4.8)	14 (35.9)	5 (29.4)	0	0	0	0	0	1 (1.1)	1			
	H	3 (3.2)	0	3 (18.8)	0	0	3 (18.8)	0	0	0	0	0	0	0	0	0	0			
	I	2 (2.1)	4 (2.1)	0	2 (2.5)	4 (2.1)	0	0	2 (5.1)	0	0	0	0	0	0	4 (4.6)	2			
	J	6 (6.3)	3 (1.5)	3 (18.8)	3 (3.8)	3 (1.5)	3 (18.8)	0	2 (5.1)	2 (11.8)	2 (9.1)	0	0	0	0	1 (1.1)	0			
	K	20 (21.1)	4 (2.1)	0	20 (25.3)	4 (2.1)	0	20 (95.2)	0	0	0	2 (10.0)	0	1 (6.2)	0	1 (1.1)	0			
	L2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
	L2b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	<0.001	<0.001	<0.001
Gender	Female	48 (50.5)	105 (53.8)	6 (37.5)	42 (53.2)	105 (53.8)	6 (37.5)	11 (52.4)	20 (48.8)	11 (64.7)	16 (72.7)	14 (70.0)	17 (41.5)	10 (62.5)	4 (44.4)	44 (50.6)				
	Male	47 (49.5)	90 (46.2)	10 (62.5)	37 (46.8)	90 (46.2)	10 (62.5)	10 (47.6)	21 (51.2)	6 (35.3)	6 (27.3)	6 (30.0)	24 (58.5)	6 (37.5)	5 (55.6)	43 (49.4)		0.595	0.451	0.238
Age	16-20	37 (40.7)	68 (36.4)	6 (37.5)	31 (41.3)	68 (36.4)	6 (37.5)	7 (33.3)	17 (43.6)	7 (46.7)	9 (40.9)	8 (42.1)	14 (35.9)	7 (46.7)	3 (33.3)	27 (32.5)				
	21-24	32 (35.2)	67 (35.8)	4 (25.0)	28 (37.3)	67 (35.8)	4 (25.0)	9 (42.9)	14 (35.9)	5 (33.3)	5 (22.7)	10 (52.6)	12 (30.8)	3 (20.0)	3 (33.3)	34 (41.0)				
	25-30	12 (13.2)	32 (17.1)	3 (18.8)	9 (12.0)	32 (17.1)	3 (18.8)	4 (19.0)	5 (12.8)	0	6 (27.3)	1 (5.3)	9 (23.1)	3 (20.0)	2 (22.2)	11 (13.3)				
	31-40	5 (5.5)	12 (6.4)	2 (12.5)	3 (4.0)	12 (6.4)	2 (12.5)	1 (4.8)	1 (2.6)	1 (6.7)	1 (4.5)	0	3 (7.7)	2 (13.3)	0	6 (7.2)				
	41-79	5 (5.5)	8 (4.3)	1 (6.2)	4 (5.3)	8 (4.3)	1 (6.2)	0	2 (5.1)	2 (13.3)	1 (4.5)	0	1 (2.6)	0	1 (11.1)	5 (6.0)		0.884	0.782	0.809
Age in years	Median (Range)	21 (17-79)	22 (16-60)	22 (19 - 60)	21 (17-79)	22 (16 - 60)	22 (19-60)	21 (17-37)	21 (17-53)	21 (17-79)	21 (19 - 41)	21 (18-27)	22 (17- 47)	21 (17-38)	21 (18-41)	22 (16-60)				
Residence	Southampton	74 (81.3)	142 (75.9)	15 (93.8)	59 (78.7)	142 (75.9)	15 (93.8)	18 (85.7)	29 (74.4)	12 (80.0)	17 (77.3)	18 (90.0)	29 (76.3)	15 (100.0)	7 (77.8)	56 (67.5)				
	SO postcode but not in Southampton area	12 (13.2)	32 (17.10)	1 (6.2)	11 (14.7)	32 (17.1)	1 (6.2)	1 (48.8)	7 (17.9)	3 (20.0)	4 (18.2)	1 (5.0)	6 (15.8)	0	2 (22.2)	19 (22.9)				
	Outside Southampton	5 (5.5)	13 (7.0)	0	5 (6.7)	13 (7.0)	0	2 (9.5)	3 (7.7)	0	1 (4.5)	1 (0)	3 (7.9)	0	0 (0)	8 (9.6)		0.599	0.767	0.501
Asymptomatic	Yes	43 (45.3)	102 (52.3)	6 (37.5)	37 (46.8)	102 (52.3)	6 (37.5)	15 (71.4)	16 (39.0)	6 (35.3)	11 (50.0)	10 (50.0)	20 (48.8)	10 (62.5)	3 (33.3)	48 (55.2)				
	No	52 (54.7)	93 (47.7)	10 (62.5)	42 (53.2)	93 (47.7)	10 (62.5)	6 (28.6)	25 (61.0)	11 (64.7)	11 (50.0)	10 (50.0)	21 (51.2)	6 (37.5)	6 (66.7)	39 (44.8)		0.26	0.421	0.273
Ethnicity	White British	70 (76.9)	149 (79.7)	10 (62.5)	60 (80.0)	219 (78.8)	10 (62.5)	17 (81.0)	30 (76.9)	13 (86.7)	16 (72.7)	12 (63.2)	31 (81.6)	10 (66.7)	8 (88.9)	72 (85.7)				
	White other	6 (6.6)	9 (4.8)	0	6 (8.0)	15 (5.4)	0 (0)	2 (9.5)	3 (7.7)	1 (6.7)	1 (4.5)	2 (10.5)	3 (7.9)	0	0	3 (3.6)				
	Black	4 (4.4)	7 (3.7)	0	4 (5.3)	11 (4.0)	1 (6.2)	0	2 (5.1)	0	1 (4.5)	1 (5.3)	1 (2.6)	1 (6.7)	0	2 (2.4)				
	Asian	3 (3.3)	6 (3.2)	1 (6.2)	2 (2.7)	9 (3.2)	5 (31.2)	0	2 (5.1)	1 (6.7)	3 (13.6)	4 (21.1)	0	1 (6.7)	1 (11.1)	7 (8.3)				
	Any other	8 (8.8)	16 (8.6)	5 (31.2)	3 (4.0)	24 (8.6)	0	2 (9.5)	2 (5.1)	0	1 (4.5)	0	3 (7.9)	3 (20.0)	0	0		0.954	0.096	0.032

Table 5.5 Participant *omp*A genotypes and demographics, stratified by three different segment analyses.

Shown are the numbers of each type in each segment and the P values for each segment analysis for *omp*A and the various demographics

P Values ≤0.05 are significant. Significant P values are highlighted in grey. Also shown are the MLVA-*omp*A types of the subgroup founders and the corresponding figures for each segment analysis.

5.13 ANALYSIS OF PATIENT DEMOGRAPHIC INFORMATION.

5.13.1 Methods used to analyse demographic information.

The demographic data were analysed using the Pearson's Chi squared test or the Fishers exact test where applicable, to establish the significance of the demographic data in each segment analysis (Section 5.12.3). A P value of less than or equal to 0.05 was considered significant.

5.13.2 Age.

The age range of participants who had positive chlamydia samples was 16-79 for female participants and 17-60 for male participants. This age distribution was similar to the age distribution of the total specimens received, which was 16-71 for females and 16-86 for male participants (only 11 patients were older than 60 years of age). In each segment analysis (section 5.12.3) the median age was between 21 and 22 years of age (Table 5.5). The distribution of age ranges for female and male participants with positive chlamydia samples is shown in figure 5.7. The majority of specimens were collected from patients who were between the ages of 16-24 years old (n=270/375 = 72% of positive participants). The distribution of the age ranges shown in figure 5.7 follows an exponential distribution which is positively skewed with the positivity rate reducing as age increases. When the chi squared test was carried out to determine if there was any correlation between the age of the participants and the MLVA-*omp*A genotype of the chlamydia that were found in that age group, there was no significance as the P values were 0.884 (Two segment analysis), 0.782 (Three Segment analysis) and 0.809 (Ten segment analysis) (Table 5.5).

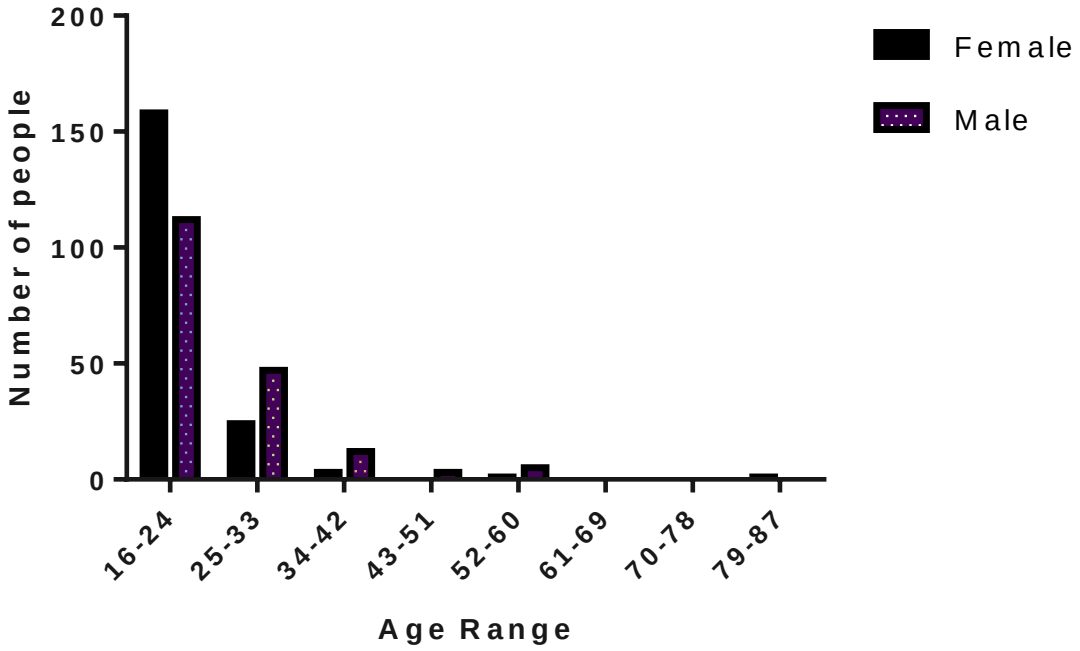


Figure 5.7 Age range of participants.

5.13.3 Gender.

In total there were more female (n=191) participants than male (n=184) participants with positive samples. Statistical analysis conducted using the chi squared test showed that there was no correlation between gender (being male or female) and the MLVA-*omp*A genotypes as in each segment analysis carried out there was an almost even distribution of both male and females in each segment. P values were 0.595 (Two segment analysis), 0.451 (Three Segment analysis) and 0.238 (Ten segment analysis) (Table 5.5).

5.13.4 Ethnicity.

The ethnic group with the highest percentage of individuals with chlamydia in this study was the white British (78%) and white other ethnic groups (8%) for both genders. This was followed by the black African and black/white Caribbean group. The lowest percentage of individuals with chlamydia was seen in the black/white African and the white Asian ethnicities (2%). To determine if the distribution was representative of the ethnicities in Southampton, the distribution of ethnicities determined in this study was compared to the office of national statistics data collected in 2010 (figure 5.8a) which showed the distribution of ethnicities in Southampton. The data showed that apart from the white population which formed the largest group, the Asians/ British Asians followed by the white-other, the Chinese and other ethnic groups formed a large proportion of those who live in Southampton. However, the Asians and British Asians were under represented in this study given that they are the second most populous group residing in Southampton (Figure 5.8a). Following chi squared analysis there was no significant difference between the ethnicity of the participant and the MLVA-*omp*A genotype ((The P values were 0.954 (Two segment analysis), 0.096 (Three Segment analysis) and 0.032 (Ten segment analysis)).

5.13.5 Occupation.

It was difficult to obtain details regarding the occupations of the participants as this was not a mandatory question that patients were required to answer when they visited the GUM clinic. Hence, occupational data (not shown) were only obtained from 188 of the 375 positive participants (103 females and 85 males); this was approximately half of the participants who were positive. In total, there were 58 different occupations amongst 188 individuals and most occupations only had one patient who worked in that particular job. It was therefore not possible to obtain an overall picture of the occupations of the participants represented in this study. In the 50% of people who did provide information on their occupation, the most prevalent groups were students and the unemployed. However, without knowing the occupations of the other 50% of participants it was not possible to draw any conclusions from these data.

a

Southampton's ethnic population breakdown

	All People	White	Irish	Other White	Black or Black British	Asian or Asian British	Mixed	Chinese or Other Ethnic Group
Number	236,700	193,384	1,894	10,178	5,207	15,149	5,207	5,441
Percent		81.70%	0.80%	4.30%	2.20%	6.40%	2.20%	2.30%

Source: Office for National statistics 2010

b

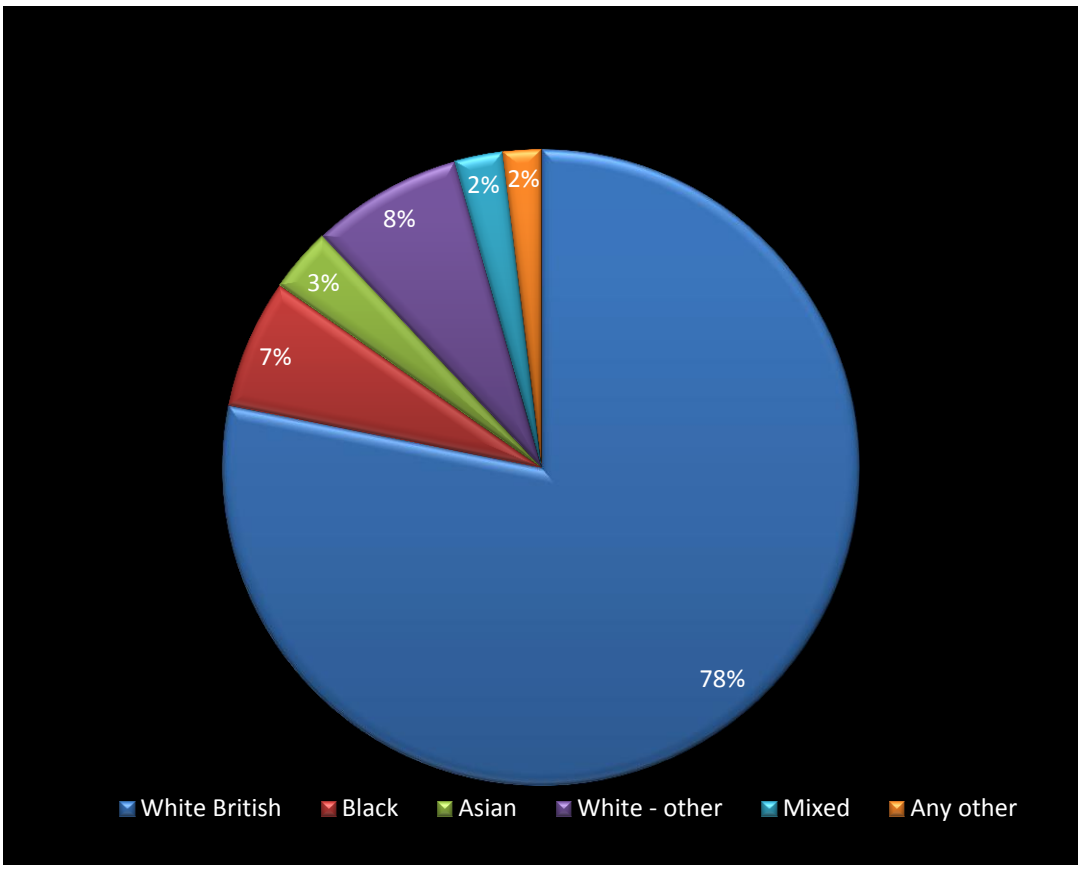


Figure 5.8 Distribution of the ethnicities of people in Southampton according to National Statistics figures (a) and a pie chart of the distribution of participants in this study (b).
Black includes individuals who are black Caribbean, black African and black other.
Asian includes Chinese, Indian and Asian other.
Mixed includes Black/white African, Black/White Caribbean, White/Asian.
Any other includes the travelling community and any other race which do not fall into the other categories.

5.13.6 Residential postcodes of participants: are there MLVA-*omp*A types of chlamydia which circulate in certain parts of Southampton?

The purpose of obtaining postcode information from the participants was to determine if there was a pattern amongst participants in relation to specific areas of Southampton and to determine for example if there was a difference between affluent and less affluent regions. The area of residence of participants who gave consent for this information to be included in this study was plotted on a map of the UK (Figure 5.9). This was done so that it would be easier to visualise the areas of Southampton where the positive participants resided. The locations were plotted on a map of the United Kingdom using a program available on the website www.spatialepidemiology.net. Figure 5.9 shows the location of the female and male participants who were positive in this study. The locations of the female participants are shown using white pins on the map whilst the locations of the male participants are shown using the red pins on the map. Most of the participants were from Southampton and the surrounding areas. However, there were some participants who attended the GUM clinic and lived afar afield as Cumbria, Lanchashire and Stafford. To check if the outliers on the MSpT (i.e the MLVA-*omp*A types which were not linked to the rest of the MSpT) were MLVA-*omp*A types from individuals who were from other areas outside Southampton, the areas in which the individuals who had these types lived were identified. All of the individuals whose samples had a MLVA-*omp*A that did not link to the rest of the network were from individuals who lived in Southampton. Table 5.6 shows the postcodes, areas and the number of male and female participants whose sample were positive for *C. trachomatis*. Above the solid black line in the table are all the SO postcodes which show the areas in Southampton and below the solid black line are the other postcodes for participants living elsewhere in the UK. These participants were included in the study as it was important to gain an idea of how far afield patients live when conducting future studies and also important when tracking new strains of chlamydia to determine exactly who has each type. In total there were 177 postcodes obtained from 187 men and 188 postcodes from 193 females who were all positive (Postcodes for ten men and five women who were positive for *C. trachomatis* were not obtained). Also included in the table are SO postcodes which include the New Forest areas. However, the new forest areas are not included on the maps showing the index of deprivation - Figure 5.10 and Figure 5.11.

To test the hypothesis that the positivity rate was higher in certain areas according to affluence, data on the index of multiple deprivation for Southampton were obtained. Although these data date back to 2010, it was assumed that the level of deprivation in each area of Southampton would have remained stable during the past three years. Figure 5.10 shows the names of the Wards in Southampton whilst Figure 5.11 shows a map of Southampton and the level of

deprivation relative to the rest of England. The index of multiple deprivation is measured by calculating the average of income deprivation, employment deprivation, health deprivation, disability, skills and training, education, crime and living environments, barriers to housing and services. The red areas of the map are indicative of the most deprived areas in Southampton whilst the dark green areas indicate the least deprived areas of Southampton.

According to the data collected, the top four regions with the highest rate of *C. trachomatis* in this study include Bargate (SO14), Freemantle (SO15), Shirley (SO16) and Bevois (SO17). According to the map showing the index of deprivation in Figure 5.11 parts of these areas are either in the 20% most deprived (coloured red) or the 20 to 40% most deprived (coloured orange) regions of Southampton. The number of male and female participants from each postcode region, who came to the clinic and tested positive are equal or close to equal numbers for both genders apart from the SO16 postcode (Shirley area) where there were 19 male and 33 females from this region who tested positive for chlamydia (Table 5.6).

To determine if there was any correlation between the place of residence and the MLVA-*omp*A genotype the different regions were grouped together into three groups. Firstly, all the Southampton postcodes were grouped together, as were the postcodes which started with SO but were not Southampton postcodes but the SO at the start of the postcode indicated that these areas where in close proximity to Southampton. The third group was composed of all the other individuals who lived outside Southampton and did not have a postcode which began with SO. Following statistical analysis there was no correlation between where individuals resided and the MLVA-*omp*A types ((The P values were 0.599 (Two segment analysis), 0.767 (Three Segment analysis) and 0.501 (Ten segment analysis)).

A

B

Figure 5.9 Regions where participants represented in this study reside.
The regions were the female participants reside are shown using white pins (A), regions were the male participants reside are shown using red pins (B). Note that if more than one participant has the same first four digits of their postcode this is represented by just one pin on the map.

Postcode	Region	Male	Female	Total
SO14	Bargate	35	35	70
SO15	Freemantle	35	37	72
SO16	Shirley	19	33	52
SO17	Bevois	18	17	35
SO18	Harefield	14	10	24
SO19	Sholing	13	17	30
SO23	St Michael, Winchester	1	0	1
SO30	Bitterne	2	2	4
SO31	Warsash, Fareham	3	4	7
SO40	Copythorne South and Ashurst, New Forest	5	6	11
SO41	Pennington, New Forest	1	1	2
SO43	Bramshaw, Copythorne North and Minstead, New Forest	0	1	1
SO45	Furzedown and Hardley Ward, New Forest	9	7	16
SO50	Eastleigh	3	3	6
SO51	Abbey Ward, Test Valley	4	3	7
SO53	Chandlers Ford, Eastleigh	1	3	4
BH12	Branksome West, poole	0	1	1
BH25	Milton, New Forest	0	1	1
BH8	Queen's Park, Bournemouth	0	1	1
BH9	Winston East, Bournemouth	0	1	1
BN3 7	Central Hove, The city of Brighton and Hove	1	0	1
BR2	Bromley	1	0	1
BS25	Banwell and Winscombe, North Somerset	1	0	1
CA3	Stanwix Urban, Carlisle	1	0	1
CM23	Bishop's Stortford All Saints, East Hertforshire	0	1	1
EX24	Coly Valley, East Devon	0	1	1
LU5	Houghton Regis, Central Bedforshire	1	0	1
NG17	kirkby in Ashfield West, Ashfield	1	0	1
OX11	Hagbourne, South Oxfordshire	0	1	1
PO12	Alverstoke, Gosport	1	0	1
PO14	Stubbington, Fareham	1	0	1
PO15	Fareham North-West, Fareham	0	1	1
PR7 2	Chorley South West, Chorley	1	0	1
RG24	Popley East, Basingstoke and Deane	0	1	1
RH2 8	Brockham, Betchworth an Buckland, Mole Valley	1	0	1

Table 5.6 Shown are the postcodes, region in the UK and numbers of males and females who reside in each region who participated in the study and were positive for *C. trachomatis.* The districts and the wards for each postcode were sourced by typing the four digit post code provided by each participant into the website www.postcodearea.co.uk which provided information on every city in the UK.

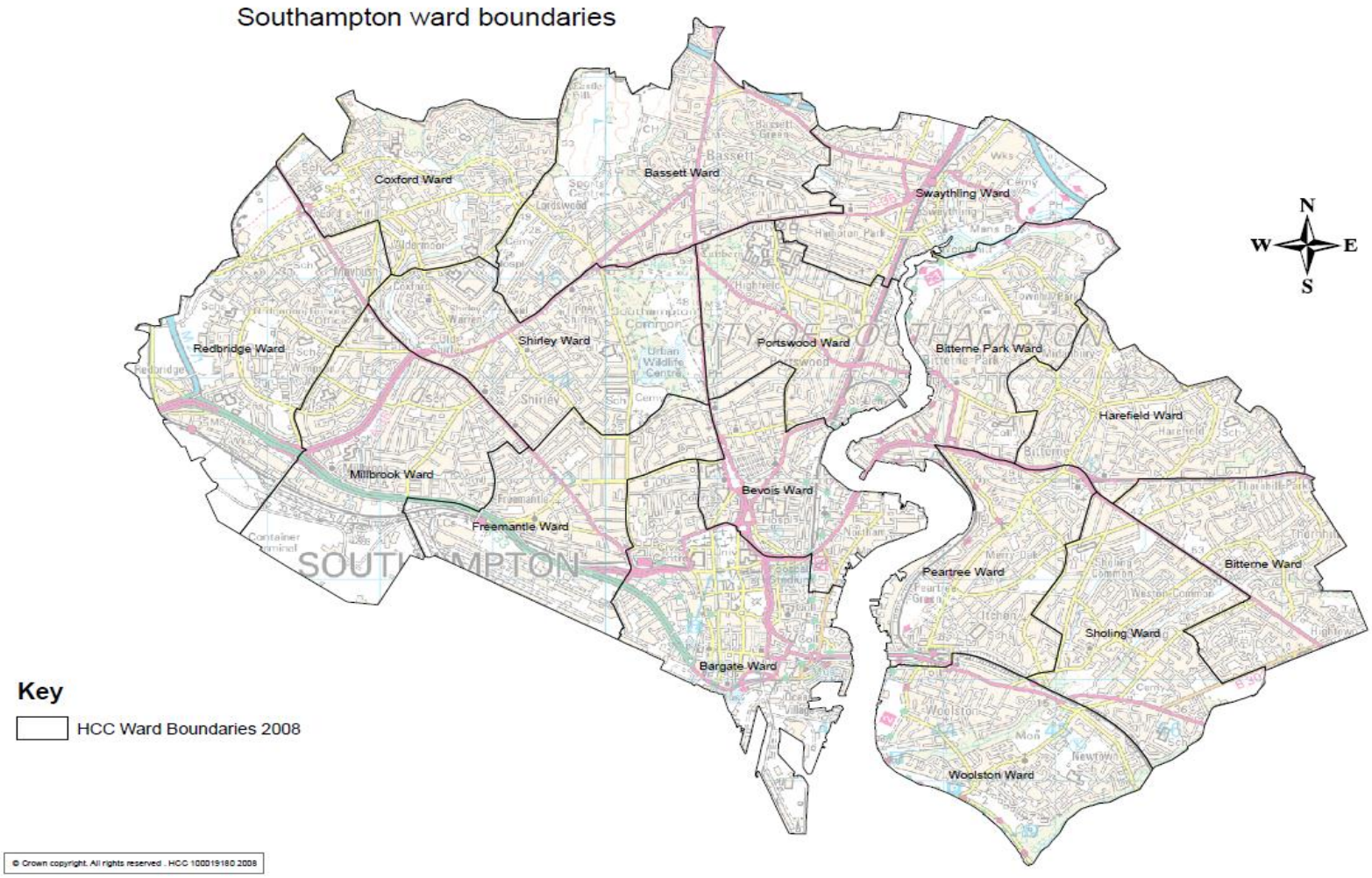


Figure 5.10 The Southampton Ward Boundaries

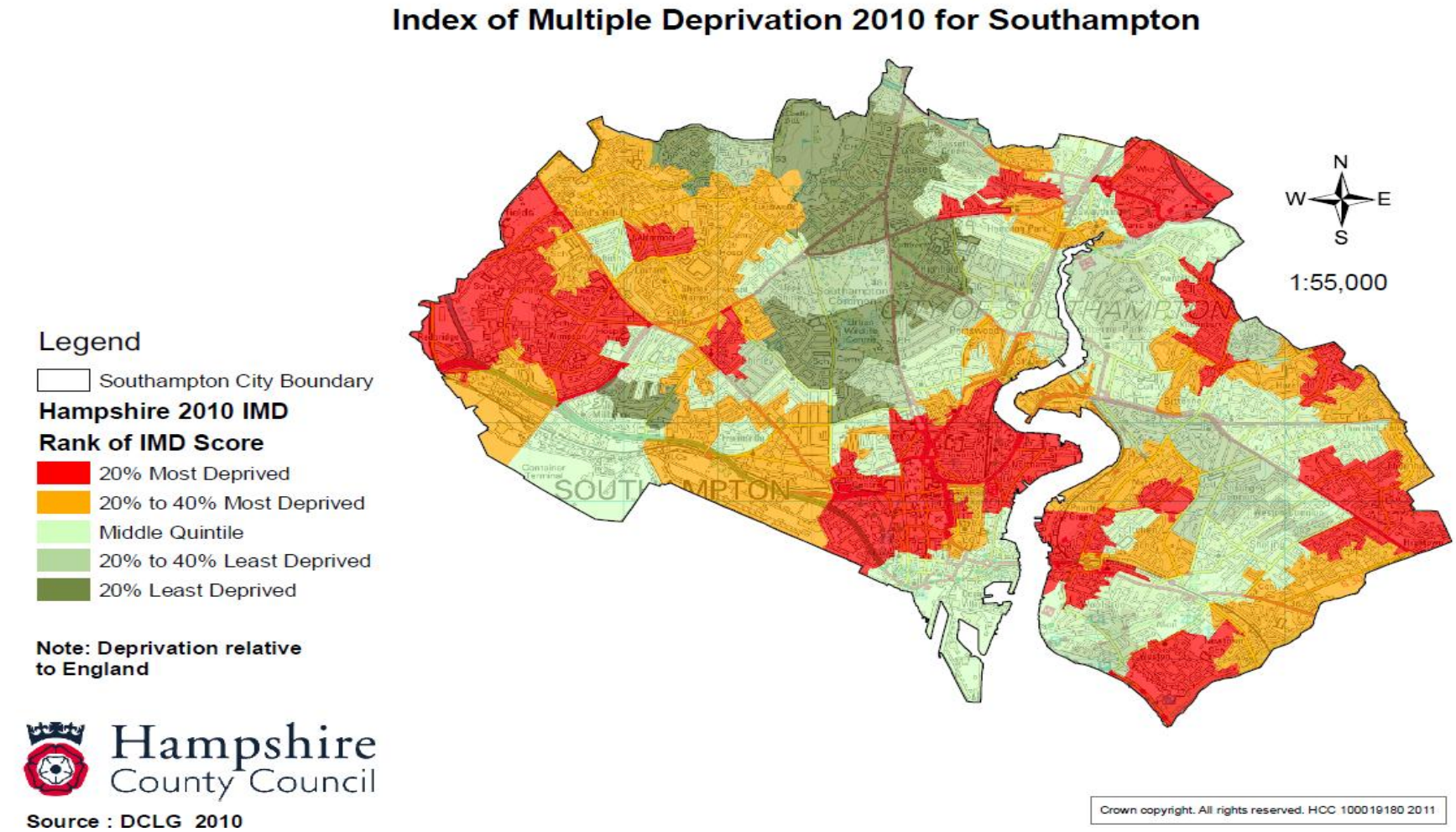


Figure 5.11 The Index of multiple deprivation for Southampton in 2010.
Source: The Department for communities and local Government website.

5.13.7 Comparison of demographic data between those who were positive and those who were negative.

One of the aims of this project was to compare demographic information from patients who attended the same GUM clinic and gave samples thatwere negative and those who were positive for *C. trachomatis*. This information could aid in targeting groups of people who are more likely to be infected with *C. trachomatis* by providing education and advice on how they can prevent themselves from contracting chlamydia. As there was approximately 90% more individuals who were negative for *C. trachomatis,* only individuals who had given demographic information were included when trying to choose a comparative negative group. In total there were 531 females and 517 males who fell into this category. The demographic data that were compared between individuals who were positive and those who were negative included age range, gender, post code, occupation and symptomatology.

There was no significant difference between gender (0.388) with regards to whether the individual was positive or negative for *C. trachomatis*. However, the age of the participants was significant with a $p<0.001$. A higher percentage (73.77%) of positive chlamydia tests were from individuals between the ages of 16-24 years old, whilst there was a higher percentage of individuals who were 25 years and above who were negative for *C. trachomatis* than were positive. These data show that individuals are more likely to be positive for *C. trachomatis* if they are between the ages of 16-24 and negative if they are 25 years and older. There was also a significant difference between symptomatology (P=0.038), area of residence (P=0.043) and ethnicity (P=0.037). Individuals who had a negative *C. trachomatis* test were more likely to be asymptomatic than those who had a positive *C. trachomatis* test. Additionally, a higher percentage (77.4%) of individuals with a Southampton postcode were positive compared to people living in regions outside Southampton. Furthermore, individuals were more likely to be positive for *C. trachomatis* if they were white British. However, this difference is possibly due to the fact that more white British individuals visited the GUM clinic compared to other ethnic groups.

5.14 SYMPTOMS PRESENTED BY PARTICIPANTS.

The majority of people who are positive for *C. trachomatis* do not present with symptoms (Meyers *et al.*, 2007, Geisler, 2010). To determine the percentage of participants who were asymptomatic in my study population, patients were asked if they had any *C. trachomatis* related symptoms and were also examined to determine if they had any signs of infection. In total, 48% of positive patients were asymptomatic. Symptoms recorded include dysuria, increased discharge, post-coital bleeding, irregular bleeding, menorrhagia and dyspareunia/pelvic pain in females. For males, the symptoms were dysuria, discharge, odd sensation or irritation in urethra and testicular or epididymal pain. 110 out of 191 (57.6%) females who took part in the study were asymptomatic for chlamydia. The most common symptoms found in females were increased discharge, dyspareunia/pelvic pain, dysuria and irregular bleeding which had 18 people who presented with these symptoms. For men the main symptoms were dysuria and discharge. The percentage of men who were asymptomatic was 46.8% (86 out of 184). These figures are slightly lower than figures seen in other studies, where up to 70% of women and 50% of men were asymptomatic (Meyers *et al.*, 2007). Table 5.7 shows the various symptoms which were reported for this study and the number of males and females that presented with each symptom. The number of males/females who were asymptomatic is also shown in the table. Following statistical analysis there was no significance between the MLVA-*omp*A type and if the patients were symptomatic or not ((The P values were 0.260 (Two segment analysis), 0.421 (Three Segment analysis) and 0.273 (Ten segment analysis)). There was also no significant difference between the MLVA-*omp*A type and the type of symptom the individuals presented with, with exception to dysuria which showed a significant difference when the two segment analysis was conducted (P=0.038). A higher percentage of individuals in segment one presented with dysuria compared to segment two. P values are shown in table 5.7.

	Symptoms	No of. People	P values for each segment analysis		
			Two Segment	Three segment	Ten Segment
Female Symptoms	Dysuria	15	0.038	0.501	0.631
	Increased Discharge	38	0.402	0.605	0.105
	Post-Coital Bleeding (PCB)	7	0.976	0.421	0.275
	Irregular bleeding	15	0.532	0.097	0.563
	Menorrhagia	0	0.452	0.156	0.461
	Dyspareunia/Pelvic pain	18	0.978	0.595	0.238
	Asymptomatic	110	-	-	-
Male Symptoms	Dysuria	58	0.564	0.951	0.876
	Discharge	41	0.324	0.245	0.762
	Odd sensation or irritation in urethra	9	0.710	0.625	0.445
	Testicular or epididymal pain	2	0.565	0.132	0.299
	Asymptomatic	86	-	-	-

Table 5.7 Symptoms found in male and female patients and the number of patients with those symptoms. Also shown are the P values for each segment analysis.

5.15 DETERMINING IF THERE WAS ANY SIGNIFICANT DIFFERENCE BETWEEN THE MONTH THE SPECIMEN WAS COLLECTED AND THE MLVA-*OMP*A TYPE.

