

***Chlamydia trachomatis* from Individuals in a Sexually Transmitted Disease Core Group Exhibit Frequent Sequence Variation in the Major Outer Membrane Protein (omp1) Gene**

Robert Brunham, Chunlin Yang, Ian Maclean, Joshua Kimani,* Gregory Maitha,* and Frank Plummer*

Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba, R3E 0W3 Canada; and *Department of Medical Microbiology, University of Nairobi, Kenya

Abstract

60 cervical *Chlamydia trachomatis* infections identified by antigen detection from 51 prostitute women in Nairobi, Kenya were evaluated for sequence polymorphism in the major outer membrane protein (omp1) gene. DNA from clinical specimens was amplified by the polymerase chain reaction and cycle sequenced through variable domains (VD) 1, 2, and 4. 37 (63%) samples had variant VD sequences, 19 (32%) samples had prototype VD sequences, and 4 (6%) samples contained omp1 sequences from two or more *C. trachomatis* strains. Among the 37 variant strains, 18 had two or fewer nucleotide substitutions in one or two VDs and represented point mutational drift variants. 19 strains had a larger number of nucleotide changes and displayed mosaic omp1 sequences that may have been generated by omp1 VD recombination. We conclude that the prevalence of *C. trachomatis* omp1 DNA polymorphism is substantial among prostitute women in Nairobi, Kenya and that this is the likely result of immune selection pressure. (*J. Clin. Invest.* 1994. 94:458–463.) **Key words:** *Chlamydia trachomatis* • omp1 • antigenic variation

Introduction

The major outer membrane protein (MOMP)¹ gene (omp1) of *Chlamydia trachomatis* is a single copy gene with multiple alleles. Each omp1 allele specifies a specific serologic variant or serovar of the organism (1). Sequence analysis shows that omp1 has four variable regions or domains (VD1 to VD4) interspersed with five constant regions (1, 2). Antibody responses to the *C. trachomatis* MOMP are primarily directed to the VDs (3). Furthermore, monoclonal antibodies to an epitope

sequence within a VD can neutralize the organism in cell culture and can distinguish between VD sequences which vary by only one or two amino acids (4, 5).

Limited epidemiologic data and phylogenetic analysis suggest that omp1 polymorphism is relatively stable and the tempo for genetic change is slow (4, 6, 7). However, MOMP-specific monoclonal antibody typing of *C. trachomatis* isolated from cases of trachoma and sexually transmitted chlamydial infections has identified a surprisingly large number of examples of atypical immunoreactivity (8–11). Detailed sequencing studies of the omp1 gene by Dean et al. (8) identified that atypical patterns of monoclonal antibody reactivity patterns mapped entirely or predominantly to amino acid sequence changes in the MOMP VDs. Lampe et al. (11) molecularly studied three variants of serovar D, two of serovar I, and one of serovar L₂ and suggested that sequence variation arises by point mutation or, possibly, by recombination of MOMP VDs. Recently, we reported that 38% of 49 clinical samples of *C. trachomatis* identified in Winnipeg from individuals with sexually acquired chlamydial infections also had sequence polymorphism in the omp1 gene (12). In aggregate, these observations suggest that the omp1 gene may exhibit a far greater degree of polymorphism than is currently appreciated.

Polymorphism in antigen-specifying genes for sexually transmitted pathogens may be most apparent among isolates originating from a sexually transmitted disease (STD) core group. This hypothesis derives from the assumption that immune responses that occur among STD core group members who are repeatedly exposed to an STD pathogen will produce a significant barrier to persistence of the pathogen in the core group (13). Thus, strain diversification of a major antigen specifying gene (omp1 in the case of *C. trachomatis*) will be necessary to allow ecologic success for the pathogen in the face of strain specific immunity in the host population (14).