To determine if there was any significance between the date the specimens were taken and the MLVA-*omp*A types, the dates on which the specimen were sent were grouped into the months the specimens were collected. The assessment was carried out for each of the segment analyses described in section 5.12.3 to determine if the genotypes in that segment were collected in a particular month as over time new MLVA-*omp*A types may be introduced in to the community and also new sexual networks may form. This is useful information as it may suggest that the samples in that part of the MSpT are from individuals that are part of the same sexual network. The Pearsons Chi squared test was conducted to determine if there was any correlation between the date the specimen was sent and the genotype of the sample. For the analysis where the MSpT was split into two and ten segments the P values were P=0.006 and P=0.050 respectively. This showed that there was significant correlation between the date the specimen was taken and the segment the MLVA-*omp*A genotype for each sample occurred. On the other hand, there was no significance when the three segment analysis was assessed.

To visualise this information, the month each sample was collected was shown using a different colour on the MSpT (Figure 5.12). Notably, there were some MLVA-*omp*A types which were assigned to samples collected in the same month or in the subsequent month. An example of this is shown using a box on figure 5.12. The demographic data from these four heterosexual individuals (three males and one female) were analysed (Table 5.8) to determine if this potential network, identified as such on the basis of all the samples having the same MLVA-*omp*A type could be confirmed as a sexual network using the demographic information provided by the individuals. The age range of the four individuals was 17-31 and they were from different ethnic backgrounds. Three of the individuals were asymptomatic and the other symptomatic, whilst three individuals lived in Southampton and the other in Fareham. This shows that although the typing data suggest that these individuals are part of the same sexual network, no conclusions can be drawn using only the demographic data without partner identification.

There were also some MLVA-*omp*A types that formed large founder groups with several samples collected in different months. Within these MLVA-*omp*A types there were some samples collected in the same month. An example of this is shown on figure 5.12 with an arrow. These data may indicate that the individuals with these MLVA-*omp*A types are part of the same sexual network. This example includes 21 samples with the MLVA-*omp*A type 8.5.1-E (the primary founder type of

the complex) (Table 5.9) from white British male and female participants. Some potential partners were identified from the data. For instance, patients with study numbers 95 (Male) and 134 (female) were tested in November, are both the same age and live in close proximity to each other. There were also six samples collected in March with this MLVA-*omp*A type from individuals who were between 18- 24. These individuals all attended the clinic over a one week period in March (19th March 2013 – 27th March 2013), all of whom resided in areas with the postcodes beginning with SO15 – SO16 (for one individual there is no postcode data). These data show that although it cannot be guaranteed due to the lack of partner identification , it is very likely that these individuals are part of the same sexual network.

Additionally, there were various SLVs of some MLVA-*omp*A types which were collected in the same month. An example of this is shown using stars on figure 5.12. In the example the samples collected in October are shown in red. There are SLVs of the MLVA-*omp*A type which were also assigned to samples collected in October, suggesting that these individuals may be part of a sexual network where the MLVA-*omp*A type has begun to evolve, or the network contains a collection of very closely related genotypes. The demographic data for individuals in this potential network is shown in table 5.10. All of the individuals in this example are of white British and white other ethnicity. They are all heterosexual male between the ages of 19-23 and all but one of the individuals lived in close proximity (SO14-SO18). As this example only contains demographic data from seven heterosexual males it shows that if this is data from a potential sexual network then there are missing members of the network as there are no females in this example and all the men in this example claim to be heterosexual which again highlights the need for partner identification to identify sexual networks.

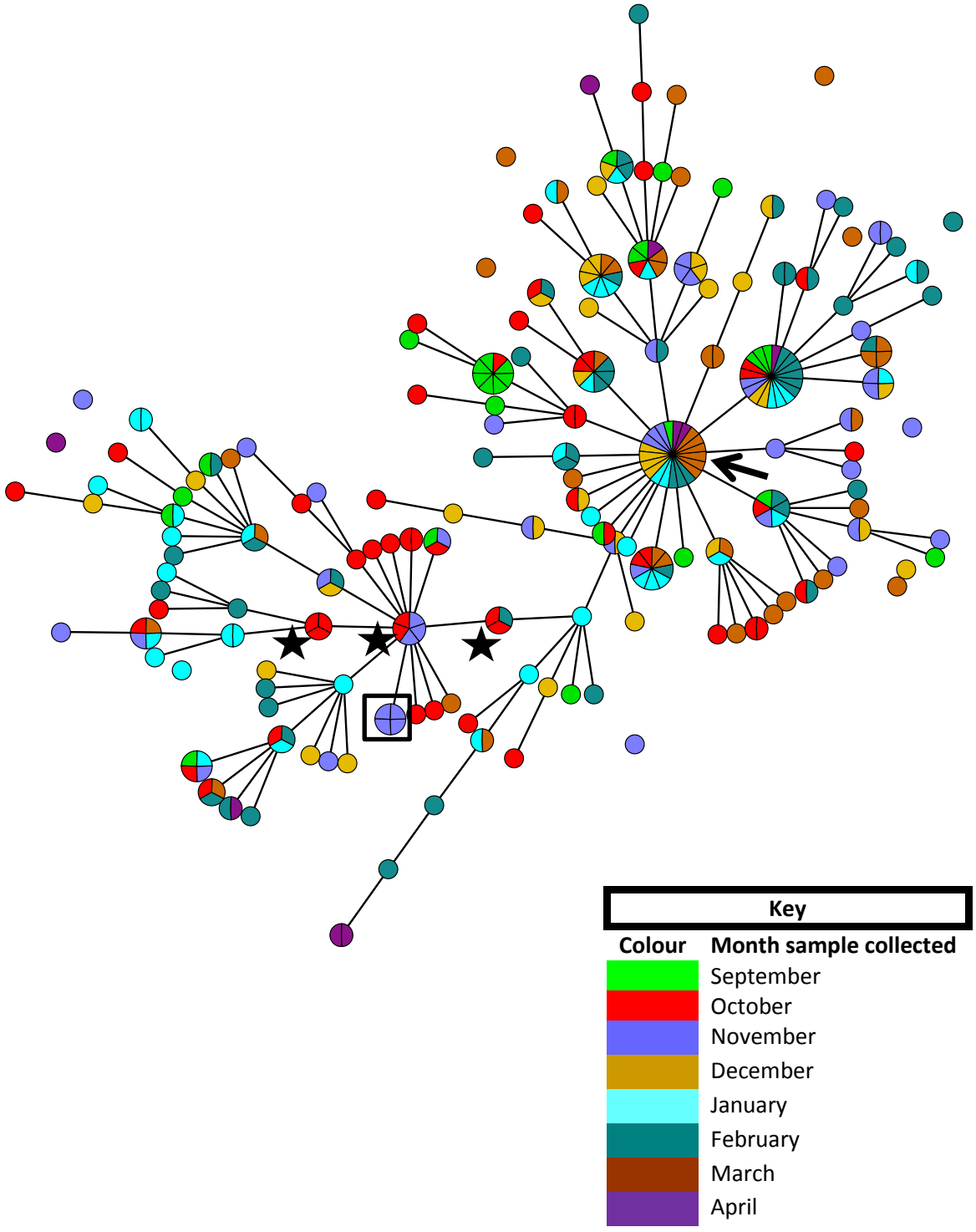


Figure 5.12 Minimum spanning tree showing the months the samples were collected.
The box shows a MLVA-*omp*A type where all of the samples with this type were collected in the same month.
The stars show single locus variants of samples collected in October.
The arrow points to a MLVA-*omp*A type with several samples collected in different months. Within this MLVA-*omp*A type there are also some samples collected in the same month.

Study ID No.	Gender	Postcode	Ethnicity	Asymptomatic	Age	Sample collection date
89	Male	SO15	Black African	Symptomatic	31	01-Nov-12
99	Male	SO19	White British	Asymptomatic	29	12-Nov-12
101	Male	SO15	White/ Black Caribbean	Asymptomatic	21	12-Nov-12
111	Female	PO15	White British	Asymptomatic	17	19-Nov-12

Table 5.8 Demographic data from four individuals who had the same MLVA-*omp*A type of chlamydia and who were all tested in November.

Study ID No.	Gender	Postcode	Ethnicity	Asymptomatic	Age	Sample collection date
3	Female	SO31	White British	Symptomatic	19	07-Sep-12
95	Male	SO15	White British	Symptomatic	23	02-Nov-12
107	Female	SO53	White British	Symptomatic	21	09-Nov-12
134	Female	SO16	White British	Asymptomatic	23	09-Nov-12
133	Female	CM23	White British	Symptomatic	23	05-Dec-12
162	Male	SO15	White British	Symptomatic	24	10-Dec-12
174	Male	CA3	White British	Symptomatic	23	19-Dec-12
175	Male	SO16	White British	Asymptomatic	23	20-Dec-12
186	Female	SO16	White British	Asymptomatic	22	03-Jan-13
226	Female	SO14	White British	Asymptomatic	18	22-Jan-13
283	Female	SO14	White British	Symptomatic	21	19-Feb-13
289	Female	SO15	White British	Asymptomatic	24	21-Feb-13
299	Male	SO45	White British	Symptomatic	41	21-Feb-13
359	Male			Symptomatic	22	19-Mar-13
353	Female	SO16	White British	Asymptomatic	18	20-Mar-13
360	Male	SO18	White British	Asymptomatic	19	20-Mar-13
356	Female	SO17	White British	Asymptomatic	18	22-Mar-13
357	Female	SO16	White British	Symptomatic	24	26-Mar-13
370	Male	SO15	White British	Asymptomatic	24	27-Mar-13
376	Female	SO18	White British	Symptomatic	20	11-Apr-13
379	Female	SO14	White British	Asymptomatic	20	12-Apr-13

Table 5.9 Demographic data of 21 individuals who had the same MLVA-*omp*A type (8.5.1-E).
Samples are ordered according to the date the sample was collected. The table is sectioned according to the month the sample was collected.

Study ID No.	MLVA-ompA type	Gender	Postcode	Ethnicity	Asymptomatic	Age	Sample collection date
32		Male	SO14	White British	Symptomatic	19	02-Oct-12
77	3.6.1-D	Male	SO15	White British	Symptomatic	22	19-Oct-12
78		Male	SO14	White Other	Symptomatic	22	22-Oct-12
34	3.3.4 - D	Male	BS25	White British	Symptomatic	21	09-Oct-12
73		Male	SO15	White British	Asymptomatic	21	17-Oct-12
74	3.3.4-E	Male	SO14	White British	Asymptomatic	23	18-Oct-12
79		Male	SO16	White British	Symptomatic	20	23-Oct-12

Table 5.10 Demographic data from individuals with chlamydia with single locus variants of a MLVA-*ompA* type all collected in October.

5.16 PATIENTS WHOSE PARTNERS ALSO HAD *C. TRACHOMATIS* POSITIVE SAMPLES.

There were 32 females and 54 males whose partners were also positive for *C. trachomatis*. However, there was not enough data to identify if these partners had given consent to be part of this study. Additionally, there were no data recorded on the patient numbers of the partners therefore no links could be made as to who these individuals were. For future studies these will be useful data to collect so that genotypes of chlamydia can be checked between partners to determine if they have the same types and this can be added to the analysis of sexual networks and used to analyse how certain types are transmitted through the community.

5.17 PARTICIPANTS WHO ATTENDED THE GUM CLINIC MORE THAN ONCE WITH REPEATED INFECTIONS DURING THE STUDY PERIOD.

During the study period only two participants had *C. trachomatis* positive samples at two visits. These patients had received treatment after the first diagnoses and presented with chlamydia at a subsequent visit. This could mean that the patients had been re-infected with *C. trachomatis* (from either the same partner who had not been treated –this can be determined by the MLVA-*omp*A type at both visits or from the same partners who had contracted chlamydia from someone who had a different MLVA-*omp*A type of *C. trachomatis*) or that they had persistent *C. trachomatis*. The patients who returned for a subsequent visit during the study period were both female (one returned one month after the first infection, whilst the other returned after two months), one of which was a white British female and the other was of black African descent. For both participants the MLVA-*omp*A types assigned were different at both visits which suggests re-infection from either a different partner to the one that they contracted chlamydia from in the first visit or from the same partner who had contracted chlamydia from someone else. Interestingly, the females both had symptoms at both of their visits and this could be the reason for their return during a short period of time.

5.18 CO-INFECTION WITH OTHER STIS.

Data on co-infection was also recorded to identify the patients who were infected with chlamydia and other STIs. The reason for recording other infections was to determine if co-infection could be linked to symptomatology. For instance being infected with another bacteria or virus as well as chlamydia may lead to the presentation of symptoms. These data were also useful in showing the other infections that are found in *C. trachomatis* infected patients. In total 207 positive participants had no previous history of chlamydia or any other STI recorded as there was no previous record of them attending the clinic. This figure is compared to 168 people who had visited the clinic previously, of these people, 59 participants had previously presented with *C. trachomatis* on one or more visits. In total nine patients had presented with other infections

during past visits these infections include Ebstein barr virus (EBV), Cytolomegavirus (CMV), Hepatitis C (Hep C) and Herpes Simplex virus 1 (HSV-1).

Data on other infections that patients had at the time of visit were also recorded. In this study patients were co-infected with gonorrhoea (n=6), HIV (n=5), HepB (n=3), HSV-2 (n=2) and syphilis (n=1) at the time of testing. The number of participants co-infected during this study was low, nevertheless, the symptomatology of all seventeen individuals who were co-infected was analysed. The hypothesis that symptoms could be associated to co-infection was not proved as the majority of patients testing positive for *C. trachomatis* and co-infected with other pathogens did not have any symptoms.

To determine if the individuals who were co-infected with the same STI were part of the same network the MLVA-*omp*A types for each individual with each of the STIs were checked. As the MLVA-*omp*A types differed between individuals who were co-infected with the same STI it was concluded that these individuals were not in the same sexual network.

5.19 Comparison of genotypes between samples collected in Southampton in 2009 and samples collected in Southampton in 2012/2013.

In 2009 a pilot study was conducted to evaluate the practicality of the MLVA-*omp*A typing system (Wang *et al.*, 2011b). As the typing systems used in the Wang *et al* (2011b) study and my study were identical and as both studies were conducted in Southampton the distribution of *omp*A genotypes were compared between the two studies to determine if there were any differences in the *omp*A distribution in Southampton in 2013 compared to 2009. An in depth analysis was conducted between the *omp*A genotypes in both studies but could not be conducted in the same way for MLVA-*omp*A genotypes because in the Wang *et al* (2011b) study there were only 47 unique MLVA-*omp*A genotypes whereas in my study there were 151 unique genotypes. Therefore there were at least 104 MLVA-*omp*A genotypes in my study that were not found in the Wang *et al* study. However, most of the *omp*A genotypes were found in both years. For comparison of the MLVA-*omp*A types an MSpT was drawn including both samples from the study in 2009 and the study in 2013 (figure 5.13). In the Wang *et al* study, 93 samples were assigned full MLVA-*omp*A genotypes whereas in my study 302 samples were assigned a full *omp*A genotype. The most prevalent genotypes remained unchanged over the years in both studies and these were *omp*A genotypes D, E and F. Three genotypes which were not detected in 2009 were detected in this current study and these include *omp*A genotypes B3/IU-FQ0279, genotype L2 and genotype L2b. Table 5.11 shows the number and percentage of each *omp*A genotype for both studies of each

type found in both years. The percentage difference for each genotype was also calculated to analyse if there were any differences in *omp*A genotypes between the two years. The largest percentage difference was for genotype F/IC-CAL8, where 5.87% less samples were assigned this type in 2013 than in 2009. There was also a 4% increase in genotype E. All other differences between the two years were very small (<2%).

An MSpT was generated to investigate whether chlamydia MLVA-*omp*A types differ in Southampton in 2009 and 2013 (figure 5.13). The MLVA-*omp*A types for both sample sets formed a single network. There were only 23 MLVA-*omp*A types found in both years and the majority of types were unique to the years. The most common MLVA genotypes were the same in both years. In both years the most common MLVA genotypes were 8.5.1, followed by MLVA types 8.5.2 and 8.6.1. The most common MLVA-*omp*A types were also the same in both years and these were 8.5.1-E followed by 8.6.1-E. The majority of MLVA-*omp*A types with large number of samples (n≥5) were found in both years, these types are likely to be founder groups. There were some MLVA-*omp*A types with greater than four samples that were only found in samples collected in 2013 which shows new subgroup founders may be emerging over the years.

ompA types	2009 Number of each type	2009 (%)	2013 Number of each type	2013 (%)	Percentage difference from 2009 to 2013
B3/IU-FQ0279	0	0	3	0.99	0.99
D/UW-3	11	11.83	33	10.93	-0.90
D/IC-CAL8	8	8.60	28	9.27	0.67
E/Bour	38	40.86	136	45.03	4.17
F/IC-CAL3	15	16.13	31	10.26	-5.87
G/392	6	6.45	25	8.28	1.83
H	0	0	3	0.99	0.99
Ia/870	4	4.30	8	2.65	-1.65
J/UW-36	4	4.30	10	3.31	-0.99
K/UW-31	7	7.53	24	7.95	0.42
L2	0	0	0	0	0
L2b	0	0	1	0.33	0.33
Total	93	100	302	100	

Table 5.11 Comparison of the *ompA* genotypes recorded from 2009 and 2013.
Also shown are the percentage of each of the *ompA* genotypes in comparison to the total number of fully genotyped samples in each of the studies and the percentage difference between both years.

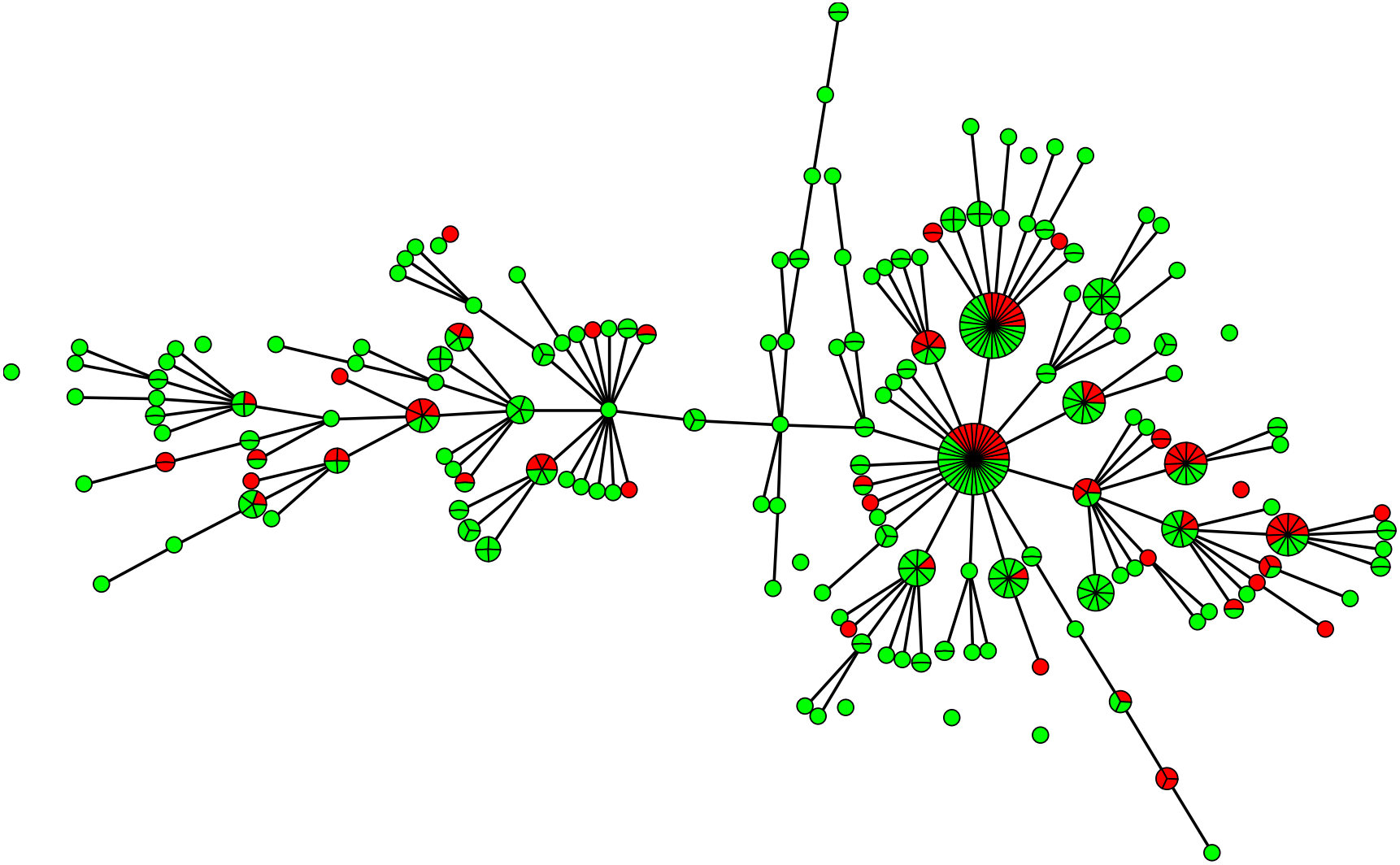


Figure 5.13 Minimum spanning tree generated using MLVA-*omp*A typing data of samples collected in Southampton in 2009 and in 2013.
Samples collected in 2009 are shown in red. Samples collected in 2013 are shown in green. The size of the spheres is indicative of how many samples make up each sphere. All samples within a sphere have the same MLVA-*omp*A type. The branches show the relationship between single locus variants.

5.20 DISCUSSION

In 2009 a pilot study was conducted to evaluate the MLVA-*omp*A typing scheme (Wang *et al.*, 2011b). The aim of this study was to evaluate a high resolution typing system for *C. trachomatis* on a population of women in Southampton who were positive for *C. trachomatis*. In that study a total of 157 samples were collected, 59% of those samples were fully genotyped using the MLVA-*omp*A typing scheme. In addition, *C. trachomatis* was isolated in cell culture from 43% of the samples collected. This pilot study showed that it was feasible to obtain samples from the GUM clinic in Southampton and type them using the MLVA-*omp*A typing scheme and that this approach was suitable for a detailed epidemiological study. Following completion of the evaluation carried out in 2009, which showed that the approach was feasible and it was possible to do detailed studies and address new questions using the MLVA-*omp*A typing scheme, it was hypothesised that there may be MLVA subtypes which may reflect particular demographics such as age group, location, and high-risk sexual behaviour. The identification of genotypes associated with particular demographic groups may lead to the identification of sexual networks within communities. This led to the study described in this chapter. The main aims of my study were to assess the distribution of *C. trachomatis* genotypes, to define sexual networks in Southampton, and to study the epidemiology of *C. trachomatis* by analysing demographic data from participants in Southampton. The ability to identify sexual networks for *C. trachomatis* enables the determination of specific strains which are being transmitted through a population and also allows for the identification of how many individuals are transmitting a particular genotype. Due to unpublished data from the study in 2009 which showed that there were samples with the same MLVA-*omp*A types collected around the same date it was initially thought that we could identify partners based on the MLVA –*omp*A type and by collecting demographic data.

Differences in the sequence of the *omp*A gene have commonly been used to assign different genotypes to samples in studies (Jurstrand *et al.*, 2001, Geisler *et al.*, 2002). The MLVA-*omp*A typing system which includes genotyping of the *omp*A gene was used to assign genotypes to samples in my study. As *omp*A on its own is commonly used for genotyping, it was important to compare the data from this study with that of other studies. Additionally, the MLVA-*omp*A typing system has only been used to assign genotypes to samples in two other studies (Pedersen *et al.*, 2008, Wang *et al.*, 2011b).

To establish how many samples would be required to fulfil the aims of the study and to make it statistically significant the help of a Medical Statistician, Mr Scott Harris was sought. After consulting with him, he suggested that during the study period of eight months, 100 positive samples could be collected per month and that would be an adequate number of samples to base

the statistical analysis on based on the calculated precision with which the prevalence of *omp*A genotypes could be estimated (section 5.4). The figure of 600 to 800 samples was attained by analysing previous data on positivity rates of chlamydia from the GUM clinic in Southampton during May 2011 to April 2012. This data showed that chlamydia was detected in 1262 patients during that period and a further 113 patients had equivocal samples. Therefore, it was estimated that 100 positive participants per month was a reasonable number of samples to expect. However, it was not taken into consideration that not all of the positive samples would be obtained or that some patients would not give consent for their sample to be included in this study. There were also several complications with obtaining some of the samples from patients who had given consent for their samples to be included in this study and this included; positive samples being discarded after diagnostic tests but before collection could take place, collection of some samples was not always possible because it was sometimes necessary to repeat diagnostic tests due to some samples having an equivocal result during diagnostic tests, due to sample quality and ocassionally due to the consultants wanting the test repeated on behalf of the patient. After repeating the diagnostic test, at times there was no transport media in which the sample had been placed to extract DNA from and for this reason some samples were not stored and were immediately discarded. For all these reasons the target of obtaining between 600 to 800 samples was not met. In total 380 positive samples were obtained during the eight month study period which is approximately half of the number of samples expected initially. As the expected confidence intervals, estimated for each type was not applicable as the target number of samples was not obtained, this study therefore acted as the basis to conduct larger full scale studies in the future.

To assign genotypes to patient samples and to include analysis of participant demographic data in my study, each participant was required to provide informed consent. However, consent from every patient who attended the GUM clinic was not obtained as some individuals did not want to participate in the study or nurses/doctors had forgotten to obtain consent from patients. This was also another reason why the estimated target of samples was not achievable. For the individuals who had given consent for their sample to be used in this study, it was not possible to obtain all the demographic data from these individuals as not all of the participants gave information for all the demographic questions. The occupation of each participant seemed to be omitted in many cases as this was not a question that was routinely asked. For future studies a question asking for the occupation of patients should be added to the questionnaire as types of chlamydia may circulate within individuals with certain type of occupations as these individuals are more likely to have met at work and formed relationships.

In the Wang *et al* (2011b) study, a proportion of the samples did not yield full genotyping data, therefore it was expected that in this study full genotyping data for all samples would also not be obtainable. However, as it was essential to have as many of the samples with full genotypes to gain a clear idea of the sexual networks in Southampton, a second set of primers were used in the PCR process as described in section 4.6.1. This ensured that a higher percentage of samples were typed than was possible in the paper by Wang *et al* (2011b) where only one set of primers were used to assign MLVA-*omp*A types to samples. One of the limitations of my study included being unable to culture from the samples because the swabs were placed in a lysis buffer which denatured the chlamydia before it was received in the laboratory for culturing, so unlike the previous study published in 2009, this meant that chlamydia could not be isolated in cell culture and typing of the samples had to be conducted directly from the clinical samples received from the GUM clinic via the HPA. This also had further implications as the samples couldn't be whole genome sequenced because for whole genome sequencing isolation of *C. trachomatis* is a pre-requisite to obtain the concentration of DNA required for this process.

Collection of samples for my study took place during an eight month period (September 2012 – April 2013). Of the 380 positive samples collected, 302 samples were assigned full MLVA-*omp*A genotypes. The most common MLVA-*omp*A types were unchanged from the study by Wang *et al* (2011b). However, the most common MLVA types were different. This demonstrates the impact of the inclusion of the *omp*A gene in the typing system and shows the difference that the addition of an additional marker makes on studying epidemiology especially when making comparisons between the presence of genotypes over time. The highest rates of chlamydia were diagnosed in individuals between the ages of 16 to 24. Positivity rates were higher in younger individuals possibly due to a range of biological and behavioural reasons. Younger individuals are more likely to be unmarried, to change partners frequently, to have multiple partners, and to be less educated on the use of condoms. The number of participants with no previous history was greater than those with a previous history of chlamydia and this may be due to the fact that 72% of patients in this study were between the age of 16-24 and may have only just become sexually active. Therefore, this may be their first visit to the GUM clinic. Interestingly, there were some patients who had given two specimens for diagnosis (i.e a swab and a urine sample) and the diagnostic test results showed that only one sample was positive and the other negative. This highlights the need to collect various specimens from patients to maximise the chance of detecting an infection if there is one present and providing the necessary treatment, if required.

In epidemiological studies of *C. trachomatis* between 50 to 88% of patients are asymptomatic (Kalwij *et al.*, 2010). In my study 48% of positive participants were asymptomatic. There were also two females who returned for a subsequent visit during the study period but presented with different MLVA-*omp*A types at both visits this was interesting but not unusual as a recent paper has shown that out of the seven participants who presented with *C. trachomatis* at two or more visits within their study period, only one of the seven participants had the same MLST type at both visits (Bom *et al.*, 2013a). This shows that participants are infected with a different strain of chlamydia possibly because they have multiple partners, may have recently changed partners or their partners are having sex with someone else.

Data obtained on where the participants in this study reside, has shown that most of the positive samples collected collected from participants who reside in the most deprived areas of Southampton. However, this is only true when considering GUM clinic data. A limitation with the data collected is that large populations of individuals are missing from the analysis. Those who were tested at private healthcare settings, those who attend their GP practices and those who were tested at other walk- in-centres in Southampton were not included in this study. Therefore, for future studies it may be important to include data from other healthcare settings as this would aid in better understanding whether *C. trachomatis* positivity is more likely to be detected in patients who live in more deprived areas. Locations of participants were plotted on a map of the UK (Figure 5.9). A limitation of using this map was that it does not show the location of every participant, so when the post code of an individual was the same as others in the study just one pin was plotted on the map representing all participants with the same first four digits of their postcode. This was not an issue as the areas where the participants lived could still be determined and the purpose of using this map was not to determine the rate of positivity in each area but rather to determine the areas in which positive chlamydia samples were obtained. The majority of individuals with chlamydia in this study lived in areas with postcodes beginning with SO14 to SO19. However, these are the locations in and around the GUM clinic from where (located in S014) the samples were collected and therefore people who live in other areas may choose to attend a centre where they can get tested that is closer to where they live. This is another reason why for future studies it would be ideal to collect samples from a range of healthcare settings in and around Southampton. When analysing these data it was also important to acknowledge that many areas in Southampton have a mixture of both affluent and less affluent people in the community, this is shown by the type of housing in these areas (some of which are privately owned and others which are owned by the council) and the occupations of individuals who reside in these areas. Additionally, chlamydia is mainly detected in individuals who are between the ages

of 16-24 who are the young adults still in college, or University or have just started their career and are therefore more likely to be living in the more deprived areas of Southampton due to their low income. Although occupations were only recorded for 188 participants in this study it was evident that the GUM clinic was mainly used by students (n=79) and the unemployed (n=38). The data (not shown) collected on occupations indicated that the individuals that use the GUM clinic are either unemployed or they are in low paid jobs. Therefore, it can be concluded that the more affluent individuals go elsewhere to get tested for chlamydia. As only the first four digits of the postcodes were obtained from participants (as we were unaware at the time that the protocol was written that the first four digits would not provide enough information) a more exact IMD score was unattainable for each postcode. Instead more general IMD scores were obtained by using the index of deprivation map provided by the Hampshire council for Southampton (Figure 5.11), this approach takes into account the averages of deprivation scores of regions in Southampton and does not show individual information on postcodes. For future studies ethical permission should be obtained to collect data on full postcodes of participants, with this information a more exact IMD score can be obtained, by inserting the postcodes into a database on the website www.geoconvert.com.

In the study described in chapter four, samples from MSM participants were genotyped using the MLVA-*omp*A typing system. During this study there was evidence that obtaining genotypes from urine samples may be difficult due to inhibitors (e.g uric acid, calcium phospasphate, calcium oxalate) in the samples. A previous study conducted here in Southampton included only women as it was difficult to obtain genotypes from urine samples with the protocols in place in the laboratory at the time the study was conducted (Wang *et al.*, 2011b). However, to attempt to define sexual networks it was necessary to include both women and men in my study. Men often only give urine samples which meant that it was important to find a suitable DNA extraction method to extract chlamydial DNA from urine samples that could be typed. Suitable DNA extraction kits which could be used to extract DNA from urine samples were researched. This research led to a kit manufactured by Qiagen called the QIAamp Viral RNA kit (Qiagen, UK) (Section 2.11.3). Despite the name of the product it was recommended by the manufacturer for the extraction of DNA from bacterial samples. This kit proved to remove residual contaminants and the DNA produced by this method could be fully sequenced from both the urine and swab samples. Another advantage of using this kit was that it only required 40µl of sample material and did not require the production of a pellet for DNA extraction. This was beneficial as on many occasions after diagnostic tests had been carried out there was only a small volume of material left over, which would have proved difficult to extract DNA from using other kits.

Following assignment of MLVA-*omp*A genotypes to the samples there were three *omp*A genotypes that were of interest in this study and these included; *omp*A genotypes B3/IU-FQ279 as this genotype is mainly found in trachoma infected individuals, *omp*A genotype Ia, as this genotype has previously been detected predominantly in black individuals and *omp*A genotype L2b which is mainly detected in samples provided by HIV+ MSM.

Three female participants were infected with the *omp*A genotype B3/IU-FQ279 (Genbank Accession number FJ261925) which was typed in samples collected from the genital tract, this was surprising as *omp*A genotype B is usually found in patients who have trachoma. To ensure these results were correct, the *omp*A gene in these three samples were re-amplified and re-sequenced and again the sequence was identified to be 99% identical to B3/IU-FQ279. This led to further research into the origin of the strain. A study published in 2010 on repeated *C. trachomatis* genital infections in adolescent women found that seven participants were infected with genotype B3/IU-FQ279 (Batteiger *et al.*, 2010). This study is not in isolation and other studies have been published where samples with *omp*A genotype B chlamydia have been typed from an MSM population (Li *et al.*, 2011, Christerson *et al.*, 2012). The typing of samples with *omp*A genotype B from samples collected from the rectum and from samples collected from the eye may be an indication that genotype B may have evolved to infect both genital and ocular cells. This could also be supportive evidence that the *omp*A gene is in fact mobile (Harris *et al.*, 2012).

Genotype Ia has previously been associated with black race (Hackstadt, 1999, Geisler *et al.*, 2006, Marangoni *et al.*, 2012). Researchers who carried out a study in Birmingham, Alabama found that out of 207 specimens collected from both female and males there were 17 specimens which were typed as having an *omp*A genotype Ia, all of which were from black individuals. There has also been other publications were there have been significant differences seen in genotypes with relation to race of individuals who participated in the studies. In a paper published by Workowski *et al* (1999) the group made observations that serovar Ia was more prevalent in black individuals than in other races and that serovar D was actually less prevalent. This difference in serovar distribution was attributed to differences in biological or behavioural factors seen in different races. For instance, there may be racial differences in sexual behaviour which includes racial differences in frequency of sexual contact. Differences may also be seen in sexual behaviour, sexual practices, partner choice and number of partners and these differences could be the reason why there are genotype differences between races. The authors also suggested that it is more likely for people of black race to have sexual intercourse with one another and this could reflect the serovar associations within relatively closed populations. As this significant difference is only seen with serovar Ia and D in this particular study, it was concluded that these serovars

may have only recently been introduced into the population and therefore have not reached an equilibrium in the population. The host's initial susceptibility to infection and the immune response to *C. trachomatis* infection could also explain why there are differences in genotype distributions between groups. It was also observed that racially associated differences in clinics where individuals are tested may also influence the results (Hackstadt, 1999). In my study there were eight samples which contained a genotype Ia, all of which were from white British individuals. This either disproves the hypothesis that serovar Ia is associated with black race or this could be evidence to suggest that there is more inter-racial sexual intercourse taking place and therefore this genotype is only now being found amongst the white-British population. Geiser *et al* (2006) also mentioned in their publication that serovar Ia was associated with older age as a greater percentage over the age of 23 had the infection. However, in my study all of the individuals whose samples were typed as having an *omp*A genotype Ia were between the ages of 18-22 which disproves their hypothesis and actually suggests that *omp*A genotype Ia is more commonly seen in chlamydial infections of individuals who are young. To conclude, unlike previous studies, genotype Ia was found in individuals who are young and of white-British ethnicity in Southampton indicating a revision of type-association will soon be needed (subject to larger studies being conducted). This also suggests the linking of sexual networks with the spread of such types between ethnicities.

C. trachomatis L2b was typed in the urine samples of two male patients. Both participants visited the GUM clinic in November but on different dates, were both white British, one was 19 and the other was 28, both presented with symptoms and one lived in Southampton and the other in Salisbury. This was not the first recording of samples collected from the urethra that have been typed as having an *omp*A genotype L2b. In 2005, L2b urethritis was reported by a group in France who typed a urine sample from a male. In contrast to the two patients who were diagnosed as having *omp*A genotype L2b chlamydia in my study the male in France was reported to be a HIV+ man who has sex with other men (Herida *et al.*, 2005). In the study by Herida *et al* (2005) a male had reported having multiple partners with whom he had unprotected sexual intercourse. For my study the sexual partners of the two patients were not recorded. However, they had both been to the clinic on previous occasions and one had been diagnosed with *C. trachomatis* whilst the other had provided a sample which was negative for *C. trachomatis* which indicated that they had been worried about being infected. The two individuals were heterosexual men and this suggests that LGV strains have spread to the heterosexual population. More recently L2b has also been diagnosed in a heterosexual female but again this was from a rectal biopsy (Peuchant *et al.*, 2011). As L2b has previously been found in rectal samples of both men and women and now in the urine

samples of heterosexuals, there is now evidence to suggest that it is only a matter of time before the first case of L2b is diagnosed in a heterosexual female who has supplied a urine sample or a cervical swab for diagnosis.

It has been suggested that an ideal Simpsons index of discrimination for a typing system should be >0.95 (van Belkum *et al.*, 2007). For the samples in this study this requirement is not satisfied by using the markers individually to assign types to samples. Individually the markers have a Simpsons index with figures between 0.63 and 0.82. However, this requirement is satisfied when a combination of the three markers –CT1335, CT1299, CT1291 are used, this gives a Simpsons index of 0.97 and when the three VNTR markers plus *omp*A are used to assign types to samples a value of 0.98 was obtained. This shows high levels of diversity as a number close to one means the sample set is totally diverse, i.e. all of the samples in the sample set have unique genotypes. This reiterates the fact that the MLVA-*omp*A typing system does offer a higher level of discrimination than just using *omp*A alone. However the use of *omp*A in the typing scheme does increase the value of the diversity index. Interestingly, the Simpsons index of diversity for each of the markers (apart from marker CT1291 which had a Simpsons index of diversity which was 0.1 higher in the Pedersen study than in this study) was higher with the sample set in my study than the sample set in the study in which the MLVA-*omp*A typing system for *C. trachomatis* was designed (Pedersen *et al.*, 2008). The overall Simpsons index of diversity was also higher with this sample set by a value of 0.2 than a similar study in Southampton (Wang *et al.*, 2011b). This is possible due to the larger sample size in which a larger diversity could be seen and because the sample set in the earlier study was derived exclusively from women.

The ability to assign a type for each marker in each isolate was also assessed. The marker CT1291 had the lowest typeabilty capabilities with only 86% of the samples being assigned a CT1291 type, whereas the *omp*A gene was typed in 96% of the samples in this study. This shows that despite the fact that the *omp*A gene has low resolution when used as a typing system on its own, samples are easily assigned *omp*A genotypes which is one of the reasons why the typing of this gene is still used to type chlamydia.

The minimum spanning tree shows the relatedness of the MLVA-*omp*A types and was drawn for the 302 samples which gave full MLVA-*omp*A profiles (Figure 5.4). The MSpT showed that the MLVA-*omp*A types formed a large network representative of the samples in Southampton. Larger clonal complexes typically contain subgroups and therefore have both primary and subgroup founders, i.e the single locus variants (SLVs) of the primary founder may increase in frequency and diversify to generate SLVs and therefore becomes a subgroup founder. The clonal complex in

figure 5.4 was therefore divided into smaller sections to analyse the demographic data. The default rule in the eburst programme states that any sequence type with at least three links to other STs, including the link to the primary founder should be identified as a subgroup founder. However, in larger complexes such as the one presented here for Southampton, the definition of a subgroup can be user defined (Feil *et al.*, 2004). Here, the MSpT was firstly sectioned into two parts, section A consisted of samples that have the least frequent genotypes in the heterosexual population, whilst section B consisted of the most predominant *omp*A genotypes detected in the heterosexual population (figure 5.4b). Two other rules were defined and these were that a subgroup founder is any MLVA-*omp*A type which has evolved greater than six times from the original MLVA-*omp*A type, this split the MST into three segments (figure 5.4c). Thirdly, a subgroup founder was defined as any sequence type which had at least six SLVs which splits the network into 10 segments (figure 5.4d). These 10 segments of the network were then compared to identify if there were any differences between *omp*A type, age, ethnicity, gender, area of residence, symptomatology and the type of chlamydia. The analysis showed that there were no specific genotypes associated with specific demographic groups. However, there were specific *omp*A genotypes which were associated with groups of closely related MLVA-*omp*A types. This shows that although recombination within and between *omp*A genes has previously been documented (Harris *et al.*, 2012), recombination of the *omp*A gene may not be occurring frequently. Therefore, this gene can still be used to study short term local epidemiology as they are conducted over a short period of time, provided that it is used in combination with other markers as the *omp*A gene only provides a limited level of resolution for genotyping.

The date the sample was collected and the MLVA-*omp*A type of the chlamydia in the sample was also shown to be significant and therefore this typing scheme is useful for temporal studies of sexual networks. However, some MLVA-*omp*A types only contained either all male or all female participants who were heterosexual and had given samples with the same MLVA-*omp*A types during the same month. This shows that this information cannot be used solely to define sexual networks. To properly define sexual networks more information is required to draw conclusions on which individuals are within sexual networks and how these individuals are linked. This will require the collection of information on sexual partners. A previous study has defined couples based on individuals who have the same postcode, have an age gap of no more than 20 years and have different family names to avoid brothers and sisters (Bom *et al.*, 2011). However, these criteria has many flaws such as a postcode identifies a group of houses and not just one house, it is possible (but rare) to have a couple with more than a 20 year age gap and if the couple is identified because they have different family names then this will eliminate married couples.

Therefore to define groups of people who are connected sexually, information on partners of positive individuals is required. There are possible sexual networks in my study identified as there were several samples (e.g swabs) with the same MLVA-*omp*A type which were collected in the same period of time. However, it is important to acknowledge that there are individuals who are missing from analysis as they may not have attended the GUM clinic, they may have refused to give consent for their sample to be included in this study, their sample may not have been collected or did not give a full MLVA- *omp*A profile and therefore was not included in the MSpT. Collecting information on partners would be useful in combination with typing data to better understand the spread of chlamydia to study local epidemiology.

5.21 FUTURE WORK.

There were many learning curves in this project which can be changed and adapted for future projects. For instance the wording on the information sheet states that a sample will be collected from the patient for the study. This raised an important question as to what happens if the person returns for another visit during the study period, could their subsequent sample be used as they have consented once before? As this was a grey area, it was concluded that for subsequent studies the word samples should be used on the consent form as a pose to sample which would enable the use of any samples collected during the study period provided that the individual has given consent for their sample to be included once during the study. This would then include all samples subsequently collected from the patient during the study period but not those collected previously even if it is within the study period.

In future studies it would be worthwhile to record information on the partners of individuals to analyse if partners have the same MLVA-*omp*A type and if the types they have are in the same segments of the minimum spanning tree. This will require recording of the patient numbers of partners who are also positive to form links on the MSpT between these individuals. A computer-based self-interview may also be useful when collecting data to increase the accuracy of the data and to reduce human error when recording data. These data can then be sent directly to the research team and thereby cut down the time needed for the consultant to enter the data. The questionnaire can be completed whilst patients are waiting for their consultation. During the consultation the consultant can check that all the sections have been filled in correctly. This system works well in places like Amsterdam were it has been used to collect epidemiological data on *Neisseria gonorrhoeae* (Heymans *et al.*, 2012). In this study paper versions were also made available for those who would rather fill in paper versions. Another advantage of having the computer-based forms is that they can be adapted so that they can be translated into various languages and there can also be audio output, which is useful for those who cannot read.

Recording of other information such as sexual behaviour habits of patients, the use of condoms, anal intercourse, fisting, drug use, sexual preference, high risk behaviour, socio-economic status, education level may also be useful information when analysing sexual networks and studying epidemiology within a community. Future aims could also include identifying specific types associated with specific conditions (i.e. pelvic inflammatory disease, infertility and ectopic pregnancy)

5.22 CONCLUSIONS

The use of high resolution genotyping methods is important as there are various subtypes which can be found within *omp*A types. This can be seen from the various MLVA-*omp*A genotypes in the sample set for this study. Amongst the 302 fully genotyped samples there were 151 unique MLVA-*omp*A types. Each *omp*A type (apart from L2b which only had one isolate with a full MLVA-*omp*A profile) had several MLVA genotypes, this shows that the diversity of samples is greater than was once thought when only *omp*A typing was used to discriminate between samples.

To conclude, this study was successful in fully assigning genotypes to 302 samples from individuals who attended a GUM clinic in Southampton over a six month period. However as there were no specific genotypes associated with specific demographic groups, sexual networks within the community could not be identified. Southampton is unlike other major cities in the UK, where there are large ethnic variations within its population. The majority of the population in Southampton are in fact of white British ethnicity and this therefore adds some level of bias to this study. As can be seen from figure 5.8a, ethnicities other than the white-British ethnicity form only a small proportion of the population of Southampton. To adjust for this bias a survey in a city such as London or Manchester would be ideal as these cities have larger populations and a higher percentage of other ethnic groups and because of this the MSpT drawn from MLVA-*omp*A genotypes of samples in these cities may form several different clusters, similar to studies carried out in Amsterdam (Bom *et al.*, 2011, Bom *et al.*, 2013a) and Sweden (Christerson *et al.*, 2012).

CHAPTER 6 ANALYSIS OF CURRENT HIGH RESOLUTION TYPING METHODS AND THE DEVELOPMENT OF A NEW TYPING SCHEME FOR *C. TRACHOMATIS.*

6.1 INTRODUCTION

Since the early 1960s when isolation of *C. trachomatis* in cell culture became a routine procedure, serological typing has been carried out based on chlamydial surface antigens (Gordon *et al.*, 1969). As immunological reagents advanced in terms of quality and specificity, monoclonal antibodies were used for serotyping mainly directed to the MOMP (Pedersen *et al.*, 2009). However, this method requires expert staff, specialist equipment and is expensive as well as time consuming (Pedersen *et al.*, 2009). Another disadvantage of serotyping is that the current antibody panel is not able to detect evolving and emerging strains (Pedersen *et al.*, 2009).

Bacterial epidemiological typing utilises specific genotypic labels that can be used to determine the source and transmission routes of the spread of strains of bacteria. As well as being a useful tool for tracking strains it can be used by forensic scientists in medico-legal cases and to identify emerging pathogenic strains, including bioterrorism agents (van Belkum *et al.*, 2007). In 1993 Yang *et al.* showed that sequence determination of *omp*A (the gene that encodes the MOMP) also known as '*omp*A typing' followed by phylogenetic analysis offered higher resolution than serotyping (Yang *et al.*, 1993). This was later confirmed by other independent research groups (Quinn *et al.*, 1996, Morre *et al.*, 1998). Pedersen *et al* have shown that the Simpson's index of diversity (D) for *omp*A is sub-optimal at only 0.69 (Pedersen *et al.*, 2008). According to the guidelines by the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) a typing system should have a D of no less than 0.95 in order to be 'ideal' for studying epidemiology (van Belkum *et al.*, 2007). Therefore the requirement for a typing system, which offers higher resolution for chlamydia other than *omp*A genotyping, has been recognised for a long time, and it is only in the past six years that attempts have been made to develop a typing system which offers such accurate discrimination and that can be used for detailed local epidemiological studies and to define sexual networks. As well as the low discriminatory power offered by *omp*A typing there is evidence of recombination between and within *omp*A genes (Dean *et al.*, 2009, Harris *et al.*, 2012), which renders its use as a sole marker unreliable.

There are currently three genetic typing schemes which by using multiple markers offer a higher level of discrimination between *C. trachomatis* samples than using *omp*A alone. These typing schemes were designed by Klint *et al* (2007), Pedersen *et al* (2008) and Peuchant *et al* (2012).

However, there are still some regions of the genome that are left uncovered by the choice and location of the markers used in these typing schemes. This means that any variations that occur within the genome in other regions remain undetected. These variations may be significant epidemiologically and genetically.

The purpose of the work carried out in this chapter was to assess the suitability of the two typing schemes used in the previous studies described in this thesis and if possible to develop a genetic typing scheme with markers which are evenly distributed around the *C. trachomatis* genome. A new typing scheme for *C. trachomatis* must also offer high discriminatory power and it must have markers which can be used to study short term and perhaps long term epidemiology (the latter is dependent on the stability of the markers chosen over a period of time).

6.2 AIMS:

- To compare the location of the current markers used for high resolution genotyping in the *C. trachomatis* genome.
- To identify all of the SNPs present between the genomes from samples collected from women who attended a GUM clinic in Southampton in 2009.
- To identify any differences between genomes of samples with the same MLVA-*omp*A type. This could help to identify new markers potentially useful for a new typing scheme.
- To identify any differences between genomes of samples with the same MLVA but different *omp*A types.
- To compare whole genome sequence data to the MLVA-*omp*A and MST data of a set of samples from Southampton to determine how close these typing systems are to the resolution afforded by whole genome sequences of the samples in terms of strain discrimination.

6.3 RATIONALE FOR STUDY

Even though there are already typing schemes available that offer a high level of discrimination to type *C.* trachomatis (Klint *et al.*, 2007, Pedersen *et al.*, 2008, Peuchant *et al.*, 2012) the markers in these systems are not evenly distributed around the *C. trachomatis* genome. This can be seen in figure 6.1 which shows the location of the markers which offer high levels of discrimination for typing *C. trachomatis*. As can be seen in figure 6.1 there are large areas of the genome which do not have any markers.

At this point in the project it was considered that this would be a reasonable starting point to assess the current typing schemes for *C. trachomatis* using bioinformatics tools and to perhaps look for new and better markers to type *C. trachomatis*.

6.4 STUDY DESIGN.

To develop a new typing scheme that can also be used by other groups, the markers used in the typing scheme must be evaluated and validated. This is required so that various laboratories can communicate information between one another and so that they can also exchange outbreak information. Achieving standardisation between labs can be difficult and time consuming but is necessary in achieving the essential requirements to share information on types circulating within communities.

The first step in the process of designing new markers for genotyping is to have a database of diverse genomes. At the time this project started only twenty *C. trachomatis* genomes were available on the NCBI database. Therefore it was important to obtain more genomes of recently collected samples from a local population. The advantage of having a large number of whole genome sequences including those collected from a defined population and location is that these genomes provide a large diversity of strains, some with only small differences between them that can be compared. This is because it is possible that closely related clonal types may arise from a single founder type. Such differences between genomes could be exploited in new typing schemes. Consequently, clinical samples obtained from the Health Protection Agency (HPA) regional laboratory in Southampton and used for a study carried out in the Molecular Microbiology Department, SGH (Wang *et al.*, 2011b) were sent to be 'whole genome sequenced' at the Wellcome Trust Sanger Institute (WTSI). In the Wang *et al* (2011b) study 157 samples were collected, 93 of which were assigned full MLVA-*omp*A genotypes. Out of the 93 fully genotyped samples, 62 samples from that study were then sent to the WTSI to have their whole genomes sequenced. This added a large number of new *C. trachomatis* genomes to the existing panel which were available online at the time this project started.

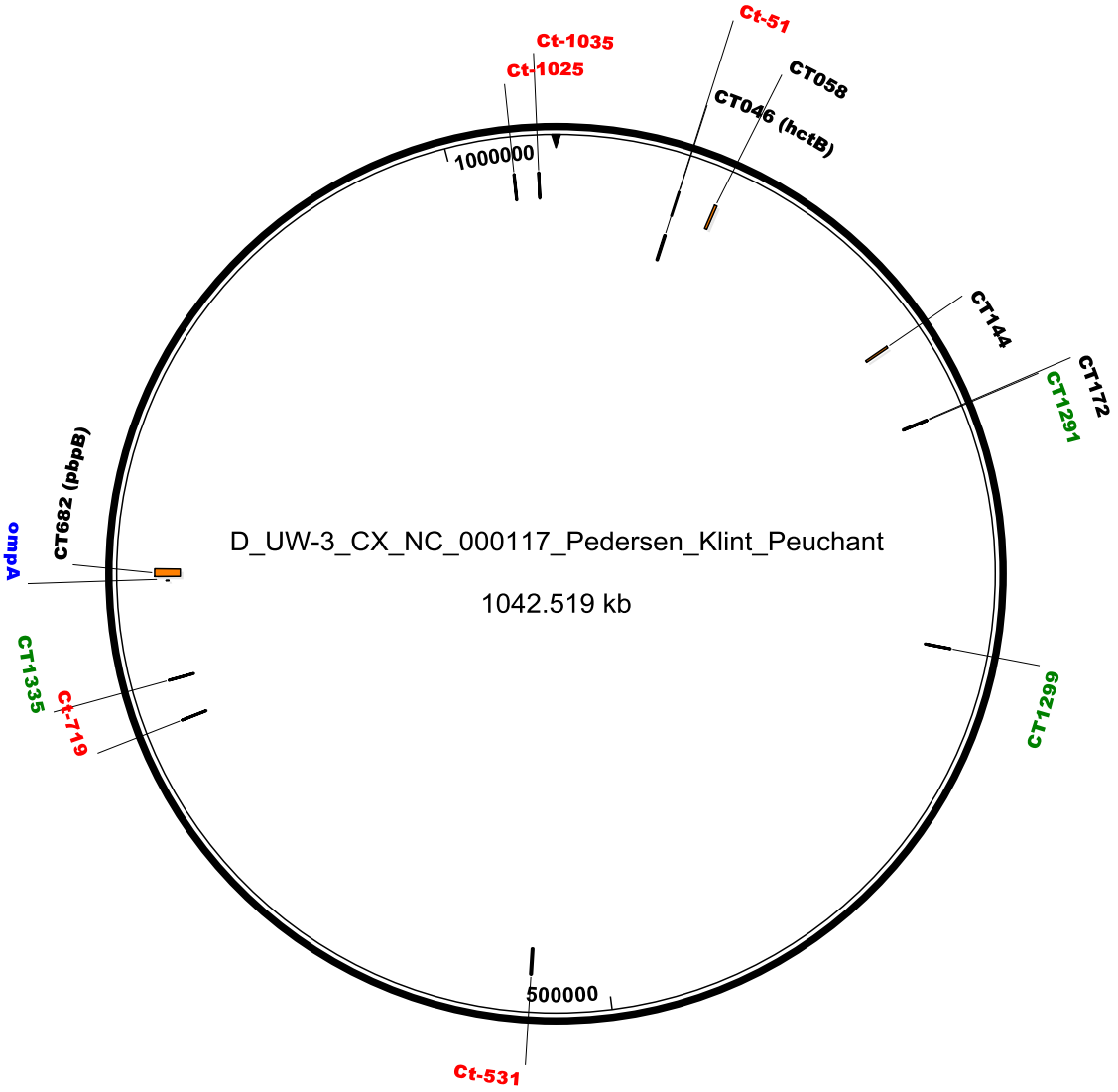


Figure 6.1 A graphical representation of the circular genome of *C. trachomatis* (reference strain D/UW-3/CX).
Indicated are the five regions described by Klint *et al* (2007) (markers in black), the three VNTR markers described by Pedersen *et al* (2008) (markers in green), the five markers described by Peuchant *et al* (2012) (markers in red) and the location of *omp*A (marker in blue). The diagram was drawn using seqbuilder (Section 2.15.2.2) by uploading the sequence of the reference strain – D/UW-3/CX into the programme. The primers for each marker were then added to the genome to locate the sequence of each of the markers in the *C .trachomatis* genome.

6.5 CONSTRUCTION OF A PHYLOGENETIC TREE FROM WHOLE GENOME SEQUENCING DATA OF 62 ISOLATES FROM SOUTHAMPTON.

6.5.1 Methods used to construct the phylogenetic tree.

A phylogenetic tree which included 62 samples from a study carried out in Southampton in 2009 (Wang *et al.*, 2011b) was constructed using whole genome sequence data generated for each of the samples. These 62 samples were chosen firstly because they were samples available in a collection held at the Molecular microbiology laboratory in Southampton, secondly because these samples were culturable and could therefore be sent for whole genome sequencing and thirdly because a full MLVA-*omp*A profile had already been determined for these 62 isolates. To generate the tree, sequence reads were firstly mapped against the complete assembled reference genome of SW4 using a script designed by Dr. Simon Harris at the WTSI called Multiple mappings to bam. SW4 is a chlamydia sample which was received from Dr. Kenneth Persson, Department of Clinical Microbiology, Malmö University Hospital, Malmö, Sweden and was genotyped as having an *omp*A genotype F. The following command was inserted into the putty terminal (section 3.4.4) to generate the tree:

```
~sh16/scripts/multiple_mappings_to_bam.py –r SW4_chrpl.fasta -g -t  *fastq.gz
```

This command instructed the script to map all of the 62 whole genome sequences (all of which were in a fastq.gz format and therefore *fastq.gz was used to inform the script to map anything which had that suffix to the whole genome sequence of SW4) which had not yet been assembled to the already assembled reference genome of SW4, using the –r option in the command line. The –g option was included in the command so that the script would create mapping plots. The –t option was added to the command so that a tabfile of the SNPs would also be produced. The script firstly identified all the homologous regions amongst the multiple genomes compared and then eliminated paralogous sequences to reduce false positive SNPs and finally multiple sequence alignments were generated and SNPs were identified between the genomes (Croucher *et al.*, 2011), i.e. homoplasic SNPs were eliminated. This was necessary so that the true phylogenetic relationship between the samples could be determined.

After the mapping process had been completed the phylogenetic tree was created by James Hadfield at the WTSI using the script Randomised A(x)ccelerated Maximum Likelihood (RAxML) to detect all the SNPs between the genomes and create the phylogenetic tree (Stamatakis, 2006). The following command was typed into the terminal to create the tree:

~sh16/scripts/run_RAxML.py –a SW4_chrpl_smalt.aln -o SW4_concatenated.tree

This command informed the script to put all the samples in the alignment file which was created from mapping the whole genome sequences of the 62 samples to the assembled genome of SW4. In this instance the alignment file (-a option in the command) was called SW4_chrpl_smalt. The output file (-o option in the command) was called SW4_concatenated.

The phylogenetic tree was then adjusted for recombination events to avoid false relationships being inferred. This was achieved using Gubbins which is a bioinformatics script used to detect recombination events within the genome (Croucher *et al.*, 2011). The script takes a multifasta alignment file, finds all of the SNP sites and then eliminates recombination and rebuilds the tree. Regions of recombination were identified as loci with a high density of polymorphisms. The regions of putative recombination events were then treated as missing data in all taxa downstream of the branch in which the recombination events occurred when redrawing the phylogeny with RAxML (Croucher *et al.*, 2011). The following command was used to remove recombination and redraw the tree:

~sh16/scripts/run_gubbins_fasttree.pl –t SW4_cocancatenated.tree –o new_SW4_concatenated

This command informed the script to remove any regions of recombination and to rebuild the tree without the regions present. To provide a measure of support for the relationships between the strains, 100 bootstrap replicates were calculated on the final tree (Harris *et al.*, 2012). The diversity between the samples can be seen from the position of the samples on the phylogenetic tree (figure 6.2).

6.5.2 Analysing the phylogenetic tree

On the phylogenetic tree the 62 samples were split into two clades –T1 and T2 (figures 6.2 and 6.3), these clades have been described previously (Harris *et al.*, 2012). The T1 clade consists of *omp*A genotypes D, E and F whilst the T2 clade consists of *omp*A genotypes D, G, I, J and K. The T1 clade was formed from the *omp*A genotypes which are the most prevalent in the heterosexual population whilst the T2 clade was formed from the *omp*A genotypes which are least prevalent in the community but perhaps more prevalent in the homosexual population. Notably, in both clades the individual *omp*A genotypes do not cluster together and *omp*A genotype D is found in both clades.

To compare the MLVA, the MST and *omp*A types to the whole genome data, a script designed by Dr. Simon Harris called reportlabtest.py was used. Only the markers in MLVA-*omp*A (Pedersen *et*

al., 2008) and the MST (Klint *et al.*, 2007) typing schemes were compared to whole genome sequence data as these two schemes offered the highest level of discrimination for *C. trachomatis*, were designed to study local epidemiology and the markers in these typing schemes had been assessed for their stability in chapter 3. The script added a different colour for each type as blocks next to the sample name to which that type corresponded to. The following command was inserted into the putty terminal to produce the figures:-

~sh16/scripts/reportlabtest.py –t MLVA_final.tre –M –L left –a 2 –m soton_MLVA.csv –c 2 –C 8,3,4,5,6,8 –O portrait

This command informed the script to use the tree file MLVA_final.tre (-t) and add the various loci as outlined in the metadata file – Soton_MLVA.csv (-m) as different colours for each type beside each sample name on the phylogenetic tree (for the comparison of MST data to the whole genome sequence tree a different data file was used – Soton_MST.csv- which included all the different types for each marker in the MST typing scheme). The –M option was included to inform the script that this was a midpoint root tree, the –L option was used to inform the script to ladderise the tree to the left of the sample names. The –a option was used to inform the script to align the taxon names to the right of the page. The –c option informed the script of the column in the metadata file to use for track names and the –C option was included to inform the script of the columns in the metadata file to use to colour the track names and block next to the name. The –O option was used to inform the script of the preferred orientation of the page. In this instance the orientation was portrait.

The comparisons can be seen to the right of the phylogenetic tree (figure 6.2B and figure 6.3B) the VNTR and MST types for each sample is shown using different colours for each VNTR and MST type. The corresponding type for each colour is shown in the key below the diagram. As well as showing the comparison for each of the individual VNTRs and markers in the MST scheme, the three VNTR and the five MST regions together which make up the MLVA and MST typing schemes is shown in the fourth column for the MLVA comparison (figure 6.2B) and the sixth column for the MST comparison (figure 6.3B) located to the right hand side of the phylogenetic tree. The *omp*A type is also shown in both figures using various colours to distinguish between each type.

6.5.3 Comparison of the whole genome sequence data to the *omp*A types.

The samples with the same *omp*A type were located on the tree and analysed, this analysis showed that samples with the same *omp*A type did not always form a cluster in the same region of the phylogenetic tree (Figure 6.2A). This showed that *omp*A typing does not show the true phylogenetic relationship between samples. All of the *omp*A genotypes occurred in multiple

branches on the tree suggesting that there has been transfer of at least part if not all of the gene (Harris *et al.*, 2012). Interestingly, *omp*A genotype D was found in both clades. There were no other *omp*A genotypes found in both clades of the tree. In this collection of samples it was already known from the initial *omp*A genotyping carried out before the samples had been whole genome sequenced that there were two types of *omp*A genotype D, these were genotypes identical to the following isolates on the NCBI website: D/IC-CAL8 and D/UW-3/CX. This analysis showed that apart from Soton D3 (Highlighted on figure 6.2) which had an identical *omp*A genotype to D/UW-3/CX all the other *omp*A genotype Ds in the T1 clade were *omp*A genotype D/IC-CAL8 which all clustered separately from Soton D3 in one region of the T1 clade. This was interesting because all of the isolates with the *omp*A genotype D in the T1 clade had the same MLVA type -8.5.2. In the T2 clade there were eight samples with the *omp*A genotype D which were all clustered together and were all identical to the reference genome D/UW-3/CX. However, unlike the T1 clade, the Ds in the T2 clade are more diverse and this may be evidence that the T1 clade is a more clonal clade than T2.

6.5.4 Comparison of whole genome sequencing data to the markers in the Pedersen MLVA-*omp*A typing system.

To assess the suitability of the MLVA-*omp*A typing system a comparison of the whole genome sequence data and the MLVA-*omp*A typing system was conducted (figure 6.2). Each VNTR type was assessed individually for its suitability as a marker in comparison to the whole genome sequence data. The T1 Clade consisted mainly of CT1335 type eight all but one sample which was CT1335 type six. However, for the T2 clade most samples were CT1335 type three, but CT1335 type three was interrupted by other CT1335 types in the T2 clade. The various types for the VNTR CT1299 are dispersed throughout the samples in both clades and also the samples are not clustered together according to VNTR type. For the VNTR CT1291 clade T1 is mainly made of two CT1291 types whereas clade T2 is made up of several different CT1291 types. This reflects a higher diversity in clade T2.

6.5.5 Comparison of whole genome sequencing data to the markers in the Klint MST typing system.

To assess the suitability of the Klint MST typing system the markers used for this typing scheme were compared to the whole genome sequence tree. Unlike the analysis conducted to compare the MLVA-*omp*A markers to the whole genome sequencing data where all the MLVA types had already been determined in a previous study by Wang *et al* (Wang *et al.*, 2011b) to assess the suitability of the MST markers, the types were assigned using the whole genome sequence data. The whole genome sequence for each of the samples was opened using Artemis (Section 2.15.3). For the genes, *hct*B and *pbp*B these were easily identified as these genes were already annotated

on to the genome. However for the hypothetical genes CT058, CT144 and CT172, *in silico* PCR was performed to locate these regions. Once all five regions were located, the sequence for each type were inputted in to the Uppsala MLST database (http://mlstdb.bmc.uu.se/cgibin/curr/mlstdbnet/mlstdbnet.pl?file=mlstdb_profiles.xml&page=oneseq) to assign types for each marker in each sample. For sequences not already on the database, these samples were assigned arbitrary numbers that had not already been used for that particular marker. This information was then inputted in to a file and the comparison was conducted using the reportlatest.py script as described in section 6.5.2.

The comparison showed that some of the markers in the MST scheme offer very high discriminatory power to the point where almost every isolate has a different type. This was confirmed by calculating the Simpson index of diversity for the five MST markers. CT046 and CT682 have the highest level of discrimination amongst the five MST markers with a D of 0.955 and 0.900 respectively. The markers for the MST typing scheme also do not show the true relationship between the samples when compared to the whole genome sequence data. Again types for the samples in the T1 and T2 clades were not clustered in one region of the tree which is what one would expect to see from a typing scheme which offers a good comparison between the markers in that scheme and whole genome sequencing data (figure 6.3).

With this information it became apparent that an improved typing system may be designed that would include markers that are evenly distributed across the genome and also a typing scheme that, when placed next to whole genome sequencing information, would produce a tree that offered similar results as the results of a tree produced from whole genome sequencing. To develop a typing system to study local epidemiology, SNPs or tandem repeats in the genome needed to be identified that offer a high discriminatory power between samples.

6.6 ANALYSIS OF GENOME DATA TO INVESTIGATE WHETHER IT WAS POSSIBLE TO DESIGN A NEW TYPING SYSTEM FOR *C. TRACHOMATIS*.

6.6.1 Identifying potential SNPs which may be useful in a SNP typing scheme for *C. trachomatis.*

Following the whole genome sequencing of the 62 samples, the tab file generated from mapping the genomes to SW4 included all the SNPs present in the genomes compared to SW4. SNPs were called if over 75% of the genome reads had the same SNP. If between 25% to 75% of reads had the same SNP then the SNP was marked as missing data (Croucher *et al.*, 2011). 10, 431 SNPs were identified between the 62 genomes. As there was a large number of SNPs between the

genomes it was difficult to establish which SNPs were useful to develop a SNP typing system. Therefore other approaches to develop a new typing scheme were investigated.