To test this hypothesis, 60 *C. trachomatis* infections identified by antigen detection among women who work as prostitutes in the Pumwani district of Nairobi, Kenya were evaluated for sequence variation in the omp1 VDs. The results show that *C. trachomatis* omp1 sequence variation is extraordinarily prevalent among core group members, presumably in response to immune selection.

Methods

Study population

In April 1991, a cohort of prostitute women working in the Pumwani district of Nairobi, Kenya was established to study the immunoepidemi-

Address correspondence to Dr. Robert C. Brunham, Department of Medical Microbiology, University of Manitoba, Room 543, 730 William Avenue, Winnipeg, Manitoba, R3E 0W3, Canada.

Received for publication 15 December 1993 and in revised form 4 April 1994.

1. Abbreviations used in this paper: MOMP, major outer membrane protein, omp1, outer membrane protein gene, VD, variable domain.

J. Clin. Invest.

© The American Society for Clinical Investigation, Inc.

0021-9738/94/07/0458/06 \$2.00

Volume 94, July 1994, 458–463

Table 1. Summary of Sequence Analysis of VD1, VD2, and VD4 of *Chlamydia trachomatis omp 1* for 60 Strains Isolated among Pumwani Prostitute Cohort between April 1991 and April 1992

	Serovar	Number	Percentage in each group
Prototype	E	15	32
	G	4	
Minor variants	C	3	30
	K	7	
	L ₁	1	
	L ₂	7	
Major variants	I-like/H-like	4	32
	L ₁ -like/L ₂ -like	11	
	L ₂ -like/L ₁ -like	2	
	L ₃ -like/H-like	2	
Mixed		4	6

ology of sexually transmitted *C. trachomatis* infection. At enrollment demographic, sexual, reproductive, and medical histories were obtained in a standard format interview and a complete physical examination including vaginal speculum and bimanual examination was performed. Sera for HIV serology, CD4 T cell enumeration, and endocervical swabs for *Neisseria gonorrhoeae* culture and *C. trachomatis* antigen detection were obtained at enrollment. Women were scheduled for routine follow-up at two weekly intervals. At each visit, a limited interview regarding condom use, number of sexual partners and STD symptoms was conducted and a general and genital examination including vaginal speculum and bimanual was performed. At each visit, endocervical swabs for gonococcal and chlamydial infection were obtained. Chlamydia infected women were treated with doxycycline 100 mg orally twice daily for 7 d and a test-of-cure performed 4 d after completion of treatment.

Laboratory tests

Chlamydia antigen detection. Cervical swabs were collected and evaluated for chlamydial antigen using a commercially available enzyme immunoassay (Clearview; Unipath, Nepean, Ontario, Canada). Specimens were processed according to the manufacturer's instructions.

***C. trachomatis omp1* gene characterization.** Samples found to contain chlamydial antigen were stored at -20°C and shipped to Winnipeg where DNA was extracted with 0.1 M Tris-HCl, pH 8.0, 0.10 M NaCl, 0.005 M EDTA, 1% sodium dodecyl sulfate, and 50 µg/ml proteinase K as previously described (12). DNA was further extracted with an equal volume of phenol and chloroform and precipitated with sodium acetate and isopropanol to remove nonspecific inhibitors of the polymerase chain reaction (PCR).

***Omp1* gene amplification and variable domain DNA sequencing.** An 879-bp product of the *omp1* gene was amplified by PCR using previously described flanking and nested primers (12). Amplification with PCR involved the use of 2 U of *Thermus aquaticus* (Taq) polymerase and annealing temperatures of 55 or 60°C and an extending temperature of 72°C for 2 to 3 min. The DNA sequence of VD1, VD2, and VD4 was determined by the double stranded DNA dideoxynucleotide chain termination method using 5' end-labeled [³²P]ATP oligonucleotide primers as described (12). Fig. 1 shows the position of the primers with

respect to the *omp1* gene VD organization and the legend describes the nucleotide sequence of each primer. Sequences for VD1, 2, and 4 among samples identified in the current study were compared with the VD sequences from the 15 prototype *C. trachomatis* serovars as reported by Yuan et al. (15).