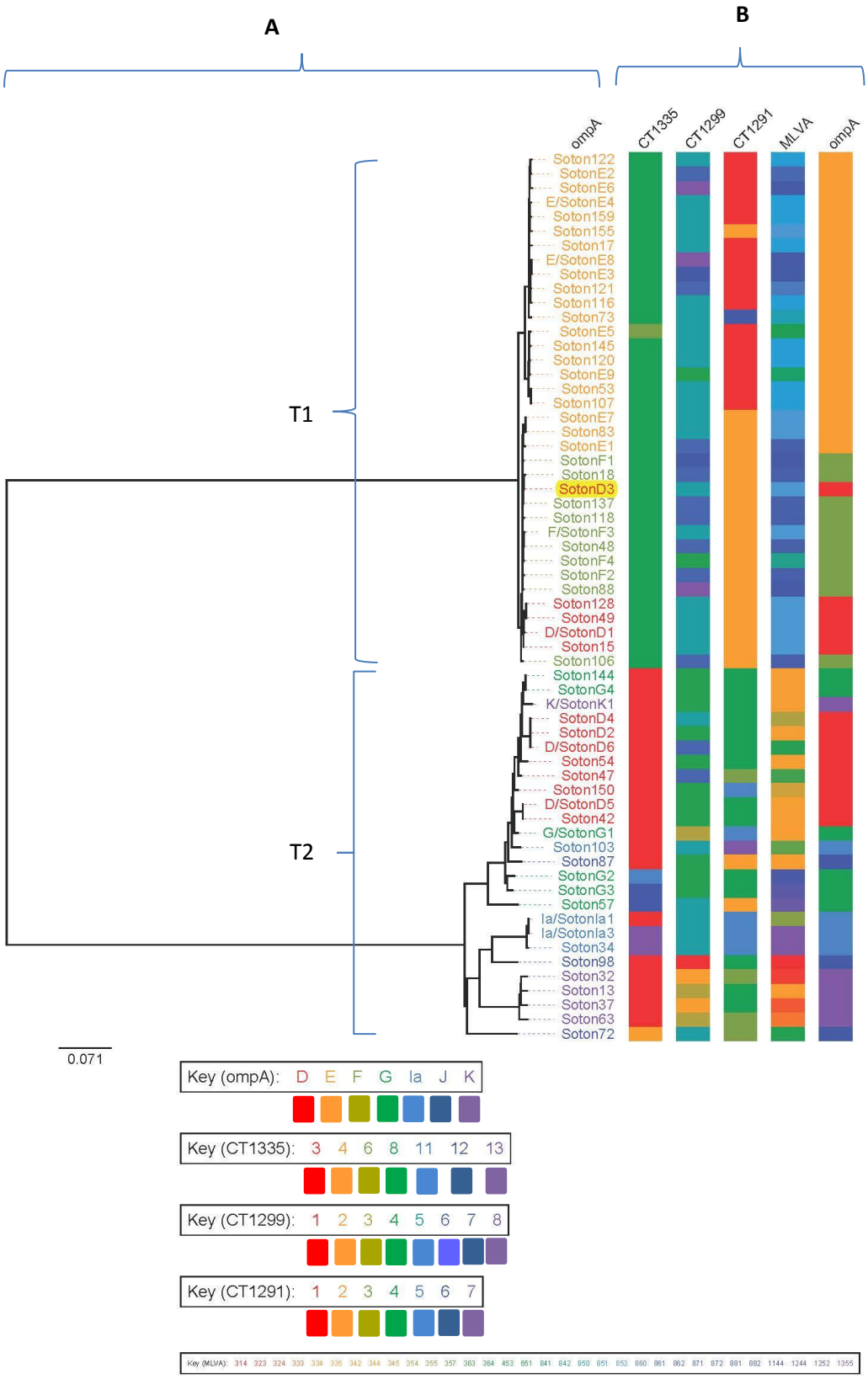


Figure 6.2 Comparison of MLVA-*omp*A genotyping to whole genome sequencing (recombination removed).

The tree shows the relationship between 62 samples from the Wang *et al* study (2011b) -A. The two clades are labelled T1 and T2. Strain names are coloured by *omp*A genotype. Branch lengths represent substitutions per site (see scale bar). The scale bar represents 71,000 SNPs. The various MLVA and *omp*A types for each sample are represented by different colours to the right of the phylogenetic tree- B. The types corresponding to each colour are shown in the key. Soton D3 is highlighted to show its location compared to the other *omp*A genotype Ds in the T1 clade.

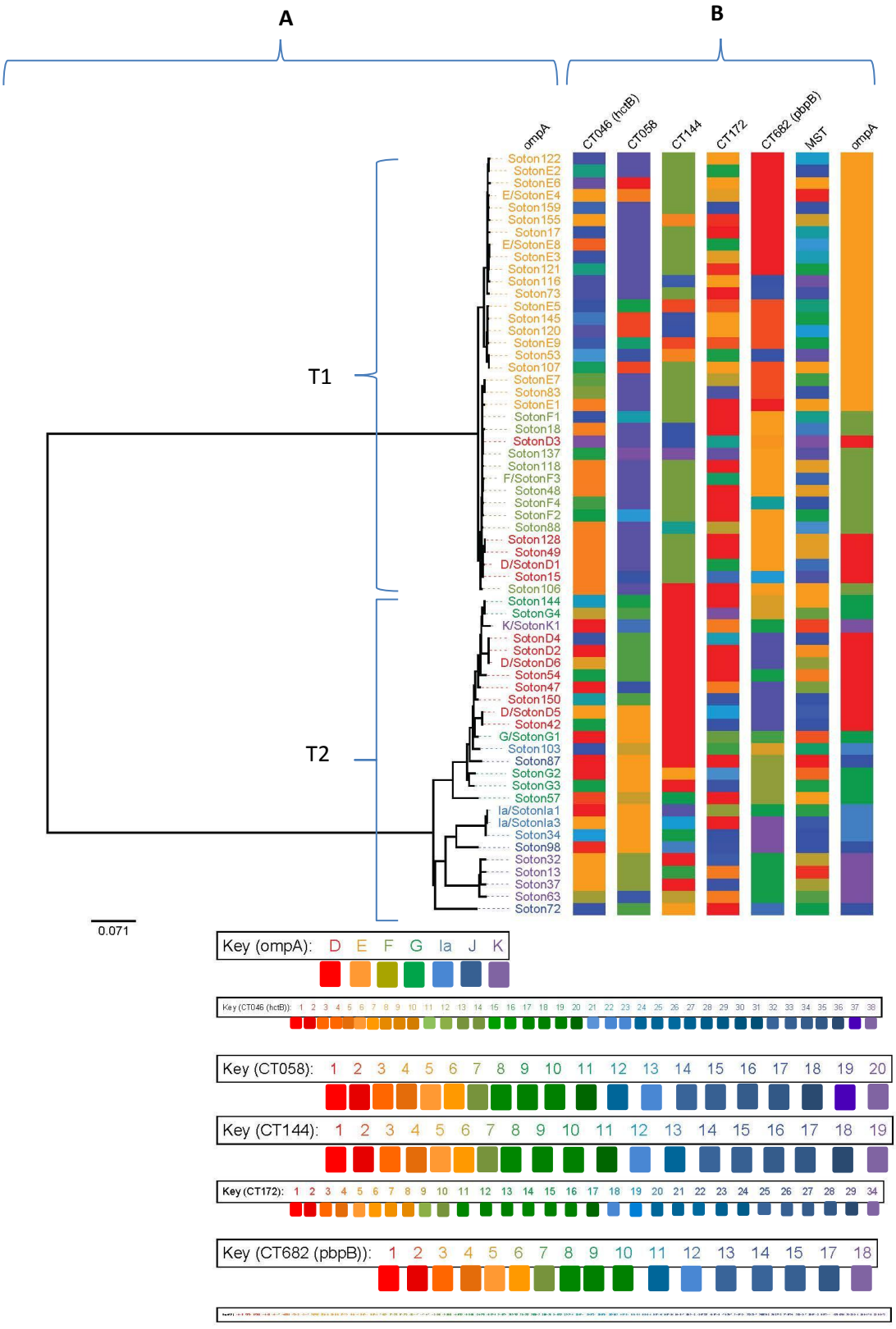


Figure 6.3 Comparison of MST and *ompA* genotyping to whole genome sequencing (recombination removed).
The tree shows the relationship between 62 samples from the Wang *et al* study (2011b) -A. The two clades are labelled T1 and T2 on the diagram. Strain names are coloured by *ompA* genotype. Branch lengths represent substitutions per site (see scale bar). The scale bar represents 71,000 SNPs. The various MST and *ompA* types for each sample are represented by different colours to the right of the phylogenetic tree- B. The types corresponding to each colour are shown in the key.

6.6.2 Identifying potential regions within samples with the same MLVA-*omp*A type which could be used in a new typing scheme for *C. trachomatis*.

Samples with the same MLVA-*omp*A types were identified from a collection of samples in Southampton. These samples were from the study described by Wang *et al* (Wang *et al.*, 2011b) and had already been whole genome sequenced, therefore the data generated from whole genome sequencing was readily available for analysis. Following the identification of these samples, another idea was developed to further investigate the whole genome sequences of the samples that had the same MLVA-*omp*A type according to the Pedersen scheme (Pedersen *et al.*, 2008) and to determine if there were any differences between these genomes and then establish whether these differences could possibly form the basis of a typing scheme where the current MLVA-*omp*A typing scheme could be further sub-divided if there where a few SNPs which are found between each sample with the same MLVA-*omp*A type. If new markers are found then this could add to the current typing scheme- providing more markers around the genome and offering a high discriminatory power to type samples.

There were three MLVA-*omp*A types which had more than one sample with the same MLVA-*omp*A type and these formed three groups; three samples had the MLVA type 8.5.2 and had *omp*A genotype D, these samples formed group one, the second group had eight samples with *omp*A type E and MLVA type 8.5.1, whilst the third group comprised of samples with *omp*A genotype F and the MLVA type 8.6.2. The sequences relating to each one of these MLVA types are shown in table 6.1. In each group there was one sample which had a fully assembled whole genome sequence. This was advantageous as all the other whole genome sequences of samples in the group did not need to be assembled and were mapped against the sequence in that group that had already been assembled. The sequences were assembled by collaborators at the WTSI - Dr Helena Seth-Smith and Dr Simon Harris. The mapping was carried out using the following command in a script designed by Dr Simon Harris, WTSI:

```
~sh16/scripts/multiple_mappings_to_bam.py –r D/SotonD1.fasta -g -t  *fastq.gz
```

This command informed the script to map the fastq files of the samples which had the same MLVA-*omp*A type to the previously assembled whole genome sequence D/SotonD1. This was subsequently repeated for the samples with MLVA-*omp*A type 8.5.1-E and MLVA-*omp*A type 8.6.2-F using the assembled genomes E/SotonE4 and SotonF2 respectively.

Interestingly, all of the samples within the three groups were only found in the T1 clade. The isolates with the same MLVA-*omp*A type did not always cluster together in same region of the tree except the three isolates that were typed as having a MLVA-*omp*A type 8.5.2-D.

6.6.2.1 Identifying SNPs within samples with the sample MLVA-*omp*A genotypes.

Firstly, SNPs within the genomes of the samples with the same MLVA-*omp*A type were identified. The script-Multiple mappings to bam produced a file called a variant call format (VCF file) which is a format currently in use at the WTSI which describes various types of sequence variations such as SNPs and indels relative to the assembled genomes in each of the MLVA groups. The number of SNPs between the genomes with the same MLVA-*omp*A type is shown in table 6.1. This table shows the total number of SNPs within each MLVA-*omp*A group when the genomes in each group were compared to one another. There were over ten times the number of SNPs in isolates with the MLVA-*omp*A type 8.6.2 –F compared to the isolates with MLVA-*omp*A type 8.5.2 –D. Additionally, even though there were more isolates with the MLVA-*omp*A type 8.5.1-E than with the MLVA-*omp*A type 8.6.2 –F there was over four times the amount of SNPs in isolates with the MLVA-*omp*A type 8.6.2 –F compared to the isolates with MLVA-*omp*A type 8.5.1-E. This shows that there is a lot of variation between samples with the MLVA-*omp*A genotype 8.6.2-F.

Figure 6.4 shows the relationship between each isolate in groups two and three (8.5.1-E and 8.6.2-F). Group one only had three isolates and therefore a tree was not drawn for the samples which had the MLVA type 8.5.2 with *omp*A genotype D. As can be observed from the trees in figure 6.3 there are several SNPs between the samples with the same MLVA-*omp*A types therefore these samples, although, when typed using the MLVA-*omp*A typing scheme are identical, actually possess a large number of differences elsewhere in the genome.

6.6.2.2 Identifying other useful tandem repeats within the genomes of the samples with the same MLVA-*omp*A type.

As it became evident that the samples with the same MLVA-*omp*A type had differences elsewhere in the genome the next aim was to identify other tandem repeats in the genome which would give a better representation of these samples than the MLVA scheme.

Tandem repeats within three different genomes (D/IC-CAL3, E/Bour, and F-IC-CAL-3) were identified using the program mreps 2.5 (http://bioweb.pasteur.fr/seqanal/interfaces/mreps.html). The genomes were searched for 'maximal repeats'(which is a repeat that cannot be extended on either side of the repeat to make a longer repeat) using the standard settings in the programme with minor modifications with repeats of ten or more nucleotides and output small repeats that can occur randomnly (-allowsmall) as described by Pedersen *et al* (Pedersen *et al.*, 2008). The software produced a list of all tandem repeats found within the genomes according to the parameters set. The intention was to compare the repeats found using this software with the genomes of the samples with the same MLVA-*omp*A types. The sequences were analysed bioinformatically using Artemis (Section 2.15.3) to determine if there were any other useful

tandem repeats in other regions of the genome that could be used to create a typing scheme. Unfortunately there were no other useful tandem repeats found in these genomes which offered a high discriminatory power for *C. trachomatis* between samples with the same MLVA-*omp*A types. This was determined by comparing the tandem repeats identified using the mreps program between the genomes with the same MLVA-*omp*A types. The tandem repeats were identical in the genomes and therefore offered no discrimination between the genomes.

	Sequence in the three VNTR types							
Group Number	**CT1335**	**CT1299**	**CT1291**	**MLVA type**	***omp*A type**	**Number of samples with this type**	**Southampton samples with the MLVA-*omp*A type**	**SNPs between genomes**
1	13T7A	11C	8C	8.5.2	D/IC-CAL8	3	D/SotonD1 Soton128 Soton49	54
2	13T7A	11C	6C	8.5.1	E/Bour	8	E/SotonE4 Soton159 Soton116 Soton17 Soton145 Soton120 Soton107 Soton53	1362
3	13T7A	12C	8C	8.6.2	F/IC-CAL3	5	Soton/F2 Soton137 Soton48 Soton118 Soton106	5719

Table 6.1 Southampton *C. trachomatis* isolates from women (Wang *et al.*, 2011b) who had more than one sample with the same MLVA-*omp*A type. Also shown is the number of SNPs between the samples that had the same MLVA-*omp*A type and the names of the Southampton isolates with the same MLVA-*omp*A types.

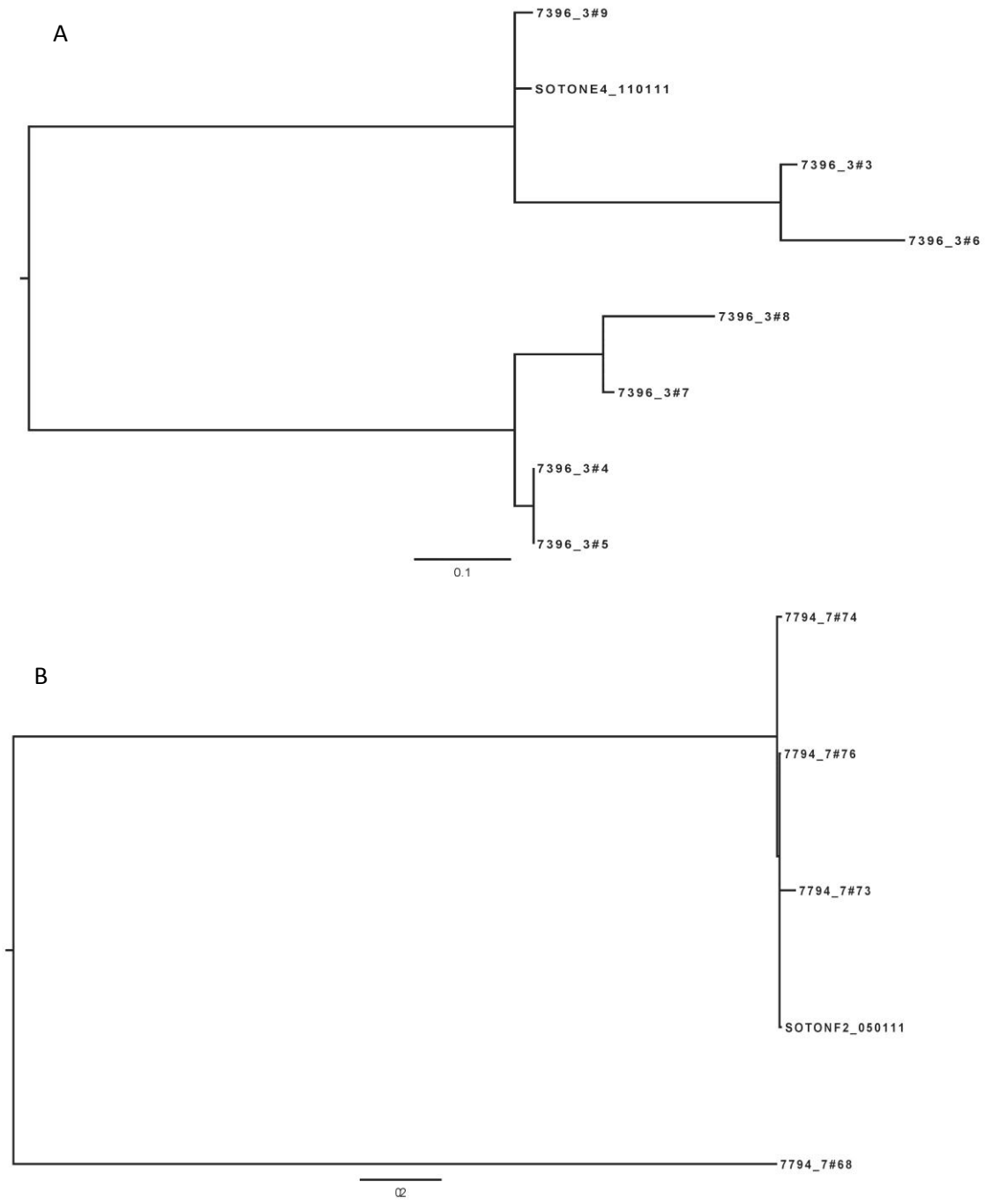


Figure 6.4 Phylogenetic tree depicting the differences dictated by the number of SNPS between samples with the same MLVA-*omp*A types.

The trees were created from whole genome sequence data from samples collected from patients in Southampton who had samples with the MLVA-*omp*A sequence type 8.5.1 –E/Bour (a) five Southampton samples with the MLVA-*omp*A sequence - 8.6.2 –F/IC-CAL3 (b). Branch lengths represent substitutions per site (see scale bar). The scale bar in (A) represents 100,000 SNPs whilst in (B) the scale bar represents 200,000 SNPs.

6.7 SAMPLES WITH THE SAME MLVA BUT DIFFERENT *OMP*A TYPES

As well as the set of samples with the same MLVA-*omp*A type there was also a set of samples which had the same MLVA but different *omp*A types, this group of *C. trachomatis* isolates were interesting because there has been recent speculation that recombination is widespread and a driving force for chlamydial diversity (Harris *et al.*, 2012). If only the *omp*A gene was different and the backbone of the genome remained the same then this could be evidence of recombination of *omp*A genes between these genomes (Harris *et al.*, 2012). There were three MLVA groups identified amongst the 62 Southampton genomes (MLVA types 8.5.2, 8.6.2 and 3.4.4) and within these groups were different *omp*A types. For instance the genome sequences of the samples with MLVA-type 8.5.2 included samples with *omp*A genotypes D, E and F. These samples were not clustered in the same region of the phylogenetic tree (Figures 6.2 and 6.3). This was also true for samples with types 8.6.2 and 3.4.4.

6.8 DISCUSSION

Typing methods are used to study the spread of bacteria and other microorganisms. They are also useful when determining how a certain type of bacteria evolves and is spread between individuals (van Belkum *et al.*, 2007).

As there is a lot of variation which can be identified between chlamydial genomes there is also scope for developing new typing schemes for *C. trachomatis.* Additionally, there are large sections of the genome which remain without markers and these areas are potentially suitable targets for developing a new typing scheme. With the vast amount of whole genome sequencing data available for *C. trachomatis* it is now timely to evaluate this new data. However, the design of new typing schemes involves a full redesign of the available genomes.

A phylogenetic tree of the 62 available isolates from Southampton was constructed. The tree was adjusted for recombination events because over longer durations of time, recombination may prevent the true relationships between distantly related isolates of species from being recognised (Croucher *et al.*, 2011). There were over 10,000 SNP differences between the 62 genomes which can be seen from the phylogenetic tree (Figure 6.2). The 62 genital tract isolates form two clades even though there are no phenotypic differences between the T1 and T2 clades. In 2006 Miyairi and colleagues showed that there are no differences in growth rates between genital tract isolates. As the T1 and T2 clades shown in figure 6.2 and 6.3 are all genital tract isolates which have been shown to grow at an intermediate rate (ocular strains are slow growing and LGV strain are often fast growing) it is unlikely that there are any differences in growth rates in the isolates which make up these two clades (Miyairi *et al.*, 2006). Notably the isolates that make up the T1

clade are the isolates which are more predominant in prevalence studies of *C. trachomatis* whilst the isolates that form the T2 clade are the rare *omp*A genotypes found in prevalence studies of heterosexuals but are more prevalent in the rectum of MSMs (Suchland *et al.*, 2003, Wang *et al.*, 2011b, Christerson *et al.*, 2012, Bom *et al.*, 2013a). This split of the phylogenetic tree into two clades may therefore be indicative of tissue tropism of the isolates. As there were thousands of SNPs between the isolates it was not possible to establish which of these SNPs would be useful in a SNP typing system for *C. trachomatis*.

Comparative studies carried out in this chapter have shown that the MLVA-*omp*A and the MST typing systems do not show the true phylogenetic relationship between samples and neither does assigning genotypes to the samples using the *omp*A gene on its own. There was no clear relationship seen when the three MLVA-*omp*A and the five MST markers were used as a typing system when compared to the whole genome sequence data. This shows the need for a new typing system which mirrors the true phylogenetic relationship between *C. trachomatis* samples. In fact the MST scheme may actually offer a higher discrimination than is required for epidemiological studies as with some of the markers almost every sample had a different type. This level of discrimination does not allow for the tracking of genotypes within the community because with such high levels of discrimination every sample becomes unique. Additionally, mapping of the various typing markers on to the circular genome of *C. trachomatis* has also shown that none of the current typing schemes include markers that are evenly distributed around the *C. trachomatis* genome (Figure 6.1). As there may be genetic differences in these areas of the genome which may have been missed by the location of the markers in the current typing scheme, there may be a requirement for a typing scheme which includes markers in these areas of the genome.

From the study by Wang *et al* (2011b) there were some samples which, when typed using the MLVA-*omp*A typing scheme, were identical. However, even though they had the same MLVA-*omp*A type it was possible they may have differences outside the markers. Thus, the reason for obtaining whole genome sequences for the set of samples with identical MLVA-*omp*A types was to determine if there were any variations elsewhere within the genome. If there were variations this could be key in identifying how similar these genomes are and might lead to possible further subdividision of these isolates and allow the design of a typing system that offers an even higher level of discrimination between samples than that offered by the current MLVA-*omp*A typing scheme (Pedersen *et al.*, 2008). If it were the case that no variations were found within the rest of the genome between these samples then it can be concluded that the MLVA-*omp*A typing scheme is suitable for typing and there would be no need for the development of a new typing scheme.

This would also mean that the whole genome sequencing of samples would not be necessary as a simple typing scheme would be adequate to differentiate samples and indeed group them together when trying to define and track sexual networks on a community based level. However, the work carried out in this chapter has shown that there are differences between samples with the same MLVA-*omp*A type. As can be seen from figure 6.4, even though these isolates have the same MLVA-*omp*A type, there are other differences which can be observed in other areas of the whole genome of the samples. These differences are in the form of SNPs and indels. These differences can be used to design a new typing scheme for *C. trachomatis*. If there were more samples with the same MLVA-*omp*A type this would have been advantageous when trying to determine which SNPs or markers would be useful in a typing scheme. However, as there were not many samples in each group (for instance there were only three isolates with the MLVA-*omp*A type 8.5.2), this was not possible to complete and could act as future work to collect more samples with the same MLVA-*omp*A type to determine which changes are most important. The need for large numbers of genomes to identify useful regions to develop a typing scheme was highlighted during the development of the MST scheme by Klint *et al* (2007). Initially, when the markers were analysed in 16 reference genomes the five MST markers were shown to offer lower discrimination than assigning genotypes to samples using the *omp*A gene. It was only when these regions were analysed in 47 clinical isolates that the regions were shown to have higher discriminatory power than typing with *omp*A (Klint *et al.*, 2007).This highlights the need to have several genomes to develop a new typing scheme. The lack of concordance between the MLVA-*omp*A typing system and whole genome sequence data suggests different approaches may be required to develop a new typing scheme for *C. trachomatis*.

The data collected here also confirmed findings by another group (Harris *et al.*, 2012) and showed that even within the genomes of samples that have the same MLVA but different *omp*A types there are several differences elsewhere in genome. Therefore it can be concluded that there are recombination events occurring within and between *omp*A genes and there are also other SNPs and possible recombination events occurring elsewhere in the *C. trachomatis* genome (Harris *et al.*, 2012).

6.9 FUTURE WORK.

Future work for this study includes identifying lineages consisting of groups of samples by analysing the samples on the phylogenetic tree. SNPs can then be identified that distinguish between these lineages to develop a new typing scheme which reflects the true phylogeny of the samples. Once the typing scheme has been developed new PCR primers can be designed to test the new markers on a set of local samples. Additionally, to find SNPs that differentiate the

lineages, new MLST or VNTR markers which reflect the positioning of the isolates on the whole genome sequence tree can be used in a new typing scheme. These can be used to set up a library system so that the typing method can be used in different laboratories and can be inputted into a database which various groups can have access to, to compare data and add new data to as time progresses.

Once potential new markers have been identified as suitable candidates for a typing scheme they should be tested on a set of samples to ensure that a high percentage of samples can be typed when using the newly developed typing scheme. There are several criteria that must be met by the markers in a typing scheme which include the following; the markers must be stable, have a high discriminatory power, offer a high level of typeability, the typing scheme must offer epidemiological concordance, and the data generated by the typing scheme must be reproducible. The markers must also be assessed for their stability to ensure that the sequence of the markers do not change rapidly. Stability should be measured by culturing a set of isolates and the sequence of the markers should be determined before and after passage. The isolates tested for stability should have been stored over different periods and under different conditions and processed in the same run to avoid variations which can be introduced during preparation in the laboratory (van Belkum *et al.*, 2007). Although the stability of the markers will be better assessed in an *in vivo* situation this is not always possible and needs to be conducted under controlled conditions which is difficult to do with *C. trachomatis.* The markers should also offer a high discriminatory power with a Simpsons index of 0.95 or above to be ideal for use in a typing system. The typing system must also reflect and agree with available epidemiological information. In addition, the location of the markers and the primers used to amplify the sequence of the markers should also be made readily available to other groups to test on their local samples to analyse if the typing results are reproducible (van Belkum *et al.*, 2007).

CHAPTER 7 FINAL DISCUSSION

An understanding of bacterial pathogen population structure is essential to monitor infections and for the development of public health strategies and medical treatments to counter the spread of infections. Despite the fact that *C. trachomatis* exists as a set of genotypes, some with specific tropisms for the genital tract, little is known about the molecular epidemiology of *C. trachomatis* in the UK. To address important questions about chlamydial infections current technologies that offer the level of resolution required to make meaningful analyses is an important requirement. *Omp*A genotyping has served extremely well and is still used by many groups to type chlamydia strains. The typing of this gene splits *C. trachomatis* into at least 15 genotypes. However, *omp*A genotyping does not offer the high level of discriminatory power that is needed to study epidemiology. Several studies conducted to investigate the genotypes circulating within a community have shown that over 40% of samples tend to have an *omp*A genotype E (Pantoja *et al.*, 2001, Pedersen *et al.*, 2008, Wang *et al.*, 2011b, Marangoni *et al.*, 2012). There is also some evidence that *omp*A is mobile (Dean *et al.*, 2009, Harris *et al.*, 2012). It has also been shown that recombination events within the genome occur (Harris *et al.*, 2012). However, for short term local epidemiological studies such as those carried out for this thesis, recombination events are unlikely to occur during the short duration of the studies (Wang and Grayston, 1963). This is why it is acceptable to use typing systems to type chlamydia and there is no true need for whole genome sequencing when answering epidemiological questions. As over short time scales recombination events are unlikely to have occurred therefore clones and clonal complexes can be recognised (Wang and Grayston, 1963). Additionally, it is acceptable to use *omp*A genotyping for short term epidemiological studies and figures 4.4 and 5.5 show that samples with the same *omp*A types often cluster closely together. Nevertheless, the discriminatory power as defined by the Simpsons index of diversity (Hunter and Gaston, 1988)which is a measure of the diversity of the samples in a study, is low for studies where *omp*A genotyping has been analysed (Bom *et al.*, 2011, Christerson *et al.*, 2011). Therefore, *omp*A genotyping alone cannot be used for determining transmission patterns in sexual networks or in cases of sexual abuse or assault. Much finer discriminatory typing systems are obviously needed if they are to be used for such purposes. To date there are five typing methods designed to type *C. trachomatis* but not all of these systems offer a higher level of discriminatory power than that offered by the *omp*A gene. These typing systems include two MLST schemes (Pannekoek *et al.*, 2008, Dean *et al.*, 2009), an MST scheme (Klint *et al.*, 2007) and two MLVA-*omp*A schemes (Pedersen *et al.*, 2008, Peuchant *et al.*, 2012). The two MLST schemes offer very low discriminatory power and therefore were unsuitable for the

project requirements. However, the other three typing schemes offer a higher level of discriminatory power and a higher level of intra-serotype variation than typing only *omp*A (Klint *et al.*, 2007, Pedersen *et al.*, 2008, Peuchant *et al.*, 2012). Therefore, it was decided that either the MST scheme or the MLVA-*omp*A typing scheme would be better suited for this project (as the scheme by Peuchant *et al* (2012) had not yet been published at the time of the study). These typing systems were analysed and assessed for their suitability and stability (chapter three). Following this analysis it was decided that the MLVA-*omp*A typing system would be adequate for the project requirements. Additionally, this typing system had already been assessed in a previous study in Southampton which showed that this typing scheme was suitable for typing samples from Southampton. The use of a typing scheme that has been evaluated previously to type samples from the city (Wang *et al.*, 2011b) allows for the ability to compare the development of the bacteria over time as we can monitor the different types over the years and study the epidemiology of chlamydia in the city as time progresses.

Even though the level of discrimination offered by *omp*A typing alone is low and yields little epidemiological information, the current high resolution typing systems for *C. trachomatis* still incorporate the *omp*A gene into the typing methods. This shows that despite the fact that the gene may be mobile and offers low resolution, inclusion of the gene into typing schemes which include other markers increases the overall resolution of the typing methods (Sections 4.7 and 5.10). The use of the *omp*A gene has also been a good way of grouping samples. *Omp*A genotypes A, B and C are generally associated with trachoma, *omp*A genotypes D-K are associated with genital tract infections, and *omp*A genotypes L1, L2 and L3 are found in patients with lymphogranuloma venereum. However, L2b (commonly found in rectal samples from HIV+ MSM) and genotype B (commonly found in trachoma cases) were found in genital tract infections in the heterosexual community (chapter five). This supports findings that *omp*A is mobile and perhaps that chlamydia has developed mechanisms where different *omp*A types can infect a range of different cells than they could previously. The ability for the various *omp*A types to infect various cells is an advantage for the bacteria as it allows the bacteria to spread at a faster rate and a disadvantage for the host who become infected with this bacterium.

This thesis describes how typing systems can be used to study epidemiology of genital tract *C. trachomatis*. The results of the stability study (chapter 3) led to the decision to use the MLVA-*omp*A typing scheme by Pedersen *et al* (2008) to study the epidemiology of *C. trachomatis* in two cities; Brighton and Southampton, as the markers in this typing scheme are stable upon adaptation to cell culture during both short and long term passaging. However, whole genome sequence analysis has shown that there may be changes elsewhere in the genome following long

term passage (chapter 3). Nevertheless, these changes are not located within the regions of the markers used in the typing scheme rendering this typing scheme suitable for typing chlamydia (chapter 3). The markers in the MST scheme are also stable following short and long term passaging but this typing scheme was not chosen for the epidemiological studies in this thesis on the basis that the markers in these typing schemes are over 500bp and therefore PCR of these regions requires nested PCR which increases the yield and specificity of amplification of the target DNA but often leads to contamination of the PCR product.

One of the main advantages of using a multi-locus typing scheme to assign types to samples is that it gives faster results than can be achieved by obtaining whole genome sequencing. Whole genome sequencing requires a lot of preparation before sequencing ensues. Once sequencing is complete knowledge of bioinformatics is required for the assembly and alignment of the genomes. One of the initial requirements of whole genome sequencing is that the DNA needs to be amplified in cell culture which requires specialist laboratory equipment and a highly skilled individual to carry out the culture. A lot of time is also required to expand chlamydia in cell culture, to obtain the concentration of DNA required for whole genome sequencing. A new method is currently under development to obtain whole genome sequences directly from clinical samples (Seth-Smith *et al.*, 2013). This method combines immunomagnetic separation (IMS) and multiple displacement amplification (MDA) (Seth-Smith *et al.*, 2013). However, depending on the volume of the initial clinical sample that this approach requires this may still prove a difficult method as some samples may not have a large enough volume of transport media in to which the swab is placed. Additionally, due to the number of repeat procedures that sometimes have to be carried out on samples (if the sample is found to produce an equivocal result) or if a repeat of the diagnostic test is requested from the doctor or patient then the volume of liquid in the tube may be inadequate for obtaining whole genome sequences directly from the swab. From the study carried out by Seth-Smith *et al* (2013) only five complete genomes were generated from 18 samples that were taken through the IMS-MDA protocol. To get a whole genome sequence directly from sequencing this will depend on the type of sample, patient-to-patient variation, sample collection and storage after collection (Seth-Smith *et al.*, 2013). The paper published by this group also suggests that the IMS-MDA protocol for *C. trachomatis* is only appropriate for swabs taken from transport medium which does not lyse the bacteria but which maintain intact EBs (Seth-Smith *et al.*, 2013) and therefore this approach would not have worked for this project as the swabs collected from the GUM clinics were placed in lysis buffer. As many clinics and GP surgeries now use swabs which are placed into a lysis buffer, this particular IMS-MDA protocol for obtaining whole genome sequences from *C. trachomatis* may not be useful for future studies,

unless prospective additional swab samples which use suitable transport media are specifically employed.

The MLVA-*omp*A typing scheme was used to assign types to MSM in Brighton and to heterosexuals in Southampton. This allowed for the comparison of two cohorts in two different cities, located in the South of England. For both studies assigning types to the samples was successful following the design of a new set of primers than the primers described in the papers by Pedersen *et al* (2008) and by Wang *et al* (2011b) study. The use of both sets of primers meant that a higher percentage of samples were assigned full MLVA-*omp*A types which is ideal as a larger set of the types of chlamydia circulating within the community can be observed.