Clonal analysis of *omp1*. *Omp1* was re-amplified from the secondary PCR reaction using 2 µl of the PCR product as template and re-designed primers. Oligonucleotide primers were re-synthesized based on primer 3 and primer 4 and incorporating a 5' KpnI restriction site and six additional unrelated bases. The sequence for modified primer 3 is: CCGTTAGGTACC-TGACTTTGTTTTTCGACCGTGTGTTT; and the sequence for modified primer 4 is: CTGCAAGGTACCTTTTCTAGATTT-CATCTTGTTCAATTG. The KpnI site is underlined. The resulting *omp1* PCR product was subjected to phenol:chloroform extraction followed by ethanol precipitation and dissolved in TE buffer. Proteinase K was added to a final concentration of 200 µg/ml, incubated for 30 min at 37°C and SDS was added to a final concentration of 0.5% and incubated at 68°C for 10 min. The reaction mixture was extracted once with phenol:chloroform and once with chloroform before ethanol precipitation. The DNA was digested with KpnI and subjected to electrophoresis. The correct sized band was excised from the gel, DNA was electroeluted and purified with phenol/chloroform extraction before ligation to dephosphorylated pUC19 restricted with KpnI. A fraction of the ligation reaction was used to transform competent DH5α *E. coli*. White colonies were picked and grown up. The plasmid was purified and the insert fragment was amplified using M13 reverse sequencing primer (AGCGGATAACAATTTACACAGGA) and M13 universal sequencing primer (CGCCAGGGTTTTCCAGTCACGAC). The PCR product was visualized on an ethidium bromide stained 1% agarose gel. The cycle sequencing for *omp1* VD1, VD2, and VD4 (using the purified plasmid that contained the correct sized insert) was done under the same condition as described above.

C. trachomatis Omp1 primers

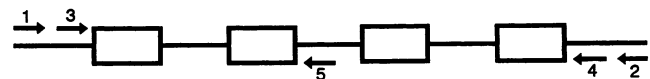


Figure 1. The relationship of the five primers used in the present study to the three variable domains (VDs), VD1, VD2, and VD4 which were sequenced and in relationship to the overall sequence organization of the *Chlamydia trachomatis omp1* gene is shown in the figure. The nucleotide sequence of the primers and their number position in the *omp1* DNA sequence is as follows: Primer 1 (sense) GCC GCT TTG AGT TCT GCT TCC TC (34 → 56); Primer 2 (antisense) ATT TAC GTG AGC AGC TCT CTC AT (1176 → 1154); Primer 3 (sense) T GAC TTT GTT TTC GAC CGT GTT TT (198 → 221); Primer 4 (antisense) TTT TCT AGA TTT CAT CTT GTT CAA T/CTG (1077 → 1051); Primer 5 (antisense) ACA TTC CCA C/GAA/G AGC TGC (621 → 604). Primers 1 and 2 were used to initially amplify, and primers 3 and 4 were used for reamplification. Primer 3 was used to sequence VD1, primer 5 VD2, and primer 4 VD4. These sequence data are available from EMBL/GenBank/DBJ under accession number M14738.