Following the assignment of types to samples, minimum spanning trees (MSpT) were generated to visualise the large amounts of typing data. As there were several MLVA-*omp*A types found in the studies described in chapters four and five rules were defined to section the MSpTs so that samples with similar MLVA-*omp*A types would be grouped to determine if there were any demographic differences between the MLVA-*omp*A groups. By generating the MSpT the primary founder of the clonal complex could easily be determined. The simplest model for emergence of clonal complexes is that the primary founder of a clonal complex is the genotype that has increased in frequency in the population due to random genetic drift or natural selection to become the predominant clone (Wang and Grayston, 1963, Wang *et al.*, 1985). Genotypes diversify from one another by occurrence of mutations that result in the formation of variants from the primary founder (Wang and Grayston, 1970). Primary founders of clonal complexes were identified for studies in chapters four and five based on this definition. It is important to mention that MSpTs do not provide information about pathways of evolutionary descent, rather they provide an understanding of recent evolutionary events (Wang and Grayston, 1970), therefore, proving suitable for studies in this thesis. It became evident that possible sexual networks could be identified but without information on partner identification it was not possible to draw links between individuals who have had sexual relations with each other. Partner notification is not compulsory in England but is important for the treatment of chlamydia. If sex partners are not treated then the infection will continue to be passed on between people in the network. In the MSpTs (figures 4.4 and figures 5.5) most *omp*A genotypes clustered together apart from *omp*A genotype D. *Omp*A genotype D was found dispersed throughout the MSpTs (figure 4.4 and figure 5.5). Interestingly it was also the only genotype found in both clades –T1 and T2 of the whole genome sequence tree (figure 6.2), suggesting that there may be something unique in chlamydia with *omp*A genotype D and this *omp*A genotype should be investigated further to determine why this is the case.

For the Brighton project (Chapter 4) the most prevalent *omp*A genotypes in MSM were G, E and D. One of the aims of the project was to determine if there were differences in the MLVA-*omp*A types of chlamydia found in HIV+ and HIV- men. Analysis of two clusters showed that there was a statistically significant difference between the MLVA-*omp*A types found in HIV+ and HIV- MSM. There was no statistically significant difference between the MLVA-*omp*A type and previously being infected with chlamydia or other STIs. However, individuals with *omp*A genotypes L2 and L2b were more likely to present with symptoms, indicating the difference between LGV and genital tract strains. Interestingly, there was no concordance in genotypes found in the urethra and rectum in men infected at both sites, this may be due to infection from different partners, or the same partner who has been infected by someone else and suggests that it is likely that high risk sexual behaviour may exist within MSM.

Culture of *C .trachomatis* is challenging due to it being an intracellular bacteria. However, despite the obstacles faced with isolating chlamydia from rectal samples, 37 rectal isolates were obtained. DNA from these isolates were sent to the Sanger Institute where the whole genome sequences of the samples will be determined. Using this data whole genome sequencing data from an MSM cohort can be compared to those from a heterosexual population to determine if there are any differences between these genomes which may help to explain the differences in tissue tropisms of various types.

In Southampton the most common *omp*A types were E, D and F amongst heterosexuals (Chapter 5). This study highlighted the need to collect data on partners so that sexual networks can be determined and confirmed using this information along with typing and demographic data. We initially hypothesised that sexual networks could be defined using the typing data in combination with demographic information provided by the participants, but this was not possible, due to lack or partner identification data. Additionally, we hypothesised that there may be certain MLVA-*omp*A types found in certain demographic groups such as age, postcode and ethnicity. However, this was not the case and the data suggested that sexual intercourse occurs between individuals from various demographic groups. There were also no differences in types found in asymptomatic and symptomatic participants.

Additionally, by comparing the circulation of selected strains from MSM and heterosexuals within a relatively closed population, the typing data may also explain why certain genotypes are more predominant in rectal isolates from the MSM population than in cervical or urethral isolates from heterosexuals (Hackstadt, 1999). A minimum spanning tree was generated (Appendix 10) to investigate whether chlamydia MLVA-*omp*A types differ between MSM in Brighton (chapter 4)

and the heterosexual population in Southampton (chapter 5). Interestingly, the most common MLVA-*omp*A type found in both cohorts was 8.5.1-E this type may represent a bridging population between heterosexuals and MSM. The types from both studies formed a single network with no clear clusters belonging solely to MSM or to heterosexuals. This was unlike studies by Christerson *et al* (2012) and Bom *et al* (2013b) where comparison of *C. trachomatis* strains from MSM and heterosexual populations were conducted using MST and the two groups formed separate clusters. There were some MLVA-*omp*A types found solely amongst MSM in Brighton and MLVA-*omp*A types found solely in the heterosexual population in Southampton. However, as there were fewer samples collected from MSM than heterosexuals in Southampton it was not possible to be certain that these types where indeed only associated with one group or the other. A better study design would perhaps be to compare types found in a city with an equal number of samples from heterosexuals and MSMs and to compare the types found in the two cohorts. This comparison shows that the individuals infected with chlamydia in the South of England are unlike those found in Amsterdam and Sweden where the genotypes collected from MSM and heterosexuals form separate clusters. Additionally, this difference could be due to the differences in the markers used in the two typing schemes. There may be differences in the regions used for assigning MST types that are unique to MSM and some unique to the heterosexual population. These differences may occur due to pressures exerted by host factors on these regions which are not exerted on the regions used in the MLVA-*omp*A typing scheme.

The MLVA-*omp*A and MST typing systems were also compared to whole genome sequence data from a study in 2009 (Wang *et al.*, 2011b) (Chapter 6). Neither the MST nor the MLVA-*omp*A typing schemes where comparable in terms of resolution to the whole genome sequencing data, therefore, the development of a new typing scheme is required. Due to the large amount of data currently available from whole genome sequencing of samples from Southampton, Brighton and those available in the public domain there is potential to develop new typing schemes with new markers that span the whole of the genome as the current markers are located in areas of the genome that leave large areas of the genome without markers. Development of such a typing scheme would require high speed computers for the assembly and alignment of the genomes and to analyse the large amounts of data. A new and better typing scheme can be developed once the level of discrimination that is required for epidemiological studies is determined. Ideally the new typing scheme should offer a high discrimination between samples but not too high that all samples are unique as the epidemiology of chlamydia cannot be studied if every sample has a unique type.

Conducting this project showed that epidemiological studies require a lot of planning and hard work to set up and to carry out the studies. There are also ethical considerations which need to be carefully thought through before commencing with these studies. It is also difficult keeping everyone (i.e. routine clinical staff, nurses etc who are involved in the research additional to their main tasks) involved in the studies motivated for the duration of the study as enthusiasm for the study declines as time progresses. Even though it is important that enough samples are collected over the study period to meet the aims of the project, it is essential to keep the study period to a minimum to avoid 'research fatigue' which can develop from carrying out the same project for an extended period of time. Nevertheless, we were able to successfully make use of a typing scheme following assessment of its suitability for the studies in the project. A large number of rectal isolates were obtained and will be used for future studies. The possibility to use the vast amounts of whole genome sequencing data to develop new typing schemes also adds to the future work that can be developed from the work in my thesis.

APPENDIX 1: MARKER STABILITY PAPER

frontiers in
CELLULAR AND INFECTION MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 17 May 2012
doi: 10.3389/fcimb.2012.00068

Genotyping markers used for multi locus VNTR analysis with *ompA* (MLVA-*ompA*) and multi sequence typing (MST) retain stability in *Chlamydia trachomatis*

Clare Labiran[1], Ian N. Clarke[1], Lesley T. Cutcliffe[1], Yibing Wang[1], Rachel J. Skilton[1], Kenneth Persson[2], Carina Bjartling[3], Björn Herrmann[4], Linus Christerson[4] and Peter Marsh[5]*

[1] Molecular Microbiology and Infection, School of Medicine, University of Southampton, Southampton, Hampshire, UK
[2] Department of Clinical Microbiology, Malmö University Hospital, Malmö, Sweden
[3] Department of Obstetrics and Gynaecology, Malmö University Hospital, Malmö, Sweden
[4] Department of Medical Sciences, Section of Clinical Bacteriology, Uppsala University, Uppsala, Sweden
[5] Health Protection Agency, Microbiology Services, Public Health Laboratory Southampton, Southampton General Hospital, Southampton, Hampshire, UK

Edited by:
Yasuko Rikihisa, Ohio State University, USA

Reviewed by:
Ted Hackstadt, Rocky Mountain Laboratories/NIAID/NIH, USA
Yajun Song, Beijing Institute of Microbiology and Epidemiology, China
João P. Gomes, National Institute of Health, Portugal

***Correspondence:**
Peter Marsh, Health Protection Agency, Microbiology Services, Public Health Laboratory Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, Hampshire, UK.
e-mail: peter.marsh@uhs.nhs.uk

We aimed to evaluate the stability of the *Chlamydia trachomatis* multi locus VNTR analysis (MLVA-*ompA*) and multi sequence typing (MST) systems through multiple passages in tissue culture. Firstly, we analyzed the stability of these markers through adaptation of *C. trachomatis* to tissue culture and secondly, we examined the stability of a four-locus MLVA-*ompA* and a five-locus MST system after multiple passages in tissue culture. Marker sequences were monitored through successive chlamydial developmental cycles to evaluate the stability of the individual DNA markers through many bacterial divisions and this, in turn, informed us of the usefulness of using such typing systems for short and long-term molecular epidemiology. Southampton genitourinary medicine (GUM) clinic isolates from endocervical swabs collected from *C. trachomatis* positive women were passaged through tissue culture. MLVA-*ompA* typing was applied to primary swab samples and to the same samples after *C. trachomatis* had been passaged through cell culture (eight passages). Sequence data from time-zero and passage-eight isolates were aligned with reference sequences to determine the stability of the markers. The Swedish new variant (nvCT) underwent 72 passages in cell culture and the markers of the two schemes were similarly analyzed. Analysis of genetic markers of the MLVA-*ompA* typing system before and after the isolates were introduced to tissue culture showed no change in the dominant sequence. The nvCT that had been passaged 72 times over the duration of a year also showed no variation in the dominant sequence for both the genotyping schemes. MLVA-*ompA* and MST markers are stable upon adaptation of *C. trachomatis* to tissue culture following isolation of strains from primary endocervical swab samples. These markers remain stable throughout multiple rounds of cell-division in tissue culture, concomitant with the incubation period and appearance of symptoms normally associated with host-infection. Both genotyping schemes are, therefore, suitable for epidemiology of *C. trachomatis*.

Keywords: *Chlamydia trachomatis*, MLVA-*ompA*, multi sequence typing, marker stability, tissue culture

INTRODUCTION

Chlamydia trachomatis, an intracellular bacterial pathogen is the leading cause of sexually transmitted infections (STIs) worldwide. The majority of diagnoses are in young people between the ages of 15 and 24. *C. trachomatis* is currently divided into at least 15 serotypes. Genital infections are caused by the non-invasive serotypes D to K and by the invasive lymphogranuloma venereum (LGV) which infects the lymphatic system. If left untreated, which is the case for up to 50% of men and 70% of women due to the asymptomatic nature of this STI, a range of complications may persist, although opportunistic screening and partner notification detect a lot of asymptomatic cases. Individuals infected with LGV can progress to develop proctitis, lymphadenopathy, and genital ulcerations, whereas disease outcome for individuals infected with serotypes D to K range from infertility, ectopic pregnancy, and pelvic inflammatory disease in women and epididymitis in men.

The availability of whole genome sequences for *C. trachomatis* has led to the development of several prototype genotyping systems. Multilocus sequence typing (MLST) is one of these typing techniques and it is usually based on 6–8 housekeeping genes which are specifically chosen because they are not under immune selection, and because of their role in cell survival. Therefore, MLST is useful when exploring long-term and

global epidemiology and two such systems have been developed for *Chlamydiacae* (Pannekoek et al., 2008; Dean et al., 2009). To achieve a higher discriminating capacity a system based on five non-housekeeping genes was created (Klint et al., 2007). It is aimed for partner tracing and other short-term epidemiology. This scheme was originally named as MLST, but the term "MLST" strictly refers to the exclusive sequencing of small amplicons from non-selectable housekeeping genes. In this paper the designation of this scheme is reassigned to "multi sequence typing" (MST) as it encapsulates large and small amplicons, and includes selectable non-housekeeping genes (Wang et al., 2011). Pedersen et al. (2008) developed a high resolution typing scheme which encompassed multi locus variable number tandem repeat (VNTR) analysis and *omp*A (MLVA-*omp*A) (Pedersen et al., 2008, 2009; Wang et al., 2011). VNTR are defined as short regions of nucleotide repeats or motifs. There are several MLVA schemes practiced in bacteriology with various end-point measurement methods including fluorescent amplicon size-discrimination and sequence analysis of the VNTR loci (Pedersen et al., 2008; Schouls et al., 2009). For *C. trachomatis* Pedersen et al. (2008) described three VNTR regions; CT1335, CT1299, CT1291. Both the MST system of Klint et al. (2007) and the Pedersen et al. (2008) MLVA-*omp*A system provide a Simpson index of diversity between 0.94 and 0.96 as opposed to 0.64–0.83 when only *omp*A is typed (Ikryannikova et al., 2010; Bom et al., 2011; Wang et al., 2011). The guidelines outlined by the European Society of Clinical Microbiology and Infectious Disease (ESCMID) states that a typing system should have a discriminatory power of no less than 0.95 in order to be "ideal" for studying epidemiology (van Belkum et al., 2007). The MLVA DNA markers were selected for typing because of their variability therefore, it is also particularly important to assess the stability of these markers when the study involves sample collection by isolation of *C. trachomatis* in continuous cell culture. In the study by Wang et al. (2011) a sub panel of samples could not be typed initially and were subsequently assigned a MLVA-*omp*A type after culture (Wang et al., 2011).

Pedersen et al. (2008) have tested the stability of the three markers and *omp*A *in vivo* in 24 patients with persistent or recurrent infection. Their results show that the markers are stable between 70 and 394 days. However, this is very difficult to control as these patients were assessed under the assumption that it was the same isolate the individual was infected with and also that each of the 24 patients had not contracted chlamydia from a different sex partner and could have, therefore, been re-infected.

The aim of our study was to evaluate the stability of the *C. trachomatis* typing systems. The MLVA-*omp*A system was assessed after multiple passages in the controlled conditions of cell culture. By following potential changes through many developmental cycles it was possible to establish the stability of the individual markers within a numerical framework of bacterial divisions and this in turn informed us of the usefulness of using such a typing system for short-term and longer-term molecular epidemiology. In this study we analyzed firstly, the stability of these markers through adaptation of *C. trachomatis* to cell culture, in doing so the initial genetic impact on the *C. trachomatis* genome resulting from changing environmental conditions was analyzed. Secondly, the stability of the VNTR markers with *omp*A as well as the MST system (Klint et al., 2007) after multiple passages in cell culture was evaluated.

METHODS

CELLS AND CULTURE OF *C. trachomatis*

McCoy cells were seeded into wells or flasks 24 h pre-infection. Cells were grown overnight in Dulbecco's Modified Eagle medium (DMEM) containing 10% foetal calf serum (FCS) and then incubated at 37°C with 5% CO_2. On the day of infection the medium was removed and replaced with inoculum containing the *C. trachomatis* isolates. This was then centrifuged at 754×g for 30 min. Subsequently the inoculum was replaced with DMEM containing cycloheximide (1 μg/mL), gentamicin (20 μg/mL), and vancomycin (10 μg/mL), and the cells incubated at 37°C for 48 h. After 48 h the inclusions were large enough and visible under a light microscope so that the extent of infection could also be determined visually to decide whether a larger surface area was required for the next passage.

HARVESTING OF *C. trachomatis* ISOLATES FROM WELLS AND FLASKS

Two days after infection tissue culture flasks were harvested, for this a cell scraper was used to detach the cells from the plastic of the flask. The sample was then centrifuged at 2851×g for 10 min, the supernatant was then discarded and the pellet resuspended in 1:10 cold phosphate buffered saline (PBS). The sample was then added to glass beads and agitated for 1 min to release the elementary bodies from the cell, centrifuged at 110×g for 5 min to remove any cell debris and the supernatant was then added to an equal volume of 4 Sucrose Phosphate (4SP) and stored at −80°C. To harvest from wells the cells were scraped up into the medium using a sterilized pipette tip, the 10 min centrifugation step was omitted as was the addition of PBS, instead the harvest was immediately beaded and then the protocol for harvesting was continued as for harvesting from flasks.

Chlamydia trachomatis ISOLATES

In 2009, 157 endocervical swabs were collected over a six month period from women who had attended GUM clinics and GP surgeries in Southampton (Wang et al., 2011). The samples tested positive for *C. trachomatis* using a commercial real time PCR (Cobas Taqman 48, Roche), performed in the Southampton Health Protection Agency (HPA) Molecular Diagnostics Laboratory. To test the stability of *omp*A and the VNTR markers (CT1335, CT1299, and CT1291), seven of the archived isolates were chosen (**Table 1**). Locations of these loci within the chlamydial genome are shown in **Table 2**. The isolates were selected based on their *omp*A type, so that all available *omp*A types could be included in this study. The isolates were initially typed using the MLVA-*omp*A typing system i.e., before inoculation of the primary sample into tissue culture. All seven isolates were then taken through eight serial passages in cell culture and were subsequently re-typed.

The stability of the Swedish new variant in cell culture was also determined, an isolate from Malmo in Sweden was sent to Southampton (**Table 1**), where it was passaged 72 times in cell culture over the period of a year. The sample was obtained from the urethra of a male patient who had contracted the new variant

Table 1 | Isolates included in all studies and their sites of isolation, also included are the repeat sequences for the three VNTR loci, their associated variant number, and the corresponding *ompA* NCBI database strain type for each isolate before culture.

ompA type (Determined by comparison to NCBI database sequences)	Source	Repeat sequence for each VNTR loci before culture			Variant number
		CT1335	CT1299	CT1291	
D/IC-CAL8	Cervix	13T7A	11C	8C	8, 5, 2
D/UW-3	Cervix	13T7A	11C	8C	8, 5, 2
E/Bour	Cervix	13T7A	11C	8C	8, 5, 2
F/IC-CAL3	Cervix	13T7A	12C	8C	8, 6, 2
G/392	Cervix	10T8A	10CT3C	10C	3, 4b[a], 4
Ia/870	Cervix	10T8A	11C	11C	3, 5, 5
K/UW-31	Cervix	10T8A	3C2T6C	10C	3, 9[b], 4
E/Bour (SW2)	Urethra	13T7A	14C	6C	8, 8, 1

[a] *VNTR region modified in the Wang et al. paper.*
[b] *New VNTR region found in Wang et al. (2011).*

Table 2 | Location and putative function of the MLVA-*ompA* and MST markers within the genome, also included is the strand (negative or positive) in which the MST and *ompA* markers are located.

Marker (Scheme)	Location and function
CT1335 (MLVA-*ompA*)	Located at −44 to −66 of DNA topoisomerase I/SWI domain fusion protein, Hypothetical protein
CT1299 (MLVA-*ompA*)	Located in a non-coding region of the genome, Hypothetical protein
CT1291 (MLVA-*ompA*)	Located within the gene for the hypothetical protein CT172.1
ompA (MLVA-*ompA*)	Negative strand (−780,237: −778,702), Outer membrane protein A
CT058 (MST)	Negative strand (−68,693: −67,260), Hypothetical protein
CT144 (MST)	Positive strand (160,568: 161,678), Hypothetical protein
CT172 (MST)	Negative strand (−195,655: −195,016), Hypothetical protein
hctB (MST)	Positive strand (51,410: 52,205), DNA binding protein
pbpB (MST)	Positive strand (780,178: 784,392), Penicillin binding protein B

The exact position (nucleotide number) within that strand is shown in brackets.

infection in Sweden in 2006. The Swedish new variant is a strain which has a 377 bp deletion in the cryptic plasmid and 44 bp duplication immediately upstream of CDS3, and is, therefore, a mutable strain and most suitable for studying marker stability (Ripa and Nilsson, 2006, 2007).

DNA EXTRACTION

DNA was extracted from isolates that had been expanded to a level of infection were there was enough to infect 100% of the cells in a T75 tissue culture flask. DNA extraction was performed using the Promega Wizard genomic DNA purification kit (Promega, Southampton, UK) and was carried out according to manufacturer's instructions. The presence of DNA in the sample was verified on an agarose DNA gel. The concentration of DNA present in each isolate was determined using the Nanodrop 2000 (Thermo Scientific, Delaware, USA). PCR was conducted with high fidelity DNA polymerase, Phusion high fidelity DNA polymerase mastermix (New England Biolabs UK, Hitchin, UK), to ensure that the sequence data obtained was of the best quality possible.

PCR AMPLIFICATION OF VNTR AND *ompA* SEQUENCES

MLVA-*ompA* analysis was carried out at Molecular Microbiology and Infection, School of Medicine, University of Southampton. VNTR and *ompA* sequences were amplified using PCR. *Omp*A was amplified using primers P1F (Frost et al., 1991) and CT5 (Rodriguez et al., 1991), whilst the three VNTR regions were amplified using primers outlined in Pedersen et al. (2008). PCR reactions were carried out in 20 μl volumes with the components consisting of 10 μl of Phusion high fidelity DNA polymerase mastermix (New England Biolabs UK, Hitchin, UK), 250 nM of the forward and reverse primers and 2 μl of DNA.

PCR conditions for the three VNTR sequences and *ompA* were as follows: 10 s at 98°C, and 35 cycles of 2 s at 98°C, 5 s at 59°C, and 10 s at 72°C, followed by an elongation step for 1 min at 72°C. The temperature was then maintained at 10°C. All strains amplified had a CT1291 band that was 225 bp in size.

Purification of the PCR products was carried out using Wizard SV gel and PCR clean-up system (Promega, Southampton, UK) and this was carried out using the manufacturer's instructions.

SEQUENCING OF MLVA-*ompA* MARKERS AND DETERMINATION OF SEQUENCE STABILITY

Sequencing of PCR products were carried out by a commercial company (Source Bioscience, Nottingham, UK). Sequencing primers are described in Wang et al. (2011). *Omp*A sequences were initially (before chlamydial passage) compared with sequences available on the NCBI database using a BLAST search and in doing so the *ompA* genotype was assigned to each isolate. The VNTR sequences were compared to the whole genome

of D/UW-3/CX and thereafter assigned MLVA-types based on the dominant sequence in the chromatograms according to rules outlined in Pedersen et al. (2008). After the isolates were taken through eight passages, the chlamydial *omp*A and VNTR sequences were compared to sequence data received before the isolates had been passaged.

The Swedish new variant, which was passaged 72 times, was sequenced before passaging began and then every twelfth passage up to 72 passages. These sequences were compared to the database sequences and also to the sequence before the new variant had been passaged.

SEQUENCING OF MST MARKERS AND DETERMINATION OF SEQUENCE STABILITY

MST was carried out at the Section of Clinical Bacteriology, Department of Medical Sciences, Uppsala University. The MST scheme comprised five highly variable target regions and was performed as previously described (Klint et al., 2007) except that the *pbpB* region was amplified as two separate fragments according to Jurstrand et al. (2010), and for sequencing of the longer PCR products CT058 and *pbp*B internal sequence primers were used in the scheme (Herrmann, personal communication). DNA samples of chlamydia cell cultures passaged 72 times, sent from Southampton to Uppsala as part of a batch of samples that had been random coded so that sequence analysis was unbiased. Sequences were compared to the sequence at each locus to all known corresponding alleles available in the Uppsala University *C. trachomatis* MLST database (http://mlstdb.bmc.uu.se), 178 sequence types comprising 183 allele variants in March 2011. Sequences were thereafter compared to the original sequence of the *C. trachomatis* strain in question, as obtained before starting passage.

RESULTS

DETERMINATION OF THE STABILITY OF THE MLVA-*ompA* TYPING SYSTEM

Seven endocervical *C. trachomatis* isolates from the Southampton women's study (Wang et al., 2011) were selected to analyse the stability of the MLVA-*omp*A markers in cell culture. The criteria for choosing these isolates were that they had to be capable of at least two passages in cell culture and were, therefore, capable of being further passaged for this study. Secondly all isolates selected for marker stability had to have different *omp*A genotypes and they all had to have gone through the same number of passages in order for the data to be comparable. The isolates included in this study comprised of seven *omp*A genotypes, which corresponded to the following NCBI database strains: D/IC-CAL8, D/UW-3, E/Bour, K/UW-31, F/IC-CAL3, G/392, Ia/870 (**Table 1**). For this study we wanted to include samples with shared MLVA types (i.e., samples that had the same MLVA types but different *omp*A types, these included local isolates which had identical *omp*A sequences to the database sequence of D/IC-CAL8, D/UW-3, and E/BOUR) (**Table 1**) so we could identify if there were any differences in stability of the MLVA markers between different *omp*A types.

The passaging of the isolate initially began with each isolate being grown in a 24-well tray, then in a six-well tray and then into flasks (T25 and T75). By the eighth passage all isolates were now growing in a T75 flask where there were enough elementary bodies available to infect all the cells. DNA was then extracted from the elementary bodies and PCR reactions conducted for sequencing purposes. The chromatograms and sequence data received were aligned to sequences obtained before the isolates were taken through eight passages. Analysis of the alignments of the loci showed that the markers were stable upon adaptation of the chlamydia to cell culture as the sequence was unchanged from the sequence before the sample had been cultured in mouse cells and also that there is no change in the dominant sequence after eight passages in cell culture.

PASSAGING OF THE SWEDISH NEW VARIANT AND DETERMINATION OF THE STABILITY OF MLVA-*ompA* MARKERS IN CELL CULTURE

The Swedish new variant (serotype E with a 377 bp deletion in the cryptic plasmid) isolated in Malmo, Sweden was taken through a further 72 passages in cell culture (**Figure 1**) in Southampton, UK, the process took one year to complete. The Swedish new variant was chosen because this isolate shows restricted cell tropism (Unemo et al., 2010) and there is also greater possibility that the MLVA-*omp*A markers are unstable in this "mutant" because the strain was evidently mutable based on the plasmid sequence. Sequences for markers from every twelfth passage were determined up to and including passage 72 (passage 0, 12, 24, 36, 48, 60, and 72). All the sequence data were aligned and there were no changes seen in the MLVA-*omp*A profile at any passage that had been sequenced.

DETERMINATION OF THE STABILITY OF THE MST TYPING SYSTEM IN CELL CULTURE

The stability of the *C. trachomatis* MST markers in cell culture was also determined for the Swedish new variant which had been passaged 72 times (**Figure 1**). For this study five MST regions were analyzed these included; hctB, CT058, CT144, CT172, and pbpB and were first described by Klint et al. (2007). The five regions were sequenced and then aligned against corresponding sequences for each marker that was available from the sequence data of the Swedish new variant. The MST profile obtained was 21-19-1-2-1 and this was identical to the MST profile of the Swedish new variant. This data shows that these five MST markers are also stable in tissue culture up to 72 passages.

DISCUSSION

There are currently five DNA typing systems published for genotyping *C. trachomatis* (Klint et al., 2007; Pannekoek et al., 2008; Pedersen et al., 2008; Dean et al., 2009; Bom et al., 2011). These typing systems use a number of regions within the *C. trachomatis* genome to assign distinct labels to clinical specimens containing *C. trachomatis* based upon the order of nucleotides in the sequence. The repeating sequences of VNTR loci build up due to an increased error rate during DNA replication at these sites, therefore, this raised error rate means that the stability of such markers over multiple bacterial cell-divisions must be verified to validate their use in epidemiology studies. To be useful for epidemiological studies the markers need to remain stable over the duration of the study allowing the data collected to be comparable

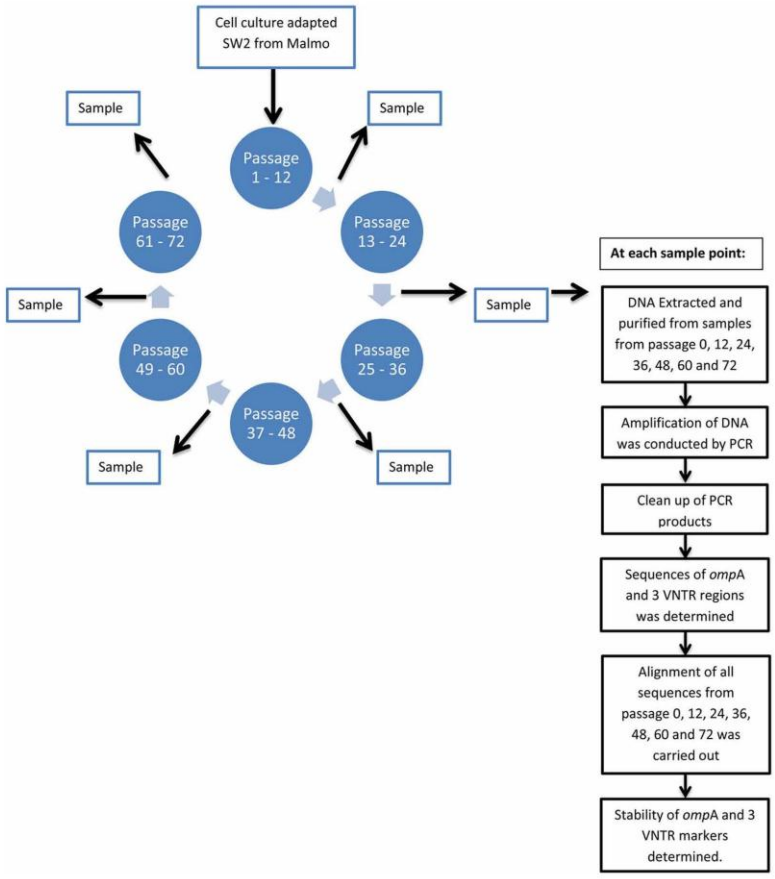


FIGURE 1 | Flow diagram showing the passage process of Sweden 2 and how the sample was processed to determine the stability of *ompA* and the VNTR markers.

(van Belkum et al., 2007). Use of genomic regions which are unstable may over represent the true distribution of bacterial genotypes within a community. It is, therefore, imperative to evaluate and understand the stability of the markers chosen for epidemiological studies.

Utilization of the gene for the major outer membrane protein A (*ompA*) has, for a long time been one of the main typing methods for distinguishing *C. trachomatis*. However, this method actually provides little discrimination (Pedersen et al., 2008; Bom et al., 2011) and masks the true phylogenetic relationships of the strains (Harris et al., 2012). Three "MLST" schemes have also been developed, two of which use seven housekeeping genes to determine overall population structure of the whole chlamydiae family (Pannekoek et al., 2008; Dean et al., 2009) and lastly a scheme which is based on five highly variable genes, some of which are under selection (which is designed to ensure high resolution) (Klint et al., 2007). Multi sequence typing (MST) best describes the typing scheme designed by Klint et al. (2007) and is the most discriminatory of the three "MLST" typing schemes available for *C. trachomatis*. The fifth typing system available for *C. trachomatis* is the MLVA-*ompA* scheme which uses three variable regions within the genome to assign types to each isolate (Pedersen et al., 2008). These DNA markers were selected for typing because of their variability, therefore, it is particularly important to assess the stability of these markers in continuous cell culture.

The stability of the DNA markers used in the two typing methods with the highest discriminatory power was assessed; the Pedersen et al. MLVA-*omp*A typing system Pedersen et al. (2008) and the Klint et al. (2007) MST typing system. Although *omp*A had previously been shown to be stable through adaptation of *C. trachomatis* to cell culture (Stothard et al., 1998), more recent epidemiological studies suggest that within a single clone (e.g., nvCT—according to the MST scheme), minor single base substitutions occasionally appear (Herrmann et al., 2008); a similar trend has been observed in LGV isolates (Christerson et al., 2010).

Our hypothesis was that adaptation of *C. trachomatis* to cell culture conditions could affect marker stability, however, on the contrary our data shows that the dominant MLVA sequences are stable on adaptation of the *C. trachomatis* to cell culture. The MLVA-*omp*A sequences were analyzed from DNA extracts taken from primary swab samples (Wang et al., 2011) and after eight passages in cell culture which is equivalent to approximately 80 divisions. To ensure our study covered a wide spread of *C. trachomatis* types all the isolates chosen were selected on the basis that they had different MLVA-*omp*A types. Three isolates with different *omp*A types but the same MLVA profile were also included. This could be indicative of recombination of the *omp*A gene. Choosing isolates with different *omp*A types also acted as a verification that no contamination between strains had occurred during passaging of the isolates. Although the *in vivo* conditions in which *C. trachomatis* persists are much more complex (involves the immune system etc.) than that of the *in vitro* system that is described here, the basis for taking the samples through eight passages to analyse the initial impact in cell culture was to mimic the actual time it takes for symptoms to appear within an individual. Symptoms usually arise in individuals seven to twenty-one days after infection occurs (Black, 1997), therefore, the isolates were taken through twenty-four days from infection to harvesting in cell culture to ensure the total time from being infected to symptoms persisting in individuals was covered over the eight passages.

A review by Bjorn-Arne Lindstedt states that it is imperative that the VNTR loci selected for MLVA typing are thoroughly checked before being applied to real-life situations, this includes long-term passaging and re-typing to ensure the stability and the reproducibility of each locus (Lindstedt, 2005). Pedersen et al. (2008) previously assessed the markers in patients with recurrent or persistent *C. trachomatis*. Others have also evaluated the MLVA system and found that some VNTR-markers may vary with replication of single clones and cause difficulties in interpretation (Bom et al., 2011). Analysis of the stability in a controlled environment by long term passaging of isolates was, therefore, important thus the second part of our study was performed to determine the stability of the MLVA-*omp*A and MST markers (Klint et al., 2007) over seventy two passages using the Swedish new variant. This isolate had already been subject to change *in vivo* as illustrated by both a deletion and duplication within the plasmid and may, therefore, be an unstable strain (or perhaps a strain which has encountered an unknown selective pressure causing the mutation), making this a useful isolate to test stability over a long duration. Three studies carried out in Sweden have shown that the Swedish new variant is clonal. The first study by Herrmann et al. (2008) typed 48 specimens from different regions in Sweden and three other countries. The study not only showed that all isolates were genotype E according to the *omp*A typing system but also that 96% of the isolates had identical sequence types when typed using a highly discriminatory MST method (Klint et al., 2007). Later studies by Jurstrand et al. (2010) also showed that all specimens typed ($n = 41$) had the same *omp*A type (genotype E) and the same MST sequence types. The sequence type assigned to all samples was unique for the Swedish new variant. The third study included nvCT cases from 2009 and although only 13 specimens were analyzed it supports the stability of the MST target regions over time (Klint et al., 2011). Although this evidence suggested that the Swedish new variant is clonal and may be stable *in vivo* the transmission of the new variant from patient to patient in these studies remains unknown and, therefore, it is also unknown as to how many passages the new variant has been through. Our *in vitro* studies were conducted under carefully controlled conditions so the number of continuous passages were defined.

In conclusion MLVA-*omp*A markers are stable through adaptation of *C. trachomatis* to growth *in vitro* and are also stable over eight passages in cell culture. We have also cultured a mutable isolate (the Swedish new variant of *C. trachomatis*) through 72 passages and showed that the MLVA-*omp*A and MST markers are stable, however, complete definition of evolutionary adaptation requires more isolates and extensive further passaging.

ACKNOWLEDGMENTS

We acknowledge the financial support from the HPA PhD Fund for Clare Labiran. We would like to thank Prof. Nicholas Thomson and Dr. Helena Seth-Smith of The Wellcome Trust Sanger Institute for their expert advice. We also acknowledge Dr. Stuart C. Clarke for his help in securing funding for the PhD project.

REFERENCES

Black, C. M. (1997). Current methods of laboratory diagnosis of *Chlamydia trachomatis* infections. *Clin. Microbiol. Rev.* 10, 160–184.

Bom, R. J., Christerson, L., Schim van der Loeff, M. F., Coutinho, R. A., Herrmann, B., and Bruisten, S. M. (2011). Evaluation of high-resolution typing methods for *Chlamydia trachomatis* in samples from heterosexual couples. *J. Clin. Microbiol.* 49, 2844–2853.

Christerson, L., de Vries, H. J., de Barbeyrac, B., Gaydos, C. A., Henrich, B., Hoffmann, S., Schachter, J., Thorvaldsen, J., Vall-Mayans, M., Klint, M., Herrmann, B., and Morre, S. A. (2010). Typing of lymphogranuloma venereum *Chlamydia trachomatis* strains. *Emerg. Infect. Dis.* 16, 1777–1779.

Dean, D., Bruno, W. J., Wan, R., Gomes, J. P., Devignot, S., Mehari, T., de Vries, H. J. C., Morre, S. A., Myers, G., Read, T. D., and Spratt, B. G. (2009). Predicting phenotype and emerging strains among *Chlamydia trachomatis* infections. *Emerg. Infect. Dis.* 15, 1385–1394.

Frost, E. H., Deslandes, S., Veilleux, S., and Bourgaux-Ramoisy, D. (1991). Typing *Chlamydia trachomatis* by detection of restriction fragment length polymorphism in the gene encoding the major outer membrane protein. *J. Infect. Dis.* 163, 1103–1107.

Harris, S. R., Clarke, I. N., Seth-Smith, H. M., Solomon, A. W., Cutcliffe, L. T., Marsh, P., Skilton, R. J., Holland, M. J., Mabey, D., Peeling, R. W., Lewis, D. A., Spratt, B. G., Unemo,

M., Persson, K., Bjartling, C., Brunham, R., de Vries, H. J., Morre, S. A., Speksnijder, A., Bebear, C. M., Clerc, M., de Barbeyrac, B., Parkhill, J., and Thomson, N. R. (2012). Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat. Genet.* 44, 413–419.

Herrmann, B., Torner, A., Low, N., Klint, M., Nilsson, A., Velicko, I., Soderblom, T., and Blaxhult, A. (2008). Emergence and spread of *Chlamydia trachomatis* variant, Sweden. *Emerg. Infect. Dis.* 14, 1462–1465.

Ikryannikova, L. N., Shkarupeta, M. M., Shitikov, E. A., Il'ina, E. N., and Govorun, V. M. (2010). Comparative evaluation of new typing schemes for urogenital *Chlamydia trachomatis* isolates. *FEMS Immunol. Med. Microbiol.* 59, 188–196.

Jurstrand, M., Christerson, L., Klint, M., Fredlund, H., Unemo, M., and Herrmann, B. (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.

Klint, M., Fuxelius, H. H., Goldkuhl, R. R., Skarin, H., Rutemark, C., Andersson, S. G. E., Persson, K., and Herrmann, B. (2007). High-resolution genotyping of *Chlamydia trachomatis* strains by multilocus sequence analysis. *J. Clin. Microbiol.* 45, 1410–1414.

Klint, M., Hadad, R., Christerson, L., Lor, B., Anagrius, C., Osterlund, A., Larsson, I., Sylvan, S., Fredlund, H., Unemo, M., and Herrmann, B. (2011). Prevalence trends in Sweden for the new variant of *Chlamydia trachomatis*. *Clin. Microbiol. Infect.* 17, 683–689.

Lindstedt, B. A. (2005). Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis* 26, 2567–2582.

Pannekoek, Y., Morelli, G., Kusecek, B., Morre, S. A., Ossewaarde, J. M., Langerak, A. A., and van der Ende, A. (2008). Multi locus sequence of Chlamydiales: clonal groupings within the obligate intracellular bacteria *Chlamydia trachomatis*. *BMC Microbiol.* 8, 42.

Pedersen, L. N., Herrmann, B., and Moller, J. K. (2009). Typing *Chlamydia trachomatis*: from egg yolk to nanotechnology. *FEMS Immunol. Med. Microbiol.* 55, 120–130.

Pedersen, L. N., Podenphant, L., and Moller, J. K. (2008). Highly discriminative genotyping of *Chlamydia trachomatis* using omp1 and a set of variable number tandem repeats. *Clin. Microbiol. Infect.* 14, 644–652.

Ripa, T., and Nilsson, P. (2006). A variant of *Chlamydia trachomatis* with deletion in cryptic plasmid: implications for use of PCR diagnostic tests. *Euro Surveill.* 11, E061109.

Ripa, T., and Nilsson, P. A. (2007). A *Chlamydia trachomatis* strain with a 377-bp deletion in the cryptic plasmid causing false-negative nucleic acid amplification tests. *Sex. Transm. Dis.* 34, 255–256.

Rodriguez, P., Vekris, A., de, B. B., Dutilh, B., Bonnet, J., and Bebear, C. (1991). Typing of *Chlamydia trachomatis* by restriction endonuclease analysis of the amplified major outer membrane protein gene. *J. Clin. Microbiol.* 29, 1132–1136.

Schouls, L. M., Spalburg, E. C., van Luit, M., Huijsdens, X. W., Pluister, G. N., van Santen-Verheuvel, M. G., and de Neeling, A. J. (2009). Multiple-locus variable number tandem repeat analysis of *Staphylococcus aureus*: comparison with pulsed-field gel electrophoresis and spa-typing. *PLoS ONE* 4:e5082. doi: 10.1371/journal.pone.0005082

Stothard, D. R., van der Pol, B., Smith, N. J., and Jones, R. B. (1998). Effect of serial passage in tissue culture on sequence of omp1 from *Chlamydia trachomatis* clinical isolates. *J. Clin. Microbiol.* 36, 3686–3688.

Unemo, M., Seth-Smith, H. M., Cutcliffe, L. T., Skilton, R. J., Barlow, D., Goulding, D., Persson, K., Harris, S. R., Kelly, A., Bjartling, C., Fredlund, H., Olcen, P., Thomson, N. R., and Clarke, I. N. (2010). The Swedish new variant of *Chlamydia trachomatis*: genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* 156, 1394–1404.

van Belkum, A., Tassios, P. T., Dijkshoorn, L., Haeggman, S., Cookson, B., Fry, N. K., Fussing, V., Green, J., Feil, E., Gerner-Smidt, P., Brisse, S., and Struelens, M. (2007). Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin. Microbiol. Infect.* 13 (Suppl 3), 1–46.

Wang, Y., Skilton, R. J., Cutcliffe, L. T., Andrews, E., Clarke, I. N., and Marsh, P. (2011). Evaluation of a high resolution genotyping method for *Chlamydia trachomatis* using routine clinical samples. *PLoS ONE* 6:e16971. doi: 10.1371/journal.pone.0016971

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 January 2012; accepted: 30 April 2012; published online: 17 May 2012.

Citation: Labiran C, Clarke IN, Cutcliffe LT, Wang Y, Skilton RJ, Persson K, Bjartling C, Herrmann B, Christerson L and Marsh P (2012) Genotyping markers used for multi locus VNTR analysis with ompA (MLVA-ompA) and multi sequence typing (MST) retain stability in Chlamydia trachomatis. Front. Cell. Inf. Microbio. 2:68. doi: 10.3389/fcimb.2012.00068

APPENDIX 2: REVIEWER COMMENTARY ON MARKER STABILITY PAPER

frontiers in
CELLULAR AND INFECTION MICROBIOLOGY

GENERAL COMMENTARY
published: 06 June 2012
doi: 10.3389/fcimb.2012.00077

Genomic stability of genotyping markers in *Chlamydia trachomatis*

João Paulo Gomes*

Department of Infectious Diseases, National Institute of Health, Lisbon, Portugal
**Correspondence: j.paulo.gomes@insa.min-saude.pt*

A commentary on

Genotyping markers used for multi locus VNTR analysis with ompA (MLVA-ompA) and multi sequence typing retain stability in *Chlamydia trachomatis*
by Labiran, C., Clarke, I. N., Cutcliffe, L. T., Wang, Y., Skilton, R. J., Persson, K., Bjartling, C., Herrmann, B., Christerson, L., and Marsh, P. (2012). Front. Cell. Infect. Microbiol. 2:68. doi: 10.3389/fcimb.2012.00068

Chlamydia trachomatis is an obligate intracellular human pathogen that infects columnar epithelial cells of ocular or genital mucosae, and mononuclear phagocytes. More than 50 fully sequenced genomes are publically available (Harris et al., 2012), and despite low polymorphic character when compared with most bacteria, these data allow a better understanding of *C. trachomatis* chromosomal dynamics. Indeed, it revealed not only the chromosome regions more prone to genetic mutations, but also that recombination is highly chromosome dispersed, although the frequency and relative weight of recombination and mutation events undoubtedly posits *C. trachomatis* as a low recombining bacterium (Joseph et al., 2011; Ferreira et al., 2012).

In order to understand the epidemiology and pathogenesis of the highly prevalent chlamydial diseases, the classification of clinical isolates has been a priority of researchers for decades. However, the traditional typing system, which is based on the polymorphism of the gene (*ompA*) encoding the major outer membrane protein (Yuan et al., 1989) has been jeopardized in the last years as it does not group the isolated strains according to their cell-appetence or virulence. Thus, with the progressive release of sequenced genomes, several typing systems have been developed (Klint et al., 2007; Pannekoek et al., 2008; Pedersen et al., 2008; Dean et al., 2009; Bom et al., 2011), although none of them is consensual due to the dissimilar genetic characteristics or discriminatory power of the loci enrolled in each of those.

Although it is assumed that the typing schemes should ideally ensure that the chosen loci are stable over the course of evolution, this issue is of difficult assessment. Labiran et al. (2012) have now presented a study where they have assessed the stability of some regions of the *C. trachomatis* genome by using innovate approaches. They have used the genetic markers from the described typing schemes that provide the highest Simpson index of diversity: the multi-sequence typing (MST) system described by Klint et al. (2007) and the multi locus variable number tandem repeat (VNTR) analysis (MLVA-*ompA*) described by Pedersen et al. (2008). Labiran et al. (2012) have initially evaluated the stability of those genetic markers through the first stages of adaptation of *C. trachomatis* to cell culture. This is interesting as the first attempt to study mutational patterns over adaptation to cell culture had been performed back in 1998 (Stothard et al., 1998), by studying solely the stability of *ompA* over 20 *in vitro* passages. Now, by analyzing three VNTR besides *ompA*, the authors undoubtedly increase the chances of observing genetic alterations, as these loci possess a high discriminatory power. Furthermore, as the repeated sequences of VNTR loci are generated as a consequence of an increased error rate during DNA replication at these regions, they are hypothetically more prone to additional variation. In this first approach, the authors found no alterations in those genomic regions by studying seven clinical isolates over eight passages. The initial inoculations were properly done from primary swabs in order to avoid previous contact of the strains with the *in vitro* system. Interestingly, the eight passages represent about 16 days in culture aiming to reflect the *in vivo* estimated period for appearance of symptoms for *C. trachomatis* infection of both genitalia and conjunctivae (Black, 1997). There is no doubt that there is a radical change of environmental conditions when transiting from the *in vivo* to the *in vitro* system, mostly because the latter lacks immune system, competing flora, pH alterations, hormonal fluctuations, etc. Nevertheless, although the parallelism between these two scenarios remains speculative for these reasons, the impossibility to perform this evaluation *in vivo* and under controlled conditions makes this an interesting approach.

In a second stage, the authors went even further by studying the "long-term" genome stability under influence of *in vitro* environmental conditions. To achieve this, they used the MLVA-*ompA* and also the five-locus MST system to monitor the stability of the enrolled loci throughout 72 passages in tissue culture. The coverage and the runtime of this approach make it an unprecedented evaluation of genome stability. In this study Labiran et al. (2012) used the so called "Swedish new variant," a *C. trachomatis* isolate that, besides the higher number of pseudogenes than most sequenced genital isolates (Jeffrey et al., 2010), shows both a deletion and duplication within the plasmid, which are genomic changes that had never been described before (Unemo et al., 2010). These kind of genomic alterations are infrequent as *C. trachomatis* presents a nearly identical core and pan genome and a DNA sequence similarity of >98% (Harris et al., 2012). So, this isolate could hypothetically be considered as genomically unstable, making this a valuable strain to test stability over a long time *in vitro* passage. Similarly to the previous approach, no genetic changes were observed throughout the long-term *in vitro* maintenance, suggesting that the MLVA-*ompA* and MST systems are, on a genetic stability basis, highly accurate for both short-term (such as partner tracing) and long-term molecular epidemiology.

Moreover, the results from the Labiran et al. (2012) study indirectly bring to light a major controversy in the field, which is the uncertainty of whether the prototype

strains that have been used worldwide leading to thousands of papers, accurately represent current circulating clinical strains. In fact, the former are laboratory adapted for decades but at this stage one cannot know the extent of genomic alterations that may occur due to laboratory adaptation. On one hand, the results from Labiran et al. (2012) point to specific genomic stability, which, together with the *C. trachomatis* transmission bottlenecks (Andersson and Hughes, 1996) and the high doubling time of 2.5–4 h (Borges et al., 2010), disfavors genetic variability. On the other hand, it is known that some loci are prone to generate multi-alleles, such as the putative virulence factor CT135 (Sturdevant et al., 2010), and that *Chlamydia* undergoes adaptive changes when subjected to specific environmental alterations (Kutlin et al., 2005; Suchland et al., 2005). Ultimately, unless the entire genome is analyzed (at the stage of the strain isolation and after long-term *in vitro* passage), the controversy will certainly remain.

REFERENCES

Andersson, D. I., and Hughes, D. (1996). Muller's ratchet decreases fitness of a DNA-based microbe. *Proc. Natl. Acad. Sci. U.S.A.* 93, 906–907.

Black, C. M. (1997). Current methods of laboratory diagnosis of *Chlamydia trachomatis* infections. *Clin. Microbiol. Rev.* 10, 160–184.

Bom, R. J., Christerson, L., Schim Van Der Loeff, M. F., Coutinho, R. A., Herrmann, B., and Bruisten, S. M. (2011). Evaluation of high-resolution typing methods for *Chlamydia trachomatis* in samples from heterosexual couples. *J. Clin. Microbiol.* 49, 2844–2853.

Borges, V., Ferreira, R., Nunes, A., Nogueira, P., Borrego, M. J., and Gomes, J. P. (2010). Normalization strategies for real-time expression data in *Chlamydia trachomatis*. *J. Microbiol. Methods* 82, 256–264.

Dean, D., Bruno, W. J., Wan, R., Gomes, J. P., Devignot, S., Mehari, T., De Vries, H. J. C., Morré, S. A., Myers, G., Read, T. D., and Spratt, B. G. (2009). Predicting phenotype and emerging strains among *Chlamydia trachomatis* infections. *Emerging Infect. Dis.* 15, 1385–1394.

Ferreira, R., Borges, V., Nunes, A., Nogueira, P. J., Borrego, M. J., and Gomes, J. P. (2012). Impact of loci nature on estimating recombination and mutation rates in *Chlamydia trachomatis*. *G3 (Bethesda)*. 2:761–768. doi: 10.1534/g3.112.002923

Harris, S. R., Clarke, I. N., Seth-Smith, H. M., Solomon, A. W., Cutcliffe, L. T., Marsh, P., Skilton, R. J., Holland, M. J., Mabey, D., Peeling, R. W., Lewis, D. A., Spratt, B. G., Unemo, M., Persson, K., Bjartling, C., Brunham, R., de Vries, H. J., Morré, S. A., Speksnijder, A., Bébéar, C. M., Clerc, M., de Barbeyrac, B., Parkhill, J., and Thomson, N. R. (2012). Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat. Genet.* 44, 413–419.

Jeffrey, B. M., Suchland, R. J., Quinn, K. L., Davidson, J. R., Stamm, W. E., and Rockey, D. D. (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.

Joseph, S. J., Didelot, X., Gandhi, K., Dean, D., and Read, T. D. (2011). Interplay of recombination and selection in the genomes of *Chlamydia trachomatis*. *Biol. Direct* 6, 28.

Klint, M., Fuxelius, H. H., Goldkuhl, R. R., Skarin, H., Rutemark, C., Andersson, S. G. E., Persson, K., and Herrmann, B. (2007). High-resolution genotyping of *Chlamydia trachomatis* strains by multilocus sequence analysis. *J. Clin. Microbiol.* 45, 1410–1414.

Kutlin, A., Kohlhoff, S., Roblin, P., Hammerschlag, M. R., and Riska, P. (2005). Emergence of resistance to rifampin and rifalazil in *Chlamydophila pneumoniae* and *Chlamydia trachomatis*. *Antimicrob. Agents Chemother.* 49, 903–907.

Labiran, C., Clarke, I. N., Cutcliffe, L. T., Wang, Y., Skilton, R. J., Persson, K., Bjartling, C., Herrmann, B., Christerson, L., and Marsh, P. (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: 10.3389/fcimb.2012.00068

Pannekoek, Y., Morelli, G., Kusecek, B., Morre, S. A., Ossewaarde, J. M., Langerak, A. A., and Van Der Ende, A. (2008). Multi locus sequence of Chlamydiales: clonal groupings within the obligate intracellular bacteria *Chlamydia trachomatis*. *BMC Microbiol.* 8, 42. doi: 10.1186/1471-2180-8-42

Pedersen, L. N., Podenphant, L., and Moller, J. K. (2008). Highly discriminative genotyping of *Chlamydia trachomatis* using omp1 and a set of variable number tandem repeats. *Clin. Microbiol. Infect.* 14, 644–652.

Stothard, D. R., Van Der Pol, B., Smith, N. J., and Jones, R. B. (1998). Effect of serial passage in tissue culture on sequence of omp1 from *Chlamydia trachomatis* clinical isolates. *J. Clin. Microbiol.* 36, 3686–3688.

Sturdevant, G. L., Kari, L., Gardner, D. J., Olivares-Zavaleta, N., Randall, L. B., Whitmire, W. M., Carlson, J. H., Goheen, M. M., Selleck, E. M., Martens, C., and Caldwell, H. D. (2010). Frameshift mutations in a single novel virulence factor alter the in vivo pathogenicity of *Chlamydia trachomatis* for the female murine genital tract. *Infect. Immun.* 78, 3660–3668.

Suchland, R. J., Bourillon, A., Denamur, E., Stamm, W. E., and Rothstein, D. M. (2005). Rifampin-resistant RNA polymerase mutants of *Chlamydia trachomatis* remain susceptible to the ansamycin rifalazil. *Antimicrob. Agents Chemother.* 49, 1120–1126.

Unemo, M., Seth-Smith, H. M., Cutcliffe, L. T., Skilton, R. J., Barlow, D., Goulding, D., Persson, K., Harris, S. R., Kelly, A., Bjartling, C., Fredlund, H., Olcén, P., Thomson, N. R., and Clarke, I. N. (2010). The Swedish new variant of *Chlamydia trachomatis*: genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* 156, 1394–1404.

Yuan, Y., Zhang, Y. X., Watkins, N. G., and Caldwell, H. D. (1989). Nucleotide and deduced amino acid sequences for the four variable domains of the major outer membrane proteins of the 15 *Chlamydia trachomatis* serovars. *Infect. Immun.* 57, 1040–1049.

Received: 14 May 2012; accepted: 18 May 2012; published online: 06 June 2012.

Citation: Gomes JP (2012) Genomic stability of genotyping markers in Chlamydia trachomatis. Front. Cell. Inf. Microbio. 2:77. doi: 10.3389/fcimb.2012.00077

APPENDIX 3: INFORMATION SHEET FOR BRIGHTON STUDY (CHP 4).

A study to investigate the distribution of Chlamydia Trachomatis Strains in men who have sex with men attending a Sexual Health clinic in Brighton.

INFORMATION ABOUT THE RESEARCH

You are being invited to take part in a research study.
Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

Why is this study being carried out?
Chlamydia is a common cause of infection in the rectum (bum) and the urethra (water passage) and it is routinely tested for as part of a sexual health check. We know that different types of Chlamydia exist and we would like to see which types of Chlamydia are commonest in men who have sex with men.

What are the benefits of the study?
The results of this study will help us to understand how Chlamydia is transmitted and therefore how we can take measures to reduce the amount of infection.

Why have I been invited?
You have been invited to participate because we are only studying men who have sex with men.

Do I have to take part?
No. It is up to you to decide. We will describe the study and go through this information sheet, which we will then give to you. We will then ask you to sign a consent form to show you have agreed to take part.

What will happen today if I take part?
As part of your routine care, you will be seen by a nurse or a doctor who will ask you some questions and examine you. During the examination swabs are taken; usually we take two swabs from the rectum but for the study we would take one additional swab. As it is taken at the same time as the other swabs, you won't feel it. You will also be asked to provide a urine sample to test for Chlamydia and gonorrhoea, as is routine. For the study we will reserve some of the leftover urine which will only be used if your routine Chlamydia test is positive. So you only need to give us one urine sample.

What will happen if my chlamydia test result is positive?
As is normal practice, we would call you to tell you your result is positive and invite you to come back to clinic for antibiotic treatment.
For the study we want to send urine and swabs from all men with Chlamydia to Southampton for further testing (to look at the type of Chlamydia) so the extra samples you gave us will be sent away. You will not be informed of the results of the study because at the moment we do not know what the implications of these results are. We would also like to go back to your notes to gather

information about your age, ethnic status, symptoms, HIV status and sexual behaviour to see if there is any association with the results. All the information will be kept strictly confidential and there will be no link between the study information and your notes (anonymised).

What will happen if my Chlamydia test result is negative?
The extra specimens you gave us will be discarded and you will no longer be involved in the study. You will be informed of your negative results by a text message received 3 weeks after your clinic visit.

Will my taking part in the study be kept confidential?
Yes. We will follow ethical and legal guidelines and all information about you will be handled in confidence.

We will also collect information from your notes but we will only record your hospital or clinic number. No names or addresses will be kept on the database and only the research team will have access to the data

What happens if I don't want to take part?
If you decide not to participate it will not affect the care you will receive.

The team will be able to provide answers to any further questions you may have after reading this information leaflet and will be able to provide any support you need.

If you have any complaints about the conduct of the study you can contact the Trust's Patient Advisory Liaison Service (PALS) on 01273 696955.

Thank you very much for taking time to read this leaflet. If you have any further questions please do not hesitate to contact Dr. Judith Zhou or Dr Suneeta Soni.

Brighton Investigators:

Dr. Judith Zhou and Dr. Suneeta Soni
Lawson Unit, Royal Sussex County Hospital, Brighton BN2 5BE
Tel 01273 664718

Southampton Investigators:

Professor Ian Clarke and Miss Claire Labiran
University of Southampton, Southampton General Hospital, Southampton SO16 6YD
Tel 023 80796975

Dr Peter Marsh
Health Protection Agency, Regional Microbiology Network Southampton Laboratory

APPENDIX 4 : CONSENT FORM FOR BRIGHTON STUDY (CHP 4).

Brighton and Sussex University Hospitals NHS Trust
Elton John Centre
HIV/GUM Research
Sussex House
1 Abbey Road
Brighton BN2 1ES

Tel: 01273 523079
Fax: 01273 523080

CONSENT FORM

Title of Project: A study to determine the distribution of Chlamydia trachomatis in types in men who have sex with men attending a GUM clinic in Brighton.

Name of Researcher: Dr. Suneeta Soni

Please initial box

1. I confirm that I have read and understand the information sheet dated 22 August 2011 (version 3.0) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐

3. I give permission for my medical records to be looked at by representatives from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. ☐

4. I give permission for researchers to access my medical records to obtain information about sexual behaviour ☐

4. I agree to take part in the above study. ☐

Name of Patient	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

When completed, 1 for patient; 1 for researcher site file (original)

APPENDIX 5 : BD PROBETEC DATA SHEET.

BD Male Urethral Specimen Collection Kit for the BD ProbeTec™ *Chlamydia trachomatis/ Neisseria gonorrhoeae* (CT/GC) Q^x Amplified DNA Assays

English: pages 1 – 2
Dansk: side 2 – 3
Deutsch: Seiten 3 – 4
Español: páginas 5 – 6
Français : pages 6 – 7
Norsk: sider 7 – 8
Svenska: sidan 9 – 10

CE 0086

L009851
2010/12

Pokyny vám poskytne místní zástupce společnosti BD. / Kontakt den lokale BD repræsentant for at få instruktioner. / Kasutusjuhiste suhtes kontakteeruge oma kohaliku BD esindajaga. / Επικοινωνήστε με τον τοπικό αντιπρόσωπο της BD για οδηγίες. / A használati utasítást kérje a BD helyi képviseletétől. / Naudojimo instrukcijų teiraukitės vietos BD įgaliotojo atstovo. / Kontakt din lokale BD-representant for mer informasjon. / Aby uzyskać instrukcje użytkowania, skontaktuj się z lokalnym przedstawicielstwem BD. / Contacte o seu representante local da BD para obter instruções. / Inštrukcie získate u miestneho zástupcu spoločnosti BD. / Kontakta lokal Becton Dickinson-representant för anvisningar. / Свържете се с местния представител на BD за инструкзии. / Contactaţi reprezentantul dumneavoastră local BD pentru instrucţiuni. / Talimatlar için yerel BD temsilcilerinize danışın. / Obratite se svom lokalnom predstavniku kompanije BD za uputstva. / Для получения инструкций свяжетесь с местным представителем компании BD. / Өзіңіздің жергілікті БД өкіліне жүгініп нұсқау алыңыз. / Kontaktiraj lokalnog predstavnika BD za upute.

INTENDED USE

The Male Urethral Specimen Collection Kit for the **BD ProbeTec** CT/GC Q^x Amplified DNA Assays is used to collect and transport male patient urethral specimens to the laboratory for testing with the **BD ProbeTec** CT/GC Q^x Amplified DNA Assays on the **BD Viper™** System in Extracted Mode.

SUMMARY AND EXPLANATION

The Male Urethral Specimen Collection Kit for the **BD ProbeTec** CT/GC Q^x Amplified DNA Assays consists of a sterile swab and a tube of Q^x Swab Diluent for specimen transport. The rayon-tipped collection swab consists of a stainless steel wire in a blue-colored plastic shaft that is scored for breaking. After the specimen is collected, the cap of the diluent tube is removed, the swab is placed into the Q^x Swab Diluent tube and the shaft is broken at the score mark. The tube is recapped and transported to the testing facility.

REAGENTS

Materials Provided: Each Male Urethral Specimen Collection Kit for the **BD ProbeTec** CT/GC Q^x Amplified DNA Assays contains one sterile rayon-tipped collection swab and one tube of Q^x Swab Diluent. The Swab Diluent contains approximately 2 mL of potassium phosphate/potassium hydroxide buffer with DMSO and preservative.

Materials Required But Not Provided: Gloves

Storage Instructions: Store collection kit at 2 – 33°C. Do not use beyond the expiration date. Swab sterility is guaranteed if primary swab container is intact.

Warnings and Precautions:

For *in vitro* Diagnostic Use.

1. **Do not pre-wet the swab with the Q^x Swab Diluent prior to collecting specimen.**
2. Pathogenic microorganisms including hepatitis viruses and Human Immunodeficiency Virus may be present in clinical specimens. "Standard Precautions"[1-4] and institutional guidelines should be followed in handling all items contaminated with blood and other body fluids.
3. Optimal performance of the **BD ProbeTec** CT/GC Q^x Amplified DNA Assays requires proper specimen collection, handling and transport.
4. Use only collection and transport systems labeled for use with the **BD ProbeTec** CT/GC Q^x Amplified DNA Assays.
5. Proper labeling should accompany each specimen to the laboratory.
6. The Q^x Swab Diluent tube contains dimethyl sulfoxide (DMSO). DMSO is harmful by inhalation, in contact with skin and if swallowed. Keep container in a well-ventilated place. Avoid contact with eyes. In case of contact with eyes, rinse immediately with plenty of water and seek medical advice. After contact with skin, wash immediately with plenty of water. Wear suitable protective clothing and gloves.
7. When breaking the shaft of the swab, take care to avoid splashing of contents.
8. Dispose of unused Q^x Swab Diluent tubes according to the Material Safety Data Sheet.
9. Take care to avoid cross-contamination during the specimen handling steps. Specimens can contain high levels of organisms. Ensure that specimen containers do not contact one another, and discard used materials without passing over open containers. If gloves come in contact with specimen, change gloves to avoid cross-contamination.
10. Specimens must be collected and tested before the expiration date of the Q^x Swab Diluent tube.
11. ⊗ Male Urethral Specimen Collection Kit for the **BD ProbeTec** *Chlamydia trachomatis/Neisseria gonorrhoeae* (CT/GC) Q^x Amplified DNA Assays is for single use only; reuse may cause a risk of infection and/or inaccurate results.

APPENDIX 6 : VCM VALIDATION SHEET FOR CHLAMYDIA

Σ-VCM™ Validation with Chlamydia

Requirements for Validation

To show recovery of Chlamydia from specimens

The following study was carried out as part of an ongoing project at Addenbrookes Hospital, Cambridge.

200 specimens taken from patients for testing by conventional means were also transferred to VCM, tested by culture onto McCoys cells, and confirmed by immunofluorescence. Positive specimens were also further tested by quantitative real-time PCR.

10 specimens yielded positive cultures, which matched the results of the other methods.

Chlamydia results for Medical Wire transport media

1. Extraction of 10 positive samples was performed on routine corbett platform using

200ul of sample volume and elution volume of 100ul

2. The Quantitative real-time PCR was performed targetting CT - Pol (gene)

3. The CT Genotyping was performed using our in-house real-time assay targeting the MOMP gene

No	Sample I.D	Ct Value	Quantitation cp/ml	CT Genotype
1	1	31.6	39,203	H
2	3	33.3	12,319	E
3	6	32.4	23,108	E
4	22	29.3	208,388	J
5	29	30.5	88,175	F
6	36	35.2	3,309	F
7	38	35	3,061	G
8	73	28.9	274,711	G
9	85	40.5	78	NO TYPE
10	236	26.9	1,116,617	J

Ct = Cycle threshold

All of these specimens were grown on McCoy cells all were culture positive by IF Staining.

These results show that VCM gives good recovery of Chlamydia, matching other methods, and that specimens are of good quality allowing acceptable performance using a molecular method (RT-PCR)

APPENDIX 7 : INFORMATION SHEET FOR SOUTHAMPTON STUDY (CHP 5).

Medicine

UNIVERSITY OF Southampton

Mailpoint 814, Level C, South Block,
Southampton General Hospital,
Tremona Road,
Southampton,
SO16 6YD,
UK.
Tel:02380798886
Fax: 02380796992

INFORMATION ABOUT THE RESEARCH

Title of study: Study on the Epidemiology of Chlamydia in Southampton

You are being invited to take part in a research study.
Before you decide if you want to take part in this study you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is unclear or if you would like more information and we will be happy to answer these questions.

Thank you for taking time to read this information leaflet.

What is the purpose of the study?
Chlamydia is the most common sexually transmitted infection in the UK therefore we would like to conduct an educational study on the epidemiology of chlamydia in Southampton. From previous studies we know that there are different types of chlamydia circulating in the community and we would like to find out what types are associated with people from different age-groups, backgrounds, genders and the types associated with different symptoms.

What are the benefits of the study?
The results of the study will help us to understand how chlamydia is transmitted and why certain types of chlamydia persist in the community to a higher extent than other types.

Why have I been chosen?
All individuals who are undergoing chlamydia tests at this clinic are being invited to participate in this study.

Do I have to take part?
It is up to you to decide whether or not to take part. If you decide to take part you will be given this information sheet to keep and asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What happens if I don't want to take part?
If you decide not to participate it will not affect your normal healthcare. The team will be able to provide answers to any further questions you may have after reading this information leaflet.

Thank you very much for taking time to read this leaflet. If you have any further questions please do not hesitate to contact the doctors/nurses at the GUM clinic or the research team (contact details below).

What if there is a problem?
Any complaints about the way you have been dealt with during the study will be addressed. Please do contact the clinic or a member of the research team (contact details below), who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this by contacting the Independent Complaint Advocacy Service (ICAS) on 0845 600 8616.

It is very unlikely that anything will go wrong, however in the event that something does go wrong and you are harmed during the research and this is due to someone's negligence then you may have grounds for legal action for compensation against the University of Southampton but you may have to pay legal costs. The normal National Health Service complaints will still be available to you (if appropriate).

Contacts for Further Information

Researcher:
Miss Clare Labiran
University of Southampton, School of Medicine, Mailpoint 814, Level C, South Block, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, UK
Tel: 02380 798886

Head of research governance, University of Southampton:
Dr Martina Prude
University of Southampton, University road, Room 4055, George Thomas Building 37, Southampton, SO17 1BJ, UK
Tel: 02380 595058

APPENDIX 8 : CONSENT FORM FOR SOUTHAMPTON STUDY (CHP 5)

Medicine

UNIVERSITY OF Southampton

Mailpoint 814, Level C, South Block,
Southampton General Hospital,
Southampton,
SO16 6YD,
Tel:02380798886
Fax: 02380796992

Centre Number:
Study Number:
Patient Identification Number for this study:

CONSENT FORM

Title of Study: Study on the Epidemiology of Chlamydia in Southampton

Name of Researcher: Miss Clare Labiran

Please **initial** box

1. I confirm that I have read and understand the information sheet dated 20th June 2012 (Version 2) for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected. ☐

3. I give permission for my medical records to be accessed by representatives from regulatory authorities or from the NHS Trust and for the researchers to obtain anonymised data from my medical records for this study. ☐

4. I give permission for my sample to be stored as a scientific resource in a Human Tissue Act licensed Tissue Bank. ☐

5. I give permission for my sample to be used for future ethically approved research. ☐

6. I know how to contact the research team if I need to. ☐

7. I agree to take part in the above study. ☐

Name of Patient	Date	Signature
Name of Person taking consent	Date	Signature

When completed, 1 for patient; 1 for GUM clinic file.