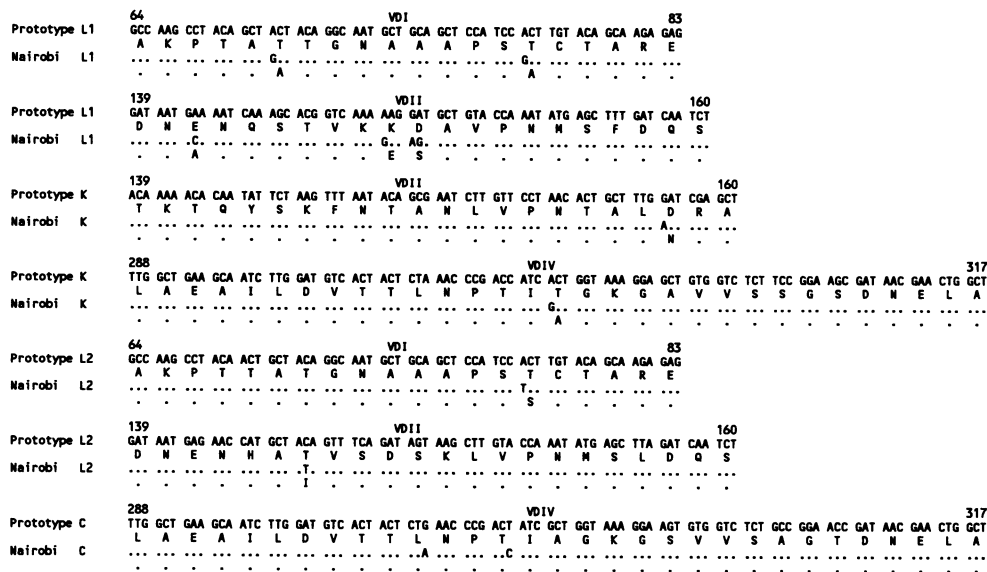


Figure 2. Nucleotide and amino acid sequence comparison of prototype and seven clinical isolates that varied at one or two sites in one or two of three sequenced variable domains (VDs).

Results

Study population. 259 women have been enrolled and have completed a mean length of follow-up of 13.6 ± 7.9 mo with 10.6 ± 7.8 visits. 164 (63%) women were HIV seropositive at enrollment with a mean CD4 T cell count of 437 ± 273 . 231 positive chlamydial isolates were identified from 2575 specimens. 113 women had 231 chlamydial infections with 56 women having a single positive chlamydia test and 57 women having two or more positive tests.

VD sequence analysis of the *C. trachomatis* omp1 gene. The first 60 *C. trachomatis* infections occurred among 51 women and these samples were used in the current study. DNA was extracted from antigen-positive specimens and was characterized by DNA sequencing of omp1 VD1, 2, and 4. Table I summarizes the data. 11 distinct genotypes were identified and 10 of the 11 genotypes were identified on two or more occasions from two or more different individuals. Four genotypes (serovar E, L₁-like/L₂-like, K, and L₂) accounted for 67% of the isolates. 19 (32%) isolates had prototype VD sequences and 37 (62%) isolates had variant VD sequences. Four (6%) samples contained omp1 DNA sequences from multiple serovars and presumably represent mixed infection with two or more *C. trachomatis* strains.

Among the 37 isolates with variant sequences, 18 (30%)

had two or fewer nucleotide substitutions in one or two VDs and were classified as minor variants. Fig. 2 shows the DNA and inferred amino acid sequences in the relevant VDs for each type of minor variant.

19 (32%) isolates had substantial sequence change and were classified as major omp1 variants. Each major omp1 variant was isolated on more than one occasion from more than one person. Figs. 3, 4, 5, and 6 illustrate the DNA and inferred amino acid sequence for VD1, 2, and 4 for each of the major variants. These strains appear to represent omp1 mosaics mixing VDs from different serovars.

Among the major and minor omp1 variant genotypes, 21 nucleotide substitutions occurred in 18 codons. 19 nucleotide substitutions were in nonsynonymous codons. 17 nucleotide substitutions were transitional types in which purine transition predominated (7 A → G, 6 G → A, 2 C → T, and 2 T → C).

Clonal analysis of omp1. To investigate whether omp1 mosaic sequences resulted from PCR artifact, two types of experiments were performed. Omp1 DNA from the second cycle amplification was reamplified with modified primers 3 and 4 containing KpnI restriction sites. Omp1 from each of the four major variant types was reamplified, ligated in pUC19 and cloned into *E. coli*. Re-sequencing of VD1, 2, and 4 from recombinant pUC19 showed that each major variant was a single molecular species with mosaic VD sequences. This excludes primer prefer-

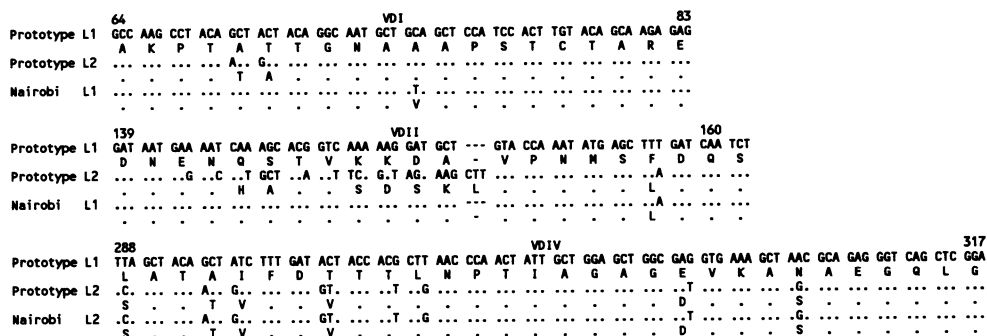


Figure 3. Nucleotide and amino acid sequence comparison of prototype and a clinical isolate of mosaic L₁-like/L₂-like sequences in variable domains (VDs) 1, 2, and 4.

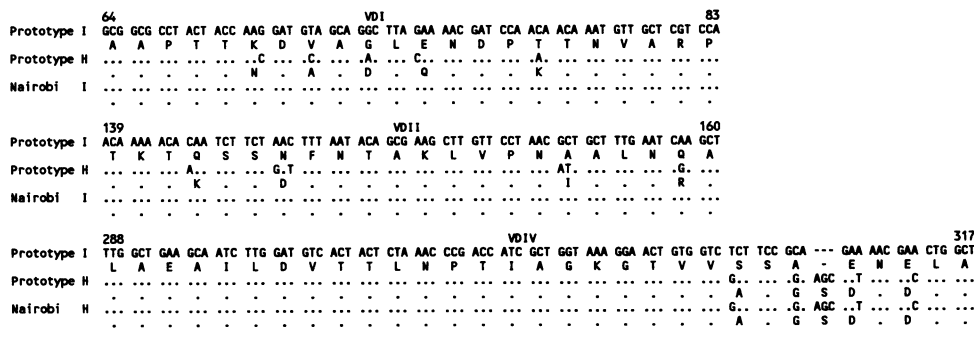


Figure 6. Nucleotide and amino acid sequence comparison of prototype and a clinical isolate of mosaic I/H sequences in variable domains (VDs) 1, 2, and 4.

37 (66%) had variant sequences in VD1, VD2, or VD4 regions of the *omp1* gene. When compared with prototype VD sequences from the 15 standard *C. trachomatis* serovars, 19 (51%) of the 37 variants appeared to involve a major antigenic shift, and 18 (49%) appeared to be minor antigenic drift variants.

The genetic mechanisms for diversification at the *omp1* locus are unknown. Stephens et al. (1) suggested that accumulated clustered base substitution for closely related serovars and insertion and deletions for distantly related serovars accounts for *omp1* diversification. Detailed analysis of the nucleotide changes in the chlamydial *omp1* VD sequences observed in this study suggests that mutational drift is not random since 17 of 21 nucleotide changes were transition-type and 19 of 21 resulted in nonsynonymous codons. The former observation may reflect the specificity of a mutator mechanism and the latter suggests the presence of phenotypic selection. With other prokaryotes, genetic analysis has shown that mutations affecting proofreading, postreplication repair of copying errors, or in PolI raise the frequency of specific base substitutions and thus, enzymes acting at any of these sites may be involved in the origin of *omp1* mutations (18).