Chlamydia in Southampton Consent form
Version 2.0, 20th June 2012

APPENDIX 9 : QIAMP INFORMATION SHEET

For purification of viral nucleic acids for in vitro diagnostics in Europe, the QIAamp DSP Virus Kit is CE-IVD-marked, compliant with EU Directive 98/79/EC. This kit is not available in the USA and Canada.

Purification of DNA from urine

For preparation of cellular, bacterial, or viral DNA from urine, the QIAamp Viral RNA Mini Kit is recommended. Buffer AVL supplied with this kit is optimized to inactivate the numerous PCR inhibitors found in urine.

Purification of DNA from stool

The QIAamp DNA Stool Mini Kit is recommended for preparation of DNA from stool. Stool samples typically contain many compounds that can degrade DNA and inhibit downstream enzymatic reactions. The QIAamp DNA Stool Mini Kit removes these substances through the action of a proprietary reagent that efficiently adsorbs inhibitors, together with a lysis buffer that provides optimized conditions for inhibitor removal. DNA purification using the QIAamp DNA Stool Mini Kit can be fully automated on the QIAcube for increased standardization and ease of use.

QIAamp DNA Mini or QIAamp DNA Blood Mini Kits can also be used to purify viral DNA from stool, but removal of inhibitors is not as effective. See Appendix F, page 58.

Purification of DNA from formalin-fixed, paraffin-embedded tissues

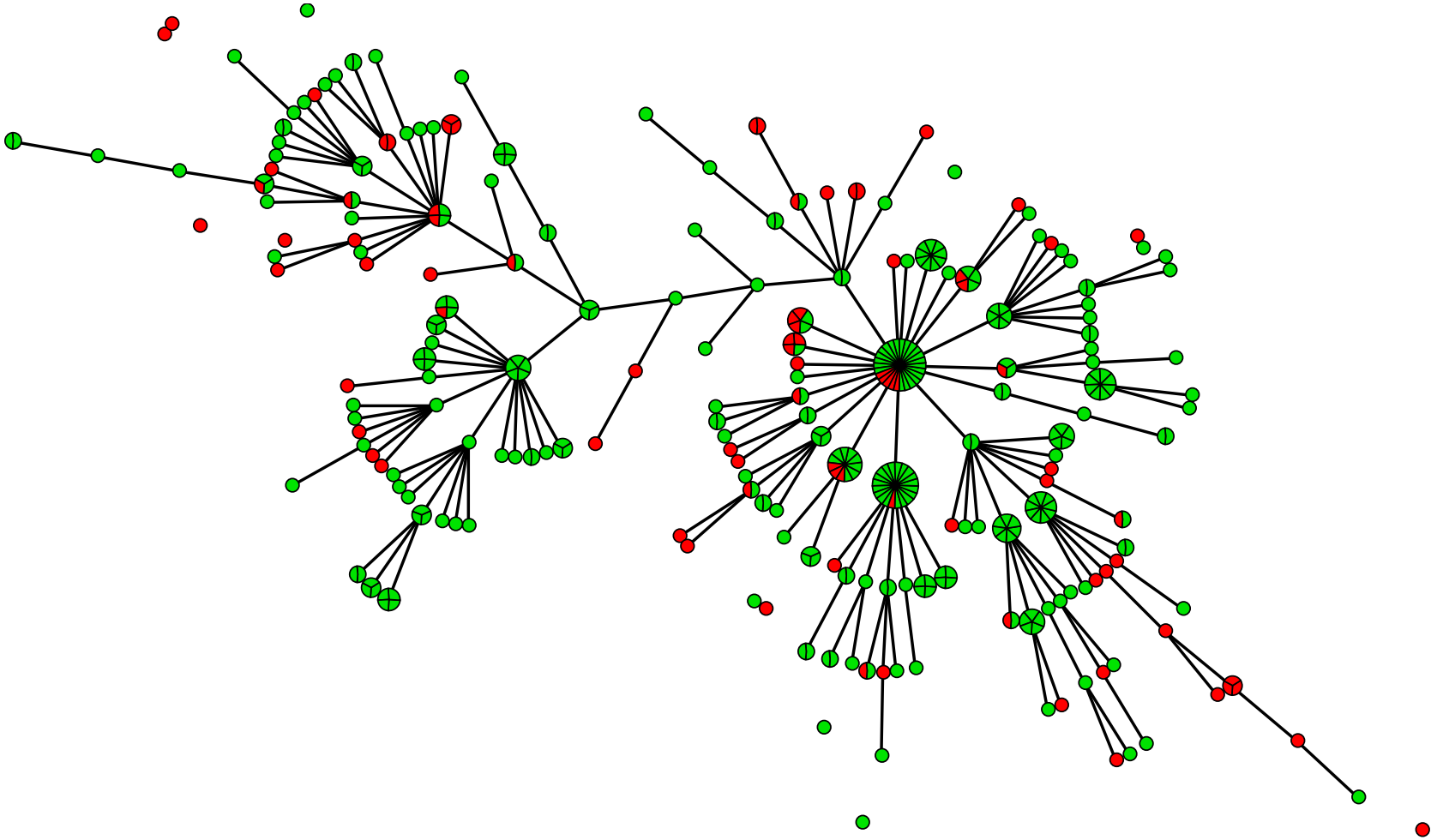
The QIAamp DNA FFPE Tissue Kit is recommended for purification of DNA from formalin-fixed, paraffin-embedded (FFPE) tissues. The kit combines the selective binding properties of a silica-based membrane with flexible elution volumes of between 20 and 100 µl. Specially optimized lysis conditions allow genomic DNA to be efficiently purified from FFPE tissue sections without the need for overnight incubation.

The QIAamp DNA Mini Kit can also be used for purification of DNA from FFPE tissues that have first been processed as described in Appendix C, page 53.

Purification of DNA from forensic and human identity samples

The QIAamp DNA Investigator Kit is recommended for purification of total (genomic and mitochondrial) DNA from a wide range of forensic and human identity samples, such as casework or crime-scene samples, dried blood, bone, and sexual assault samples, swabs, and filters. Purification is fast and efficient, and purified DNA performs well in downstream analyses, such as quantitative PCR and STR analysis, with high signal-to-noise ratios. The procedure is designed to ensure that there is no sample-to-sample cross-contamination. Purification of DNA using the QIAamp DNA Investigator Kit can be automated on the QIAcube.

APPENDIX 10: MINIMUM SPANNING TREE OF SAMPLES COLLECTED IN BRIGHTON (RED) AND SAMPLES COLLECTED IN SOUTHAMPTON (GREEN)



REFERENCES

ABDELRAHMAN, Y. M. & BELLAND, R. J. 2005. The chlamydial developmental cycle. *FEMS Microbiol Rev,* 29**,** 949-59.

AMANN, R. I., LUDWIG, W. & SCHLEIFER, K. H. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev,* 59**,** 143-69.

ANDERSEN, A. A. & ROGERS, D. G. Resistance to tetracycline and sulfadiazine in swine *C. trachomatis* isolatyes. Proceedings of the Ninth International Symposium on Human Chlamydial Infections, International *Chlamydia* Symposium, 1998. 313-316.

AUFM KAMPE, H. J., WEICKMANN, H. K. & KELLY, J. J. 1951. The influence of temperature on the shape of ice crystals growing at water saturation. *Journal of meterology,* 8.

BANNANTINE, J. P., STAMM, W. E., SUCHLAND, R. J. & ROCKEY, D. D. 1998. Chlamydia trachomatis IncA is localized to the inclusion membrane and is recognized by antisera from infected humans and primates. *Infect Immun,* 66**,** 6017-21.

BARDAKCI, F. 2000. Random Amplified Polymorphic DNA (RAPD) Markers. *Turkisk Journal of Biology,* 25**,** 185-196.

BARNES, R. C., ROMPALO, A. M. & STAMM, W. E. 1987. Comparison of Chlamydia trachomatis serovars causing rectal and cervical infections. *J Infect Dis,* 156**,** 953-8.

BARNES, R. C., WANG, S. P., KUO, C. C. & STAMM, W. E. 1985. Rapid immunotyping of Chlamydia trachomatis with monoclonal antibodies in a solid-phase enzyme immunoassay. *J Clin Microbiol,* 22**,** 609-13.

BATTEIGER, B. E., TU, W., OFNER, S., VAN DER POL, B., STOTHARD, D. R., ORR, D. P., KATZ, B. P. & FORTENBERRY, J. D. 2010. Repeated Chlamydia trachomatis genital infections in adolescent women. *J Infect Dis,* 201**,** 42-51.

BEATTY, W. L., MORRISON, R. P. & BYRNE, G. I. 1994. Persistent chlamydiae: from cell culture to a paradigm for chlamydial pathogenesis. *Microbiol Rev,* 58**,** 686-99.

BELLAND, R. J., SCIDMORE, M. A., CRANE, D. D., HOGAN, D. M., WHITMIRE, W., MCCLARTY, G. & CALDWELL, H. D. 2001. Chlamydia trachomatis cytotoxicity associated with complete and partial cytotoxin genes. *Proc Natl Acad Sci U S A,* 98**,** 13984-9.

BLACK, C. M. 1997. Current methods of laboratory diagnosis of Chlamydia trachomatis infections. *Clin Microbiol Rev,* 10**,** 160-84.

BOISVERT, J. F., KOUTSKY, L. A., SUCHLAND, R. J. & STAMM, W. E. 1999. Clinical features of Chlamydia trachomatis rectal infection by serovar among homosexually active men. *Sex Transm Dis,* 26**,** 392-8.

BOM, R. J., CHRISTERSON, L., SCHIM VAN DER LOEFF, M. F., COUTINHO, R. A., HERRMANN, B. & BRUISTEN, S. M. 2011. Evaluation of high-resolution typing methods for Chlamydia trachomatis in samples from heterosexual couples. *Journal of Clinical Microbiology,* 49**,** 2844-2853.

BOM, R. J., MATSER, A., BRUISTEN, S. M., VAN ROOIJEN, M. S., HEIJMAN, T., MORRE, S. A., DE VRIES, H. J. & VAN DER LOEFF, M. F. 2013a. Multilocus sequence typing of Chlamydia trachomatis among men who have sex with men reveals co-circulating strains not associated with specific subpopulations. *J Infect Dis*.

BOM, R. J., VAN DER HELM, J. J., SCHIM VAN DER LOEFF, M. F., VAN ROOIJEN, M. S., HEIJMAN, T., MATSER, A., DE VRIES, H. J. & BRUISTEN, S. M. 2013b. Distinct transmission networks of Chlamydia trachomatis in men who have sex with men and heterosexual adults in Amsterdam, The Netherlands. *Plos One,* 8**,** e53869.

BROWNE, K. & LIM, J. 2010. Trans lives in the 'gay capital of the UK'. *Gender, Place and Culture: A Journal of Feminist Geography,* 17**,** 615-633.

BRUNHAM, R. C. & REY-LADINO, J. 2005. Immunology of Chlamydia infection: implications for a Chlamydia trachomatis vaccine. *Nat Rev Immunol,* 5**,** 149-61.

BYRD, T. F. & HORWITZ, M. A. 1993. Regulation of transferrin receptor expression and ferritin content in human mononuclear phagocytes. Coordinate upregulation by iron transferrin and downregulation by interferon gamma. *J Clin Invest,* 91**,** 969-76.

BYRNE, G. I. & OJCIUS, D. M. 2004. Chlamydia and apoptosis: life and death decisions of an intracellular pathogen. *Nat Rev Microbiol,* 2**,** 802-8.

CDC. 2011. *Chlamydia* [Online]. Available: http://www.cdc.gov/std/stats10/chlamydia.htm.

CHEN, B., STOUT, R. & CAMPBELL, W. F. 1996. Nitric oxide production: a mechanism of Chlamydia trachomatis inhibition in interferon-gamma-treated RAW264.7 cells. *FEMS Immunol Med Microbiol,* 14**,** 109-20.

CHERNESKY, M. 2005. The laboratory diagnosis of Chlamydia trachomatis infections. *Canadian STI best practice laboratory guidelines,* 16**,** 39-44.

CHRISTERSON, L., BOM, R. J., BRUISTEN, S. M., YASS, R., HARDICK, J., BRATT, G., GAYDOS, C. A., MORRE, S. A. & HERRMANN, B. 2012. Chlamydia trachomatis strains show specific clustering for men who have sex with men compared to heterosexual transmission in Sweden, the Netherlands and the United States. *J Clin Microbiol*.

CHRISTERSON, L., DE VRIES, H. J., DE BARBEYRAC, B., GAYDOS, C. A., HENRICH, B., HOFFMANN, S., SCHACHTER, J., THORVALDSEN, J., VALL-MAYANS, M., KLINT, M., HERRMANN, B. & MORRE, S. A. 2010. Typing of lymphogranuloma venereum Chlamydia trachomatis strains. *Emerg Infect Dis,* 16**,** 1777-9.

CHRISTERSON, L., DE VRIES, H. J., KLINT, M., HERRMANN, B. & MORRE, S. A. 2011. Multilocus sequence typing of urogenital Chlamydia trachomatis from patients with different degrees of clinical symptoms. *Sex Transm.Dis.,* 38**,** 490-494.

CLARKE, S. C. 2002. Nucleotide sequence-based typing of bacteria and the impact of automation. *Bioessays,* 24**,** 858-62.

CREIGHTON, S., TENANT-FLOWERS, M., TAYLOR, C. B., MILLER, R. & LOW, N. 2003. Co-infection with gonorrhoea and chlamydia: how much is there and what does it mean? *Int J STD AIDS,* 14**,** 109-13.

CRESSWELL, P., ACKERMAN, A. L., GIODINI, A., PEAPER, D. R. & WEARSCH, P. A. 2005. Mechanisms of MHC class I-restricted antigen processing and cross-presentation. *Immunol Rev,* 207**,** 145-57.

CROUCHER, N. J., HARRIS, S. R., FRASER, C., QUAIL, M. A., BURTON, J., VAN DER LINDEN, M., MCGEE, L., VON GOTTBERG, A., SONG, J. H., KO, K. S., PICHON, B., BAKER, S., PARRY, C. M., LAMBERTSEN, L. M., SHAHINAS, D., PILLAI, D. R., MITCHELL, T. J., DOUGAN, G., TOMASZ, A., KLUGMAN, K. P., PARKHILL, J., HANAGE, W. P. & BENTLEY, S. D. 2011. Rapid pneumococcal evolution in response to clinical interventions. *Science,* 331**,** 430-4.

CROY, T. R., KUO, C. C. & WANG, S. P. 1975. Comparative susceptibility of eleven mammalian cell lines to infection with trachoma organisms. *J Clin Microbiol,* 1**,** 434-9.

DAROUGAR, S., FORSEY, T., BREWERTON, D. A. & ROGERS, K. L. 1980. Prevalence of antichlamydial antibody in London blood donors. *Br J Vener Dis,* 56**,** 404-7.

DARVILLE, T., O'NEILL, J. M., ANDREWS, C. W., JR., NAGARAJAN, U. M., STAHL, L. & OJCIUS, D. M. 2003. Toll-like receptor-2, but not Toll-like receptor-4, is essential for development of oviduct pathology in chlamydial genital tract infection. *J Immunol,* 171**,** 6187-97.

DEAN, D., BRUNO, W. J., WAN, R., GOMES, J. P., DEVIGNOT, S., MEHARI, T., DE VRIES, H. J. C., MORRE, S. A., MYERS, G., READ, T. D. & SPRATT, B. G. 2009. Predicting Phenotype and Emerging Strains among Chlamydia trachomatis Infections. *Emerging Infectious Diseases,* 15**,** 1385-1394.

DESSUS-BABUS, S., DARVILLE, T. L., CUOZZO, F. P., FERGUSON, K. & WYRICK, P. B. 2002. Differences in innate immune responses (in vitro) to HeLa cells infected with nondisseminating serovar E and disseminating serovar L2 of Chlamydia trachomatis. *Infect Immun,* 70**,** 3234-48.

DUECK, D., ABHA, C, MOORE, C, SARICH, J, GOODRIDGE, D, HECKERMAN, SAYER, D, MALLAL, S, JOJIC, N 2009. Population Sequencing from Chromatogram Data. *In:* TORONTO, U. O. (ed.).

DUMONT, F., MARECHAL, P. A. & GERVAIS, P. 2006. Involvement of two specific causes of cell mortality in freeze-thaw cycles with freezing to -196 degrees C. *Appl Environ Microbiol,* 72**,** 1330-5.

DWYER, R. S., TREHARNE, J. D., JONES, B. R. & HERRING, J. 1972. Chlamydial infection. Results of micro-immunofluorescence tests for the detection of type-specific antibody in certain Chlamydial infections. *Br J Vener Dis,* 48**,** 452-459.

EB, F., ORFILA, J. & LEFEBVRE, J. F. 1976. Ultrastructural study of the development of the agent of ewe's abortion. *J Ultrastruct Res,* 56**,** 177-85.

ECDC 2011. Sexually transmitted infections in Europe, 1990–2009. Stockholm: ECDC.

ECDC 2012. Annual Epidemiological Report 2012.: ECDC.

EISSENBERG, L. G. & WYRICK, P. B. 1981. Inhibition of phagolysosome fusion is localized to Chlamydia psittaci-laden vacuoles. *Infect Immun,* 32**,** 889-96.

EMRE, U., SOKOLOVSKAYA, N., ROBLIN, P. M., SCHACHTER, J. & HAMMERSCHLAG, M. R. 1995. Detection of anti-Chlamydia pneumoniae IgE in children with reactive airway disease. *J Infect Dis,* 172**,** 265-7.

ENTRICAN, G., BUXTON, D. & LONGBOTTOM, D. 2001. Chlamydial infection in sheep: immune control versus fetal pathology. *J R Soc Med,* 94**,** 273-7.

EVERETT, K. D. & HATCH, T. P. 1995. Architecture of the cell envelope of Chlamydia psittaci 6BC. *J Bacteriol,* 177**,** 877-82.

EVERETT, K. D. E., BUSH, R. M. & ANDERSEN, A. A. 1999. Emended description of the order Chlamydiales , proposal of Parachlamydiaceae fam. nov. and Simkaniaceae fam. nov., each containing one monotypic genus, revised taxonomy of the family Chlamydiaceae , including a new genus and five new species, and standards for the identification of organisms. *International Journal of Systematic Bacteriology,* 49**,** 415-440.

FALK, L., LINDBERG, M., JURSTRAND, M., BACKMAN, A., OLCEN, P. & FREDLUND, H. 2003. Genotyping of Chlamydia trachomatis would improve contact tracing. *Sex Transm Dis,* 30**,** 205-10.

FEIL, E. J., LI, B. C., AANENSEN, D. M., HANAGE, W. P. & SPRATT, B. G. 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol,* 186**,** 1518-30.

FIELDS, K. A., FISCHER, E. & HACKSTADT, T. 2002. Inhibition of fusion of Chlamydia trachomatis inclusions at 32 degrees C correlates with restricted export of IncA. *Infect Immun,* 70**,** 3816-23.

FONG, I. W., CHIU, B., VIIRA, E., FONG, M. W., JANG, D. & MAHONY, J. 1997. Rabbit model for Chlamydia pneumoniae infection. *J Clin Microbiol,* 35**,** 48-52.

FORSEY, T. & DAROUGAR, S. 1984. Acute conjunctivitis caused by an atypical chlamydial strain: Chlamydia IOL 207. *Br J Ophthalmol,* 68**,** 409-11.

FREIDANK, H. M., BILLING, H. & WIEDMANN-AL-AHMAD, M. 2001. Influence of iron restriction on Chlamydia pneumoniae and C. trachomatis. *J Med Microbiol,* 50**,** 223-7.

FROST, E. H., DESLANDES, S., VEILLEUX, S. & BOURGAUX-RAMOISY, D. 1991. Typing Chlamydia trachomatis by detection of restriction fragment length polymorphism in the gene encoding the major outer membrane protein. *Journal of Infectious Diseases,* 163**,** 1103-1107.

FUKUSHI, H. & HIRAI, K. 1992. Proposal of Chlamydia pecorum sp. nov. for Chlamydia strains derived from ruminants. *Int J Syst Bacteriol,* 42**,** 306-8.

GACZYNSKA, M., ROCK, K. L. & GOLDBERG, A. L. 1993. Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature,* 365**,** 264-7.

GEISLER, W. M. 2010. Duration of untreated, uncomplicated Chlamydia trachomatis genital infection and factors associated with chlamydia resolution: a review of human studies. *J Infect Dis,* 201 Suppl 2**,** S104-13.

GEISLER, W. M., MORRISON, S. G. & BACHMANN, L. H. 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**,** 856-8.

GEISLER, W. M., SUCHLAND, R. J. & STAMM, W. E. 2006. Association of Chlamydia trachomatis Serovar Ia infection with black race in a sexually transmitted diseases clinic patient population in Birmingham, Alabama. *Sex Transm Dis,* 33**,** 621-4.

GEISLER, W. M., WHITTINGTON, W. L., SUCHLAND, R. J. & STAMM, W. E. 2002. Epidemiology of anorectal chlamydial and gonococcal infections among men having sex with men in Seattle: utilizing serovar and auxotype strain typing. *Sex Transm Dis,* 29**,** 189-95.

GEWIRTZMAN, A., BOBRICK, L., CONNER, K. & TYRING, S. K. 2011. Epidemiology of sexually Transmitted Infections. *In:* GROSS, G. & TYRING, S. K. (eds.) *Sexually Transmitted Infections and Sexually Transmitted Diseases.* Berlin: Springer-Verlag.

GHANI, A. C., ISON, C. A., WARD, H., GARNETT, G. P., BELL, G., KINGHORN, G. R., WEBER, J. & DAY, S. 1996. Sexual partner networks in the transmission of sexually transmitted diseases. An analysis of gonorrhea cases in Sheffield, UK. *Sex Transm Dis,* 23**,** 498-503.

GOLDMEIER, D. & DAROUGAR, S. 1977. Isolation of Chlamydia trachomatis from throat and rectum of homosexual men. *Br J Vener Dis,* 53**,** 184-5.

GOMES, J. O. P. 2012. Genomic stability of genotyping markers in Chlamydia trachomatis. *Frontiers in Cellular and Infection Microbiology,* 2.

GORDON, F. B., HARPER, I. A., QUAN, A. L., TREHARNE, J. D., DWYER, R. S. & GARLAND, J. A. 1969. Detection of Chlamydia (Bedsonia) in certain infections of man. I. Laboratory procedures: comparison of yolk sac and cell culture for detection and isolation. *J Infect Dis,* 120**,** 451-62.

GRAYSTON, J. T., KUO, C. C., CAMPBELL, L. A. & WANG, S. P. 1989. *Chlamydia* pneumoniae sp. nov. for *Chlamydia* sp. strain TWAR. *International Journal of Systematic Bacteriology,* 39**,** 88-90.

GRAYSTON, J. T., WOOLRIDGE, R. L., WANG, S. P., YEN, C. H., YANG, C. Y., CHENG, K. H. & CHANG, I. H. 1963. Field studies of protection from infection by experimental trachoma virus vaccine in preschool-aged children on Taiwan. *Proc Soc Exp Biol Med,* 112**,** 589-95.

GRIMWOOD, J., OLINGER, L. & STEPHENS, R. S. 2001. Expression of Chlamydia pneumoniae polymorphic membrane protein family genes. *Infect Immun,* 69**,** 2383-9.

GRIMWOOD, J. & STEPHENS, R. S. 1999. Computational analysis of the polymorphic membrane protein superfamily of Chlamydia trachomatis and Chlamydia pneumoniae. *Microb Comp Genomics,* 4**,** 187-201.

GUPTA, S., LEATHAM, E. W., CARRINGTON, D., MENDALL, M. A., KASKI, J. C. & CAMM, A. J. 1997. Elevated Chlamydia pneumoniae antibodies, cardiovascular events, and azithromycin in male survivors of myocardial infarction. *Circulation,* 96**,** 404-7.

GUPTA, S. L., CARLIN, J. M., PYATI, P., DAI, W., PFEFFERKORN, E. R. & MURPHY, M. J., JR. 1994. Antiparasitic and antiproliferative effects of indoleamine 2,3-dioxygenase enzyme expression in human fibroblasts. *Infect Immun,* 62**,** 2277-84.

HACKSTADT, T. 1999. *Chlamydia*: Intracellular Biology, Pathogenesis, and Immunity. *In:* STEPHENS, R. S. (ed.) *Cell Biology.* Washington, DC.: ASM Press.

HAFNER, L. M. & TIMMS, P. 2005. Persistent chlamydial infections: role in inflammation and challenges for vaccine development. *Mucosal Immunology,* 13**,** 4-7.

HAHN, D. L., ANTTILA, T. & SAIKKU, P. 1996. Association of Chlamydia pneumoniae IgA antibodies with recently symptomatic asthma. *Epidemiol Infect,* 117**,** 513-7.

HALBERSTAEDTER.L & VON PROWAZEK, S. 1907. *Arb Gesundhamt,* 26**,** 44-47.

HARRIS, S. R., CLARKE, I. N., SETH-SMITH, H. M., SOLOMON, A. W., CUTCLIFFE, L. T., MARSH, P., SKILTON, R. J., HOLLAND, M. J., MABEY, D., PEELING, R. W., LEWIS, D. A., SPRATT, B. G., UNEMO, M., PERSSON, K., BJARTLING, C., BRUNHAM, R., DE VRIES, H. J., MORRE, S. A., SPEKSNIJDER, A., BEBEAR, C. M., CLERC, M., DE BARBEYRAC, B., PARKHILL, J. & THOMSON, N. R. 2012. Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet*.

HERIDA, M., SEDNAOUI, P., COUTURIER, E., NEAU, D., CLERC, M., SCIEUX, C., KREPLAK, G., GOULET, V., HAMERS, F. F. & DE BARBEYRAC, B. 2005. Rectal lymphogranuloma venereum, France. *Emerg Infect Dis,* 11**,** 505-6.

HERRMANN, B., TORNER, A., LOW, N., KLINT, M., NILSSON, A., VELICKO, I., SODERBLOM, T. & BLAXHULT, A. 2008. Emergence and spread of Chlamydia trachomatis variant, Sweden. *Emerging Infectious Diseases,* 14**,** 1462-1465.

HEYMANS, R., A, A. M., BRUISTEN, S. M., HEIJMAN, T., GESKUS, R. B., SPEKSNIJDER, A. G., DAVIDOVICH, U., DE VRIES, H. J., COUTINHO, R. A. & SCHIM VAN DER LOEFF, M. F. 2012. Distinct Neisseria gonorrhoeae transmission networks among men who have sex with men in Amsterdam, The Netherlands. *J Infect Dis,* 206**,** 596-605.

HEYMER, K. J. & WILSON, D. P. 2010. Available evidence does not support serosorting as an HIV risk reduction strategy. *AIDS,* 24**,** 935-6; author reply 936-8.

HODINKA, R. L., DAVIS, C. H., CHOONG, J. & WYRICK, P. B. 1988. Ultrastructural study of endocytosis of Chlamydia trachomatis by McCoy cells. *Infect Immun,* 56**,** 1456-63.

HPA 2011. Sexually transmitted infections in men who have sex with men in the UK: 2011 report.

HPA 2012a. Health Protection Report. London: Health Protection Agency.

HPA 2012b. HIV in the United Kingdom: 2012 Report. London: Health Protection Services, Colindale.

HPA. 2013. *CTAD: Chlamydia Testing Activity Dataset* [Online]. Available: http://www.hpa.org.uk/sexualhealth/ctad.

HSIA, R., OHAYON, H., GOUNON, P., DAUTRY-VARSAT, A. & BAVOIL, P. M. 2000. Phage infection of the obligate intracellular bacterium, Chlamydia psittaci strain guinea pig inclusion conjunctivitis. *Microbes Infect,* 2**,** 761-72.

HUNTER, P. R. & GASTON, M. A. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol,* 26**,** 2465-6.

IGIETSEME, J. U., URIRI, I. M., CHOW, M., ABE, E. & RANK, R. G. 1997. Inhibition of intracellular multiplication of human strains of Chlamydia trachomatis by nitric oxide. *Biochem Biophys Res Commun,* 232**,** 595-601.

ILLUMINA. 2012. *Chemistry for Next-Generation Sequencing* [Online]. Available: http://www.illumina.com/technology/sequencing_technology.ilmn.

ISMAIL, M. A., MOAWAD, A. H., POON, E. & HENDERSON, C. 1987. Role of Chlamydia trachomatis in postpartum endometritis. *J Reprod Med,* 32**,** 280-4.

JIN, F., PRESTAGE, G. P., TEMPLETON, D. J., POYNTEN, I. M., DONOVAN, B., ZABLOTSKA, I., KIPPAX, S. C., MINDEL, A. & GRULICH, A. E. 2012. The impact of HIV seroadaptive behaviors on

sexually transmissible infections in HIV-negative homosexual men in Sydney, Australia. *Sex Transm Dis,* 39**,** 191-4.

JOHNSON, R. M. 2004. Murine oviduct epithelial cell cytokine responses to Chlamydia muridarum infection include interleukin-12-p70 secretion. *Infect Immun,* 72**,** 3951-60.

JONES, R. B., KATZ, B. P., VAN DER POL, B., CAINE, V. A., BATTEIGER, B. E. & NEWHALL, W. J. 1986. Effect of blind passage and multiple sampling on recovery of Chlamydia trachomatis from urogenital specimens. *J Clin Microbiol,* 24**,** 1029-33.

JOSEPH, S. J., DIDELOT, X., ROTHSCHILD, J., DE VRIES, H. J., MORRE, S. A., READ, T. D. & DEAN, D. 2012. Population genomics of Chlamydia trachomatis: insights on drift, selection, recombination, and population structure. *Mol Biol Evol,* 29**,** 3933-46.

JURSTRAND, M., CHRISTERSON, L., KLINT, M., FREDLUND, H., UNEMO, M. & HERRMANN, B. 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.

JURSTRAND, M., FALK, L., FREDLUND, H., LINDBERG, M., OLCEN, P., ANDERSSON, S., PERSSON, K., ALBERT, J. & BACKMAN, A. 2001. Characterization of Chlamydia trachomatis omp1 genotypes among sexually transmitted disease patients in Sweden. *J Clin Microbiol,* 39**,** 3915-9.

KALWIJ, S., MACINTOSH, M. & BARAITSER, P. 2010. Screening and treatment of Chlamydia trachomatis infections. *BMJ,* 340**,** c1915.

KARUNAKARAN, K. P., YU, H., FOSTER, L. J. & BRUNHAM, R. C. 2010. Development of a Chlamydia trachomatis T cell Vaccine. *Hum Vaccin,* 6**,** 676-80.

KENNEDY, C., BERNARD, L. J., MUESSIG, K. E., KONDA, K. A., AKL, E. A., LO, Y., GERBASE, A. & O'REILLY, K. R. 2013. Serosorting and HIV/STI Infection among HIV-Negative MSM and Transgender People: A Systematic Review and Meta-Analysis to Inform WHO Guidelines. *Journal of Sexually Transmitted Diseases,* 13.

KLAUSNER, J. D., MCFARLAND, W., BOLAN, G., HERNANDEZ, M. T., MOLITOR, F., LEMP, G. F., CAHOON-YOUNG, B., MORROW, S. & RUIZ, J. 2001. Knock-knock: a population-based survey of risk behavior, health care access, and Chlamydia trachomatis infection among low-income women in the San Francisco Bay area. *J Infect Dis,* 183**,** 1087-92.

KLINT, M., FUXELIUS, H. H., GOLDKUHL, R. R., SKARIN, H., RUTEMARK, C., ANDERSSON, S. G. E., PERSSON, K. & HERRMANN, B. 2007. High-resolution genotyping of Chlamydia trachomatis strains by multilocus sequence analysis. *Journal of Clinical Microbiology,* 45**,** 1410-1414.

KLINT, M., HADAD, R., CHRISTERSON, L., LOR,, B., ANAGRIUS, C., OSTERLUND, A., LARSSON, I., SYLVAN, S., FREDLUND, H., UNEMO, M. & HERRMANN, B. 2011. Prevalence trends in Sweden for the new variant of Chlamydia trachomatis. *Clin Microbiol Infect,* 17**,** 683-689.

KOMMEDAL, O., KVELLO, K., SKJASTAD, R., LANGELAND, N. & WIKER, H. G. 2009. Direct 16S rRNA gene sequencing from clinical specimens, with special focus on polybacterial samples and interpretation of mixed DNA chromatograms. *J Clin Microbiol,* 47**,** 3562-8.

KROPP, R. Y. & WONG, T. 2005. Emergence of lymphogranuloma venereum in Canada. *CMAJ,* 172**,** 1674-6.

KUO, C. C., GOWN, A. M., BENDITT, E. P. & GRAYSTON, J. T. 1993a. Detection of Chlamydia pneumoniae in aortic lesions of atherosclerosis by immunocytochemical stain. *Arterioscler Thromb,* 13**,** 1501-4.

KUO, C. C., GRAYSTON, J. T., CAMPBELL, L. A., GOO, Y. A., WISSLER, R. W. & BENDITT, E. P. 1995. Chlamydia pneumoniae (TWAR) in coronary arteries of young adults (15-34 years old). *Proc Natl Acad Sci U S A,* 92**,** 6911-4.

KUO, C. C., SHOR, A., CAMPBELL, L. A., FUKUSHI, H., PATTON, D. L. & GRAYSTON, J. T. 1993b. Demonstration of Chlamydia pneumoniae in atherosclerotic lesions of coronary arteries. *J Infect Dis,* 167**,** 841-9.

KUTLIN, A., ROBLIN, P. M. & HAMMERSCHLAG, M. R. 1999. In vitro activities of azithromycin and ofloxacin against Chlamydia pneumoniae in a continuous-infection model. *Antimicrob Agents Chemother,* 43**,** 2268-72.

LABIRAN, C., CLARKE, I. N., CUTCLIFFE, L. T., WANG, Y., SKILTON, R. J., PERSSON, K., BJARTLING, C., HERRMANN, B., CHRISTERSON, L. & MARSH, P. 2012a. Genotyping markers used for multi locus VNTR analysis with ompA (MLVA-ompA) and multi sequence typing (MST) retain stability in Chlamydia trachomatis. *Front Cell Infect Microbiol,* 2**,** 68.