Analysis of the 21 nucleotide substitutions observed in the present study revealed that 17 involved transition mutation. Since Taq polymerase can also produce base substitutions that are also predominantly of transitional type, it may be that the origin of some of these mutations is due to Taq polymerase (19–21). However, the following observations make this suggestion unlikely: identical mutations were detected in epidemiologically distinct infections on multiple occasions; repeated cycle sequencing of *omp1* variants yielded identical sequences; most (19/21) nucleotide changes were in nonsynonymous codons suggesting the occurrence of phenotypic selection; and most transition base substitutions involved purines (13/17) and not pyrimidines as is favored by Taq polymerase (20).

Lampe et al. (11) suggested recombination may be a genetic mechanism for *omp1* diversification. This suggestion was based upon detailed study of *C. trachomatis* strains that had serovar I sequences in VD1, 2, and 3, and serovar H sequence in VD4. Several of the major variants identified in the current study are also compatible with a recombinational mechanism including four isolates identical to the one studied by Lampe et al. (11). Logically, recombination should depend on co-infection of the host with multiple *C. trachomatis* serovars and interestingly, 6% of the identified infections in the present study involved multiple strains of *C. trachomatis*. Before concluding that the *omp1* mosaic sequences observed in the present study represent *omp1* recombinants, it is important to exclude PCR artifact. In

particular, the occurrence of multiple strain infection with two or more *omp1* alleles raises the possibility for at least two artifactual mechanisms for the generation of mosaicism during PCR amplification. Firstly, primer preference with differential amplification of *omp1* VDs during cycle sequencing may be able to generate apparent mosaic sequences. This was excluded by cloning *omp1* from the four major variants and resequencing VD1, 2, and 4. The results showed that each cloned variant *omp1* had identical mosaic sequences as initially identified in the clinical samples. Secondly, a different type of PCR artifact can give rise to mosaic molecules when related alleles are coamplified and incomplete PCR products are generated. Such incomplete products may hybridize to the alternate allelic template and be extended during subsequent PCR cycles, thus yielding mosaic molecules (21, 22). To explore this possibility, prototype serovar I and serovar H *omp1* were coamplified, subsequently cloned and 10 clones were sequenced. Six were serovar H, four were serovar I, and none was mosaic. Thus, under the experimental conditions used in the present study, mosaic *omp1* was not readily generated in the presence of two closely related *omp1* alleles.

The data from this study suggest that point mutation and perhaps recombination are involved in determining the repertoire for *omp1* polymorphism. The large number of polymorphic *omp1* isolates suggests that the tempo for genetic change in *omp1* is much greater than previously thought. We conclude that sequence variability in *omp1* may contribute to the ecologic success of *C. trachomatis*. Determination of the genetic mechanism for *omp1* diversification will be of considerable interest.

Acknowledgments

This work was supported by a grant from the Medical Research Council of Canada (SP27).

References

- Stephens, R. S., R. Sanchez-Pescador, E. A. Wagar, C. Inouye, and M. Urdea. 1987. Diversity of *Chlamydia trachomatis* major outer membrane protein genes. *J. Bacteriol.* 169:3879–3885.
- Baehr, W., Y. Zhang, T. Joseph, S. Hua, F. E. Nano, K. D. E. Everett, and H. D. Caldwell. 1988. Mapping antigenic domains expressed by *Chlamydia trachomatis* major outer membrane protein genes. *Proc. Natl. Acad. Sci. USA.* 85:4000–4004.
- Zhong, G., and R. C. Brunham. 1991. Antigenic determinants of the chlamydial major outer membrane protein resolved at a single amino acid level. *Infect. Immun.* 59:1141–1147.
- Morrison, R. P., D. S. Manning, and H. D. Caldwell. 1992. Immunology of *Chlamydia trachomatis* infections: Immunoprotective and immunopathogenic responses. In *Sexually Transmitted Diseases*. Vol. 8. T. C. Quinn, editor. Raven Press/New York, New York. 57–84.

5. Zhong, G., J. Berry, and R. C. Brunham. 1994. Antibody Recognition of a Neutralization Epitope on the Major Outer Membrane Protein (MOMP) of *Chlamydia trachomatis*. *Infect. Immun.* 62:1576–1583.
6. Kaltenboeck, B., K. G. Kousoulas, and J. Storz. 1993. Structures of and allelic diversity and relationship among the major outer membrane protein (ompA) genes of the four chlamydial species. *J. Bacteriol.* 175:487–502.
7. Fitch, W. M., E. M. Peterson, and L. M. de la Maza. 1993. Phylogenetic analysis of the outer membrane protein genes of chlamydiae and its implication for vaccine development. *Mol. Biol. Evol.* 10:892–913.
8. Dean, D., M. Patton, and R. S. Stephens. 1991. Direct sequence evaluation of the major outer membrane protein gene variant regions of *Chlamydia trachomatis* subtypes D', I', and L₂. *Infect. Immun.* 59:1579–1582.
9. Dean, D., J. Schachter, C. R. Dawson, and R. S. Stephens. 1992. Comparison of the major outer membrane protein variant sequence regions of B/Ba isolates: a molecular epidemiologic approach to *Chlamydia trachomatis* infections. *J. Infect. Dis.* 166:383–392.
10. Hayes, L. J., R. L. Bailey, D. C. W. Mabey, I. N. Clarke, M. A. Pickett, P. J. Watt, and M. E. Ward. 1992. Genotyping of *Chlamydia trachomatis* from a trachoma-endemic village in The Gambia by a nested polymerase chain reaction: Identification of strain variants. *J. Infect. Dis.* 166:1173–1177.
11. Lampe, M. F., R. J. Suchland, and W. E. Stamm. 1993. Nucleotide sequence of the variable domains within the major outer membrane protein gene from serovariants of *Chlamydia trachomatis*. *Infect. Immun.* 61:213–219.
12. Yang, C., I. Maclean, and R. C. Brunham. 1993. DNA sequence polymorphism of the *Chlamydia trachomatis* omp1 gene. *J. Infect. Dis.* 168:1225–1230.
13. Brunham, R. C., and F. A. Plummer. 1990. A general model of sexually transmitted disease epidemiology and its implications for control. *Med. Clin. North. Am.* 74:1339–1352.
14. Brunham, R. C., F. Plummer, and R. S. Stephens. 1993. Bacterial antigenic variation, host immune response and pathogen-host coevolution. *Infect. Immun.* 61:2273–2276.
15. Yuan, Y., Y. Zhang, N. G. Watkins, and H. D. Caldwell. 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.
16. Blanco, D. R., E. M. Walker, D. A. Haake, C. I. Champion, J. N. Miller, and M. A. Lovett. 1990. Complement activation limits the rate of in vitro treponemical activity and correlates with antibody-mediated aggregation of *Treponema pallidum* rare outer membrane protein. *J. Immunol.* 144:1914–1921.
17. Selander, R. K., D. A. Caugant, and T. S. Whittam. 1987. Genetic structure and variation in natural populations of *Escherichia coli*. In *Escherichia coli* and *Salmonella typhimurium*. Cellular and Molecular Biology, Volume 2. F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, and M. Schaechter, editors. American Society for Microbiology, Washington, D. C. 1625–1648.
18. Kunkel, T. A., and P. S. Alexander. 1986. The base substitution fidelity of eucaryotic DNA polymerases. *J. Biol. Chem.* 261:160–166.
19. Tindall, K. R., and T. A. Kunkel. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry.* 27:6008–6013.
20. Eckert, K. A., and T. A. Kunkel. 1990. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* 18:3739–3744.
21. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC)*. 239:487–491.
22. Meyerhans, A., Vartanian, J.-P., and S. Wain-Hobson. 1990. DNA recombination during PCR. *Nucleic Acids. Res.* 18:1687.