LABIRAN, C., CLARKE, I. N., CUTCLIFFE, L. T., WANG, Y., SKILTON, R. J., PERSSON, K., BJARTLING, C., HERRMANN, B. R., CHRISTERSON, L. & MARSH, P. 2012b. Genotyping markers used for multi locus VNTR analysis with ompA (MLVA-ompA) and multi sequence typing retain stability in Chlamydia trachomatis. *Frontiers in Cellular and Infection Microbiology,* 2.

LAN, J., OSSEWAARDE, J. M., WALBOOMERS, J. M., MEIJER, C. J. & VAN DEN BRULE, A. J. 1994. Improved PCR sensitivity for direct genotyping of Chlamydia trachomatis serovars by using a nested PCR. *J Clin Microbiol,* 32**,** 528-30.

LAURILA, A. L., ANTTILA, T., LAARA, E., BLOIGU, A., VIRTAMO, J., ALBANES, D., LEINONEN, M. & SAIKKU, P. 1997. Serological evidence of an association between Chlamydia pneumoniae infection and lung cancer. *Int J Cancer,* 74**,** 31-4.

LI, J. H., CAI, Y. M., YIN, Y. P., HONG, F. C., SHI, M. Q., FENG, T. J., PENG, R. R., WANG, B. X. & CHEN, X. S. 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**,** 143-6.

LINDSTEDT, B. A. 2005. Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis,* 26**,** 2567-2582.

LONGBOTTOM, D. 2003. Chlamydial vaccine development. *J Med Microbiol,* 52**,** 537-40.

LONGBOTTOM, D. & LIVINGSTONE, M. 2006. Vaccination against chlamydial infections of man and animals. *Vet J,* 171**,** 263-75.

LYSEN, M., OSTERLUND, A., RUBIN, C. J., PERSSON, T., PERSSON, I. & HERRMANN, B. 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. *Journal of Clinical Microbiology,* 42**,** 1641-1647.

MAASS, M. & DALHOFF, K. 1995. Transport and storage conditions for cultural recovery of Chlamydia pneumoniae. *J Clin Microbiol,* 33**,** 1793-6.

MAHONY, J. B. & CHERNESKY, M. A. 1985. Effect of swab type and storage temperature on the isolation of Chlamydia trachomatis from clinical specimens. *Journal of Clinical Microbiology,* 22**,** 865-867.

MAIDEN, M. C. 2006. Multilocus sequence typing of bacteria. *Annu Rev Microbiol,* 60**,** 561-88.

MAIDEN, M. C., BYGRAVES, J. A., FEIL, E., MORELLI, G., RUSSELL, J. E., URWIN, R., ZHANG, Q., ZHOU, J., ZURTH, K., CAUGANT, D. A., FEAVERS, I. M., ACHTMAN, M. & SPRATT, B. G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A,* 95**,** 3140-5.

MANIRE, G. P. 1966. Structure of purified cell walls of dense forms of meningopneumonitis organisms. *J Bacteriol,* 91**,** 409-13.

MARANGONI, A., FOSCHI, C., NARDINI, P., D'ANTUONO, A., BANZOLA, N., DI FRANCESCO, A., OSTANELLO, F., RUSSO, I., DONATI, M. & CEVENINI, R. 2012. Chlamydia trachomatis serovar distribution and other sexually transmitted coinfections in subjects attending an STD outpatients clinic in Italy. *New Microbiol,* 35**,** 215-9.

MATSUMOTO, A. 1973. Fine structures of cell envelopes of Chlamydia organisms as revealed by freeze-etching and negative staining techniques. *J Bacteriol,* 116**,** 1355-63.

MATSUMOTO, A. 1982. Surface projections of Chlamydia psittaci elementary bodies as revealed by freeze-deep-etching. *J Bacteriol,* 151**,** 1040-2.

MATSUMOTO, A., BESSHO, H., UEHIRA, K. & SUDA, T. 1991. Morphological studies of the association of mitochondria with chlamydial inclusions and the fusion of chlamydial inclusions. *J Electron Microsc (Tokyo),* 40**,** 356-63.

MAYER, K. H. 2011. Sexually transmitted diseases in men who have sex with men. *Clin Infect Dis,* 53 Suppl 3**,** S79-83.

MCMILLAN, A., SOMMERVILLE, R. G. & MCKIE, P. M. 1981. Chlamydial infection in homosexual men. Frequency of isolation of Chlamydia trachomatis from the urethra, ano-rectum, and pharynx. *Br J Vener Dis,* 57**,** 47-9.

MEJUTO, P. B., JA.; JUNQUERA, M.; TORREBLANCA, A.; LEIVA, P. S 2013. Genotyping *Chlamydia trachomatis* strains among men who have sex with men from a Northern Spain region: a cohort study. *BMJ Open,* 3.

MEYER, T., ARNDT, R., VON KROSIGK, A. & PLETTENBERG, A. 2005. Repeated detection of lymphogranuloma venereum caused by Chlamydia trachomatis L2 in homosexual men in Hamburg. *Sex Transm Infect,* 81**,** 91-2.

MEYERS, D. S., HALVORSON, H. & LUCKHAUPT, S. 2007. Screening for chlamydial infection: an evidence update for the U.S. Preventive Services Task Force. *Ann Intern Med,* 147**,** 135-42.

MISAGHI, S., BALSARA, Z. R., CATIC, A., SPOONER, E., PLOEGH, H. L. & STARNBACH, M. N. 2006. Chlamydia trachomatis-derived deubiquitinating enzymes in mammalian cells during infection. *Mol Microbiol,* 61**,** 142-50.

MITCHELL, C. M., HUTTON, S., MYERS, G. S., BRUNHAM, R. & TIMMS, P. 2010. Chlamydia pneumoniae is genetically diverse in animals and appears to have crossed the host barrier to humans on (at least) two occasions. *PLoS Pathog,* 6**,** e1000903.

MIYAIRI, I. & BYRNE, G. I. 2006. Chlamydia and programmed cell death. *Curr Opin Microbiol,* 9**,** 102-8.

MIYAIRI, I., MAHDI, O. S., OUELLETTE, S. P., BELLAND, R. J. & BYRNE, G. I. 2006. Different growth rates of Chlamydia trachomatis biovars reflect pathotype. *Journal of Infectious Diseases,* 194**,** 350-357.

MLOT, C. 1996. Chlamydia linked to atherosclerosis. *Science,* 272**,** 1422.

MORRE, S. A., OSSEWAARDE, J. M., LAN, J., VAN DOORNUM, G. J., WALBOOMERS, J. M., MACLAREN, D. M., MEIJER, C. J. & VAN DEN BRULE, A. J. 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**,** 345-51.

MORRE, S. A., OSSEWAARDE, J. M., SAVELKOUL, P. H., STOOF, J., MEIJER, C. J. & VAN DEN BRULE, A. J. 2000. Analysis of genetic heterogeneity in Chlamydia trachomatis clinical isolates of

serovars D, E, and F by amplified fragment length polymorphism. *J Clin Microbiol,* 38**,** 3463-6.

MOULDER, J. W. 1966. The relation of the psittacosis group (Chlamydiae) to bacteria and viruses. *Annu Rev Microbiol,* 20**,** 107-30.

MOULDER, J. W., LEVY, N. J. & SCHULMAN, L. P. 1980. Persistent infection of mouse fibroblasts (L cells) with Chlamydia psittaci: evidence for a cryptic chlamydial form. *Infect Immun,* 30**,** 874-83.

MUHLESTEIN, J. B., ANDERSON, J. L., CARLQUIST, J. F., SALUNKHE, K., HORNE, B. D., PEARSON, R. R., BUNCH, T. J., ALLEN, A., TREHAN, S. & NIELSON, C. 2000. Randomized secondary prevention trial of azithromycin in patients with coronary artery disease: primary clinical results of the ACADEMIC study. *Circulation,* 102**,** 1755-60.

MUHLESTEIN, J. B., HAMMOND, E. H., CARLQUIST, J. F., RADICKE, E., THOMSON, M. J., KARAGOUNIS, L. A., WOODS, M. L. & ANDERSON, J. L. 1996. Increased incidence of Chlamydia species within the coronary arteries of patients with symptomatic atherosclerotic versus other forms of cardiovascular disease. *J Am Coll Cardiol,* 27**,** 1555-61.

MUNDAY, P. E., CARDER, J. M. & TAYLOR-ROBINSON, D. 1985. Chlamydial proctitis? *Genitourin Med,* 61**,** 376-8.

MUNDAY, P. E., DAWSON, S. G., JOHNSON, A. P., OSBORN, M. J., THOMAS, B. J., PHILIP, S., JEFFRIES, D. J., HARRIS, J. R. & TAYLOR-ROBINSON, D. 1981. A microbiological study of non-gonococcal proctitis in passive male homosexuals. *Postgrad Med J,* 57**,** 705-11.

NAVARRO, C., JOLLY, A., NAIR, R. & CHEN, Y. 2002. Risk factors for genital chlamydial infection. *Can J Infect Dis,* 13**,** 195-207.

NEU, N., NELSON, J., RANSON, E., WHITTIER, S., SHAW, A. & COHALL, A. 2012. Urine and extragenital gonococcal and Chlamydia infections in HIV positive young men who have sex with men (YMSM): enhanced detection of rectal Chlamydia when screening with nucleic acid amplified tests versus cultures. *Journal of Men's Health,* 9**,** 89-93.

NEWCOMER, C. E., ANVER, M. R., SIMMONS, J. L., WILCKE, B. W., JR. & NACE, G. W. 1982. Spontaneous and experimental infections of Xenopus laevis with Chlamydia psittaci. *Lab Anim Sci,* 32**,** 680-6.

NGEOW, Y. F., MUNDAY, P. E., EVANS, R. T. & TAYLOR-ROBINSON, D. 1981. Taking cell cultures to the patient in an attempt to improve chlamydial isolation. *British Journal of Venereal Diseases,* 57**,** 44-46.

NHS 2012. Sexual Health in Southampton City.

NIEUWENHUIS, R. F., OSSEWAARDE, J. M., GOTZ, H. M., DEES, J., THIO, H. B., THOMEER, M. G., DEN HOLLANDER, J. C., NEUMANN, M. H. & VAN DER MEIJDEN, W. I. 2004. Resurgence of lymphogranuloma venereum in Western Europe: an outbreak of Chlamydia trachomatis serovar l2 proctitis in The Netherlands among men who have sex with men. *Clin Infect Dis,* 39**,** 996-1003.

NILSSON, U., HELLBERG, D., SHOUBNIKOVA, M., NILSSON, S. & MARDH, P. A. 1997. Sexual behavior risk factors associated with bacterial vaginosis and Chlamydia trachomatis infection. *Sex Transm Dis,* 24**,** 241-6.

O'CONNELL, C. M., IONOVA, I. A., QUAYLE, A. J., VISINTIN, A. & INGALLS, R. R. 2006. Localization of TLR2 and MyD88 to Chlamydia trachomatis inclusions. Evidence for signaling by intracellular TLR2 during infection with an obligate intracellular pathogen. *J Biol Chem,* 281**,** 1652-9.

O'CONNOR, C. M., DUNNE, M. W., PFEFFER, M. A., MUHLESTEIN, J. B., YAO, L., GUPTA, S., BENNER, R. J., FISHER, M. R. & COOK, T. D. 2003. Azithromycin for the secondary prevention of coronary heart disease events: the WIZARD study: a randomized controlled trial. *JAMA,* 290**,** 1459-66.

OLIVE, C., TOTH, I. & JACKSON, D. 2001. Technological advances in antigen delivery and synthetic peptide vaccine developmental strategies. *Mini Rev Med Chem,* 1**,** 429-38.

OLIVE, D. M. & BEAN, P. 1999. Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol,* 37**,** 1661-9.

OSTERLUND, A., PERSSON, T., PERSSON, I., LYSEN, M. & HERRMANN, B. 2005. Improved contact tracing of Chlamydia trachomatis in a Swedish county--is genotyping worthwhile? *Int J STD AIDS,* 16**,** 9-13.

PAL, S., BARNHART, K. M., WEI, Q., ABAI, A. M., PETERSON, E. M. & DE LA MAZA, L. M. 1999. Vaccination of mice with DNA plasmids coding for the Chlamydia trachomatis major outer membrane protein elicits an immune response but fails to protect against a genital challenge. *Vaccine,* 17**,** 459-65.

PANNEKOEK, Y., MORELLI, G., KUSECEK, B., MORRE, S. A., OSSEWAARDE, J. M., LANGERAK, A. A. & VAN DER ENDE, A. 2008. Multi locus sequence of Chlamydiales: clonal groupings within the obligate intracellular bacteria Chlamydia trachomatis. *Bmc Microbiology,* 8.

PANNEKOEK, Y., VAN DER ENDE, A., EIJK, P. P., VAN MARLE, J., DE WITTE, M. A., OSSEWAARDE, J. M., VAN DEN BRULE, A. J., MORRE, S. A. & DANKERT, J. 2001. Normal IncA expression and fusogenicity of inclusions in Chlamydia trachomatis isolates with the incA I47T mutation. *Infect Immun,* 69**,** 4654-6.

PANTOJA, L. G., MILLER, R. D., RAMIREZ, J. A., MOLESTINA, R. E. & SUMMERSGILL, J. T. 2001. Characterization of Chlamydia pneumoniae persistence in HEp-2 cells treated with gamma interferon. *Infect Immun,* 69**,** 7927-32.

PARR, M. B. & PARR, E. L. 1990. Antigen recognition in the female reproductive tract: I. Uptake of intraluminal protein tracers in the mouse vagina. *J Reprod Immunol,* 17**,** 101-14.

PEDERSEN, L. N., HERRMANN, B. & MOLLER, J. K. 2009. Typing Chlamydia trachomatis: from egg yolk to nanotechnology. *FEMS Immunology and Medical Microbiology,* 55**,** 120-130.

PEDERSEN, L. N., PODENPHANT, L. & MOLLER, J. K. 2008. Highly discriminative genotyping of Chlamydia trachomatis using omp1 and a set of variable number tandem repeats. *Clinical Microbiology and Infection,* 14**,** 644-652.

PELLOWE, C. & PRATT, R. J. 2006. Neonatal conjunctivitis and pneumonia due to chlamydia infection. *Sexually Transmitted Diseases,* 2**,** 16-17.

PEUCHANT, O., BALDIT, C., LE ROY, C., TROMBERT-PAOLANTONI, S., CLERC, M., BEBEAR, C. & DE BARBEYRAC, B. 2011. First case of Chlamydia trachomatis L2b proctitis in a woman. *Clin Microbiol Infect,* 17**,** E21-3.

PEUCHANT, O., LE ROY, C., HERRMANN, B., CLERC, M., BEBEAR, C. & DE BARBEYRAC, B. 2012. MLVA Subtyping of Genovar E Chlamydia trachomatis Individualizes the Swedish Variant and Anorectal Isolates from Men who Have Sex with Men. *Plos One,* 7**,** e31538.

PHE 2013. Health Protection Report: Weekly Report.

PORTER, E., YANG, H., YAVAGAL, S., PREZA, G. C., MURILLO, O., LIMA, H., GREENE, S., MAHOOZI, L., KLEIN-PATEL, M., DIAMOND, G., GULATI, S., GANZ, T., RICE, P. A. & QUAYLE, A. J. 2005. Distinct defensin profiles in Neisseria gonorrhoeae and Chlamydia trachomatis urethritis reveal novel epithelial cell-neutrophil interactions. *Infect Immun,* 73**,** 4823-33.

POTTER, M. E., KAUFMANN, A. K. & D, P. B. 1979. Psittacosis surveillance report. Atlanta: CDC.

QUINN, T. C., GAYDOS, C., SHEPHERD, M., BOBO, L., HOOK, E. W., 3RD, VISCIDI, R. & ROMPALO, A. 1996. Epidemiologic and microbiologic correlates of Chlamydia trachomatis infection in sexual partnerships. *JAMA,* 276**,** 1737-42.

QUINN, T. C., GOODELL, S. E., MKRTICHIAN, E., SCHUFFLER, M. D., WANG, S. P., STAMM, W. E. & HOLMES, K. K. 1981. Chlamydia trachomatis proctitis. *N Engl J Med,* 305**,** 195-200.

QUINT, K. D., BOM, R. J., QUINT, W. G., BRUISTEN, S. M., VAN DER LOEFF, M. F., MORRE, S. A. & DE VRIES, H. J. 2011. Anal infections with concomitant Chlamydia trachomatis genotypes among men who have sex with men in Amsterdam, the Netherlands. *BMC.Infect.Dis.,* 11**,** 63.

RAMSEY, K. H., COTTER, T. W., SALYER, R. D., MIRANPURI, G. S., YANEZ, M. A., POULSEN, C. E., DEWOLFE, J. L. & BYRNE, G. I. 1999. Prior genital tract infection with a murine or human

biovar of Chlamydia trachomatis protects mice against heterotypic challenge infection. *Infect Immun,* 67**,** 3019-25.

RASMUSSEN, S. J., ECKMANN, L., QUAYLE, A. J., SHEN, L., ZHANG, Y. X., ANDERSON, D. J., FIERER, J., STEPHENS, R. S. & KAGNOFF, M. F. 1997. Secretion of proinflammatory cytokines by epithelial cells in response to Chlamydia infection suggests a central role for epithelial cells in chlamydial pathogenesis. *J Clin Invest,* 99**,** 77-87.

REEVE, P., OWEN, J. & ORIEL, J. D. 1975. Laboratory procedures for the isolation of chlamydia trachomatis from the human genital tract. *Journal of Clinical Pathology,* 28**,** 910-914.

RIPA, T. & NILSSON, P. A. 2007. A Chlamydia trachomatis strain with a 377-bp deletion in the cryptic plasmid causing false-negative nucleic acid amplification tests. *Sex Transm Dis,* 34**,** 255-6.

ROAN, N. R., GIERAHN, T. M., HIGGINS, D. E. & STARNBACH, M. N. 2006. Monitoring the T cell response to genital tract infection. *Proc Natl Acad Sci U S A,* 103**,** 12069-74.

ROAN, N. R. & STARNBACH, M. N. 2006. Antigen-specific CD8+ T cells respond to Chlamydia trachomatis in the genital mucosa. *J Immunol,* 177**,** 7974-9.

ROAN, N. R. & STARNBACH, M. N. 2008. Immune-mediated control of Chlamydia infection. *Cell Microbiol,* 10**,** 9-19.

RODRIGUEZ, P., ALLARDET-SERVENT, A., DE BARBEYRAC, B., RAMUZ, M. & BEBEAR, C. 1994. Genetic variability among Chlamydia trachomatis reference and clinical strains analyzed by pulsed-field gel electrophoresis. *J Clin Microbiol,* 32**,** 2921-8.

RODRIGUEZ, P., VEKRIS, A., DE BARBEYRAC, B., DUTILH, B., BONNET, J. & BEBEAR, C. 1991. Typing of Chlamydia trachomatis by restriction endonuclease analysis of the amplified major outer membrane protein gene. *J Clin Microbiol,* 29**,** 1132-6.

ROTA, T. R. 1977. Chlamydia trachomatis in cell culture. II. Susceptibility of seven established mammalian cell types in vitro. Adaptation of trachoma organisms to McCoy and BHK-21 cells. *In Vitro,* 13**,** 280-92.

SCHACHTER, J., STEPHENS, R. S., TIMMS, P., KUO, C., BAVOIL, P. M., BIRKELUND, S., BOMAN, J., CALDWELL, H., CAMPBELL, L. A., CHERNESKY, M., CHRISTIANSEN, G., CLARKE, I. N., GAYDOS, C., GRAYSTON, J. T., HACKSTADT, T., HSIA, R., KALTENBOECK, B., LEINONNEN, M., OJCIUS, D., MCCLARTY, G., ORFILA, J., PEELING, R., PUOLAKKAINEN, M., QUINN, T. C., RANK, R. G., RAULSTON, J., RIDGEWAY, G. L., SAIKKU, P., STAMM, W. E., TAYLOR-ROBINSON, D. T., WANG, S. P. & WYRICK, P. B. 2001. Radical changes to chlamydial taxonomy are not necessary just yet. *Int J Syst Evol Microbiol,* 51**,** 249; author reply 251-3.

SCHAUTTEET, K., DE CLERCQ, E. & VANROMPAY, D. 2011. Chlamydia trachomatis vaccine research through the years. *Infect Dis Obstet Gynecol,* 2011**,** 963513.

SCHLEIFER, K. H. & STACKEBRANDT, E. 1983. Molecular systematics of prokaryotes. *Annu Rev Microbiol,* 37**,** 143-87.

SCIEUX, C., GRIMONT, F., REGNAULT, B., BIANCHI, A., KOWALSKI, S. & GRIMONT, P. A. 1993. Molecular typing of Chlamydia trachomatis by random amplification of polymorphic DNA. *Res Microbiol,* 144**,** 395-404.

SETH-SMITH, H. M., HARRIS, S. R., SKILTON, R. J., RADEBE, F. M., GOLPARIAN, D., SHIPITSYNA, E., DUY, P. T., SCOTT, P., CUTCLIFFE, L. T., O'NEILL, C., PARMAR, S., PITT, R., BAKER, S., ISON, C. A., MARSH, P., JALAL, H., LEWIS, D. A., UNEMO, M., CLARKE, I. N., PARKHILL, J. & THOMSON, N. R. 2013. Whole-genome sequences of Chlamydia trachomatis directly from clinical samples without culture. *Genome Res,* 23**,** 855-66.

SETH-SMITH, H. M. B., HARRIS, S. R., PERSSON, K., MARSH, P., BARRON, A., BIGNELL, A., BJARTLING, C., CLARK, L., CUTCLIFFE, L. T., LAMBDEN, P. R., LENNARD, N., LOCKEY, S. J., QUAIL, M. A., SALIM, O., SKILTON, R. J., WANG, Y. B., HOLLAND, M. J., PARKHILL, J., THOMSON, N. R. & CLARKE, I. N. 2009. Co-evolution of genomes and plasmids within Chlamydia trachomatis and the emergence in Sweden of a new variant strain. *Bmc Genomics,* 10**,** 239.

SKERMAN V.B.D, MCGOWAN V & SNEATH P.H.A 1980. Approved Lists of Bacterial Names. *International Journal of Systematic Bacteriology,* 30.

SKILTON, R. J., CUTCLIFFEN, L. T., BARLOW, D., WANG, Y., SALIM, O., LAMBDEN, P. R. & CLARKE, I. N. 2009. Penicillin induced persistence in Chlamydia trachomatis: high quality time lapse video analysis of the developmental cycle. *Plos One,* 4**,** e7723.

SODERLUND, G. & KIHLSTROM, E. 1983. Effect of methylamine and monodansylcadaverine on the susceptibility of McCoy cells to Chlamydia trachomatis infection. *Infect Immun,* 40**,** 534-41.

SPAARGAREN, J., FENNEMA, H. S., MORRE, S. A., DE VRIES, H. J. & COUTINHO, R. A. 2005. New lymphogranuloma venereum Chlamydia trachomatis variant, Amsterdam. *Emerg Infect Dis,* 11**,** 1090-2.

STAMATAKIS, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* 22**,** 2688-90.

STARNBACH, M. N., BEVAN, M. J. & LAMPE, M. F. 1994. Protective cytotoxic T lymphocytes are induced during murine infection with Chlamydia trachomatis. *J Immunol,* 153**,** 5183-9.

STEELE, L. N., BALSARA, Z. R. & STARNBACH, M. N. 2004. Hematopoietic cells are required to initiate a Chlamydia trachomatis-specific CD8+ T cell response. *J Immunol,* 173**,** 6327-37.

STEIMLE, V., SIEGRIST, C. A., MOTTET, A., LISOWSKA-GROSPIERRE, B. & MACH, B. 1994. Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA. *Science,* 265**,** 106-9.

STEPHENS, R. S., KOSHIYAMA, K., LEWIS, E. & KUBO, A. 2001. Heparin-binding outer membrane protein of chlamydiae. *Mol Microbiol,* 40**,** 691-9.

STEPHENS, R. S., MYERS, G., EPPINGER, M. & BAVOIL, P. M. 2009. Divergence without difference: phylogenetics and taxonomy of Chlamydia resolved. *FEMS Immunol Med Microbiol,* 55**,** 115-9.

STEPHENS, R. S., TAM, M. R., KUO, C. C. & NOWINSKI, R. C. 1982. Monoclonal antibodies to Chlamydia trachomatis: antibody specificities and antigen characterization. *J Immunol,* 128**,** 1083-9.

STOTHARD, D. R., VAN DER POL, B., SMITH, N. J. & JONES, R. B. 1998. Effect of serial passage in tissue culture on sequence of omp1 from Chlamydia trachomatis clinical isolates. *Journal of Clinical Microbiology,* 36**,** 3686-3688.

SU, H. & CALDWELL, H. D. 1995. CD4+ T cells play a significant role in adoptive immunity to Chlamydia trachomatis infection of the mouse genital tract. *Infect Immun,* 63**,** 3302-8.

SU, H. & CALDWELL, H. D. 1998. Sulfated polysaccharides and a synthetic sulfated polymer are potent inhibitors of Chlamydia trachomatis infectivity in vitro but lack protective efficacy in an in vivo murine model of chlamydial genital tract infection. *Infect Immun,* 66**,** 1258-60.

SU, H., RAYMOND, L., ROCKEY, D. D., FISCHER, E., HACKSTADT, T. & CALDWELL, H. D. 1996. A recombinant Chlamydia trachomatis major outer membrane protein binds to heparan sulfate receptors on epithelial cells. *Proc Natl Acad Sci U S A,* 93**,** 11143-8.

SUCHLAND, R. J., ECKERT, L. O., HAWES, S. E. & STAMM, W. E. 2003. Longitudinal assessment of infecting serovars of Chlamydia trachomatis in Seattle public health clinics: 1988-1996. *Sex Transm Dis,* 30**,** 357-61.

SUCHLAND, R. J., ROCKEY, D. D., BANNANTINE, J. P. & STAMM, W. E. 1999. Isolates of Chlamydia trachomatis that occupy nonfusogenic inclusions lack IncA, a protein localized to the inclusion membrane

Infect Immun, 68**,** 360-7.

SULAIMAN, M. Z., FOSTER, J. & PUGH, S. F. 1987. Prevalence of Chlamydia trachomatis infection in homosexual men. *Genitourin Med,* 63**,** 179-81.

TARAKTCHOGLOU, M., PACEY, A. A., TURNBULL, J. E. & ELEY, A. 2001. Infectivity of Chlamydia trachomatis serovar LGV but not E is dependent on host cell heparan sulfate. *Infect Immun,* 69**,** 968-76.

THEUNISSEN, J. J., VAN HEIJST, B. Y., WAGENVOORT, J. H., STOLZ, E. & MICHEL, M. F. 1992. Factors influencing the infectivity of Chlamydia pneumoniae elementary bodies on HL cells. *Journal of Clinical Microbiology,* 30**,** 1388-1391.

THYGESON, P. 1962. Trachoma virus: historical background and review of isolates. *Ann N Y Acad Sci,* 98**,** 6-13.

TODD, W. J. & CALDWELL, H. D. 1985. The interaction of Chlamydia trachomatis with host cells: ultrastructural studies of the mechanism of release of a biovar II strain from HeLa 229 cells. *J Infect Dis,* 151**,** 1037-44.

TROMBETTA, E. S. & MELLMAN, I. 2005. Cell biology of antigen processing in vitro and in vivo. *Annu Rev Immunol,* 23**,** 975-1028.

TUFFREY, M., FALDER, P., GALE, J. & TAYLOR-ROBINSON, D. 1986. Salpingitis in mice induced by human strains of Chlamydia trachomatis. *Br J Exp Pathol,* 67**,** 605-16.

TWIN, J., MOORE, E. E., GARLAND, S. M., STEVENS, M. P., FAIRLEY, C. K., DONOVAN, B., RAWLINSON, W. & TABRIZI, S. N. 2010. Chlamydia trachomatis Genotypes Among Men Who Have Sex With Men in Australia. *Sex Transm.Dis.*

UNEMO, M., SETH-SMITH, H. M., CUTCLIFFE, L. T., SKILTON, R. J., BARLOW, D., GOULDING, D., PERSSON, K., HARRIS, S. R., KELLY, A., BJARTLING, C., FREDLUND, H., OLCEN, P., THOMSON, N. R. & CLARKE, I. N. 2010. The Swedish new variant of Chlamydia trachomatis : Genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology,* 156**,** 1394-1404.

VAN BELKUM, A., SCHERER, S., VAN LEEUWEN, W., WILLEMSE, D., VAN ALPHEN, L. & VERBRUGH, H. 1997. Variable number of tandem repeats in clinical strains of Haemophilus influenzae. *Infect Immun,* 65**,** 5017-27.

VAN BELKUM, A., TASSIOS, P. T., DIJKSHOORN, L., HAEGGMAN, S., COOKSON, B., FRY, N. K., FUSSING, V., GREEN, J., FEIL, E., GERNER-SMIDT, P., BRISSE, S. & STRUELENS, M. 2007. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin.Microbiol.Infect.,* 13 Suppl 3**,** 1-46.

VAN DER POL, B. J. & JONES, R. B. 1992. Comparison of immunotyping of Chlamydia trachomatis by indirect fluorescent-antibody staining and radioimmunoassay. *J Clin Microbiol,* 30**,** 1014-5.

VANROMPAY, D., DE MEURICHY, W., DUCATELLE, R. & HAESEBROUCK, F. 1994. Pneumonia in Moorish tortoises (Testudo graeca) associated with avian serovar A Chlamydia psittaci. *Vet Rec,* 135**,** 284-5.

VANROMPAY, D., DUCATELLE, R. & HAESEBROUCK, F. 1995. Chlamydia psittaci infections: a review with emphasis on avian chlamydiosis. *Vet Microbiol,* 45**,** 93-119.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., VAN DE LEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M. & ET AL. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res,* 23**,** 4407-14.

WANG, S. P. & GRAYSTON, J. T. 1963. Classification of Trachoma Virus Strains by Protection of Mice from Toxic Death. *J Immunol,* 90**,** 849-56.

WANG, S. P. & GRAYSTON, J. T. 1970. Immunologic relationship between genital TRIC, lymphogranuloma venereum, and related organisms in a new microtiter indirect immunofluorescence test. *Am J Ophthalmol,* 70**,** 367-74.

WANG, S. P., GRAYSTON, J. T. & GALE, J. L. 1973. Three new immunologic types of trachoma-inclusion conjunctivitis organisms. *J Immunol,* 110**,** 873-9.

WANG, S. P., KUO, C. C., BARNES, R. C., STEPHENS, R. S. & GRAYSTON, J. T. 1985. Immunotyping of Chlamydia trachomatis with monoclonal antibodies. *J Infect Dis,* 152**,** 791-800.

WANG, Y., KAHANE, S., CUTCLIFFE, L. T., SKILTON, R. J., LAMBDEN, P. R. & CLARKE, I. N. 2011a. Development of a transformation system for Chlamydia trachomatis: restoration of glycogen biosynthesis by acquisition of a plasmid shuttle vector. *PLoS Pathog.,* 7**,** e1002258.

WANG, Y., SKILTON, R. J., CUTCLIFFE, L. T., ANDREWS, E., CLARKE, I. N. & MARSH, P. 2011b. Evaluation of a High Resolution Genotyping Method for Chlamydia trachomatis Using Routine Clinical Samples. *PLoS.One.,* 6**,** e16971.

WARD, H., MARTIN, I., MACDONALD, N., ALEXANDER, S., SIMMS, I., FENTON, K., FRENCH, P., DEAN, G. & ISON, C. 2007. Lymphogranuloma venereum in the United kingdom. *Clinical Infectious Diseases,* 44**,** 26-32.

WARD, M. E. 1983. Chlamydial classification, development and structure. *British Medical Bulletin,* 39**,** 109-115.

WARD, M. E. 2004. *The chlamydial developmental cycle in pictures.* [Online]. Available: www.chlamydiae.com 2013].

WARD, M. E. & MURRAY, A. 1984. Control mechanisms governing the infectivity of Chlamydia trachomatis for HeLa cells: mechanisms of endocytosis. *J Gen Microbiol,* 130**,** 1765-80.

WAYNE, L. G., BRENNER, D. J., COLWELL, R. R., GRIMONT, P. A. D., KANDLER, O., KRICHEVSKY, M. I., MOORE, L. H., MOORE, W. E. C., MURRAY, R. G. E., STACKEBRANDT, E., STARR, M. P. & TRUPER, H. G. 1987. International committee on Systematic Bacteriology. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic Bacteriology,* 37**,** 463-464.

WELSH, J. & MCCLELLAND, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res,* 18**,** 7213-8.

WHO. 2012. *Weekly epidemiological record* [Online]. Available: http://wwww.who.int/wer.

WILLIAMS, J. G., KUBELIK, A. R., LIVAK, K. J., RAFALSKI, J. A. & TINGEY, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res,* 18**,** 6531-5.

WILSON, D. P., REGAN, D. G., HEYMER, K. J., JIN, F., PRESTAGE, G. P. & GRULICH, A. E. 2010. Serosorting may increase the risk of HIV acquisition among men who have sex with men. *Sex Transm Dis,* 37**,** 13-7.

WILSON, D. P., TIMMS, P. & MCELWAIN, D. L. 2003. A mathematical model for the investigation of the Th1 immune response to Chlamydia trachomatis. *Math Biosci,* 182**,** 27-44.

WILSON, J. S., HONEY, E., TEMPLETON, A., PAAVONEN, J., MARDH, P. A. & STRAY-PEDERSEN, B. 2002. A systematic review of the prevalence of Chlamydia trachomatis among European women. *Hum Reprod Update,* 8**,** 385-94.

WOLKE, R. E., WYAND, D. S. & KHAIRALLAH, L. H. 1970. A light and electron microscopic study of epitheliocystis disease in the gills of Connecticut striped bass (Morone saxatilis) and white perch (Morone americanus). *J Comp Pathol,* 80**,** 559-63.

WYRICK, P. B. 2010. Chlamydia trachomatis persistence in vitro: an overview. *J Infect Dis,* 201 Suppl 2**,** S88-95.

YANG, C. L., MACLEAN, I. & BRUNHAM, R. C. 1993. DNA sequence polymorphism of the Chlamydia trachomatis omp1 gene. *J Infect Dis,* 168**,** 1225-30.

YUAN, Y., ZHANG, Y. X., WATKINS, N. G. & CALDWELL, H. D. 1989. Nucleotide and deduced amino acid sequences for the four variable domains of the major outer membrane proteins of the 15 Chlamydia trachomatis serovars. *Infect Immun,* 57**,** 1040-9.

ZHONG, G., FAN, P., JI, H., DONG, F. & HUANG, Y. 2001. Identification of a chlamydial protease-like activity factor responsible for the degradation of host transcription factors. *J Exp Med,* 193**,** 935-42.