Vol. 71, No. 3

Polymorphic Membrane Protein H Has Evolved in Parallel with the Three Disease-Causing Groups of *Chlamydia trachomatis*

Diane R. Stothard,* Gregory A. Toth, and Byron E. Batteiger

*The Department of Medicine, Division of Infectious Diseases, Indiana University
School of Medicine, Indianapolis, Indiana 46202*

Chlamydia trachomatis **is a human pathogen causing trachoma, urogenital disease, and lymphogranuloma venereum (LGV). A family of nine polymorphic membrane protein genes (*pmpA* to *pmpI*), resembling autotransporter proteins, has recently been discovered in *C. trachomatis*. *pmp* genes are large and predicted to be outer membrane proteins. We hypothesized that they would contain useful nucleotide sequence variability for epidemiologic studies. Since sequence information is available only for serovars D and L2, we sought to determine the amount of diversity within an individual *pmp* gene among serovars. We used restriction fragment length polymorphism (RFLP) analysis as a primary screen to assess the amount of sequence divergence among the *pmp* genes for serovars A to L3 of *C. trachomatis*. RFLP analysis showed little variation for some of the genes, such as *pmpA*, but substantial variation in others, such as *pmpI*. *pmpH* and *pmpE* yielded RFLP patterns that clustered the 15 serovars into ocular, urogenital, and LGV groups, and both proteins have been localized to the outer membrane. Therefore, we chose to sequence *pmpE*, *pmpH*, and *pmpI* from each of the 15 serovars. Evolutionary analysis showed three distinct divergence patterns. PmpI was least variable, resulting in an ambiguous evolutionary pattern. PmpE showed a high degree of diversity in the ocular strains compared to the other strains. Finally, the evolution of PmpH shows three groups that reflect disease groups, suggesting this protein may play a role in pathogenesis.**

Chlamydia trachomatis is an obligate intracellular human pathogen that causes trachoma, urogenital disease, and lymphogranuloma venereum (LGV). The strains of *C. trachomatis* are divided into at least 18 different serovars based on immunological cross-reactivity of the major outer membrane protein (MOMP) (30, 31, 32). Additionally, the serovars of *C. trachomatis* are classified into two groups, the trachoma biovar and the LGV biovar, based upon differences in disease presentation and in vitro cell growth characteristics. The trachoma biovar includes serovars that cause ocular trachoma (A, B, Ba, and C) and serovars that cause urogenital disease (D through K, Da, and Ia), while the LGV biovar includes serovars L1, L2, L2a, and L3.

A superfamily of genes encoding predicted polymorphic membrane proteins (*pmp*) was discovered as part of the *Chlamydia* genome project (16, 23, 26). Several proteins encoded by these genes had previously been investigated in *C. psittaci* (6, 11, 12, 20, 25). Nine genes are found in *C. trachomatis* (*pmpA* to *pmpI*) (26). The Pmp proteins are large proteins, 90 to 187 kDa in mass, and they have no obvious homologs in nonchlamydial species (29). Transcription of all nine genes has been shown for both biovars in serovars D and L2 (19). Of the nine genes, all but *pmpA* have an apparent signal peptide leader sequence and are predicted outer membrane proteins (13). While expression has not been verified for all nine proteins in all serovars, expression of PmpE, G, and H, and local-

ization of these three proteins to the outer membrane of *C. trachomatis* serovar L2 has been documented (21, 28).

The *pmp* genes are highly diverse in nucleotide sequence. None of the *pmp* genes are more than 52% similar to each other (13). However, the amount of sequence diversity within an individual *pmp* gene among the different serovars has not been investigated since sequence information is available only for serovars D and L2. Considering that the Pmp proteins are large, and thought to be expressed on the outer membrane (28, 29), we hypothesized that *pmp* genes might contain useful variability for epidemiologic studies. Currently, the MOMP gene serves as a molecular marker in epidemiologic studies because of its high degree of sequence variability. However, the MOMP gene for some serovars, like E and F, exhibits a high degree of conservation at the nucleotide level, decreasing its usefulness in epidemiologic studies (9, 27). Other markers, like the *pmp*, could improve discrimination between strains. Therefore, we examined the *pmp* genes as possible molecular epidemiologic markers by determining the amount of sequence diversity within the *pmp* genes among serovars.

We used restriction fragment length polymorphism (RFLP) analysis as an initial screen to assess the amount of sequence variability among the *pmp* genes of 15 serovars of *C. trachomatis*. RFLP analysis showed limited to no variation among serovars for some of the *pmp* genes, such as *pmpA*, but substantial variation was seen for other genes, such as *pmpI*. Interestingly, *pmpH* yielded RFLP patterns that clustered the 15 serovars into ocular, urogenital, and LGV disease groups. The RFLP patterns for *pmpE* and *pmpF* also showed potential segregation of serovars by disease group. However, unlike PmpE, expression of PmpF has not been demonstrated. Therefore, we chose to sequence *pmpE*, *pmpH*, and *pmpI* from 15

* Corresponding author. Mailing address: Department of Medicine, Division of Infectious Diseases, Indiana University School of Medicine, 435 Emerson Hall, 545 Barnhill Dr., Indianapolis, IN 46202. Phone: (317) 278-1791. Fax: (317) 274-1587. E-mail: dstothar@iupui.edu.

TABLE 1. *pmp* PCR primers

Primer	Sequence	Nucelotide location[a] and direction
pmpA1-FOR	5′-ATG AAT CGA GTT ATA GAA ATC CAT GC	145, forward
pmpA1-REV	5′-TTA GAA GCC AAC ATA GCC TCC G	3061, reverse
pmpC1-FOR	5′-GAC TGC AAT GTT AGC AAA TTA GGA TA	94, forward
pmpC1-REV	5′-CAT ACG AGC ACC GCA GTT CA	5288, reverse
pmpD1-FOR	5′-CCT CAA GCG GTT TTA TTG TTA GAC C	181, forward
pmpD1-REV	5′-CAA TCC AGT ATT CGC CTC ATA TCC T	4590, reverse
pmpE1-FOR	5′-CTT ATG CCC AAC TCA GTT CCA GA	151, forward
pmpE1-REV	5′-TCG CAG AGC AAT TTC CCC AT	2942, reverse
pmpF1-FOR	5′-ATG AAA CGG ATA CGC TAC AGT TC	83, forward
pmpF1-REV	5′-GAC CAG AGC TCC TCC TGC AT	3086, reverse
pmpG1-FOR	5′-CAA GGA ATT TAC GAT GGG GAG A	127, forward
pmpG1-REV	5′-TCC TGC ACT CAA ACC ATA ACC	3031, reverse
pmpH N1	5′-TCT CCT CAA GTG TTA ACA CCT AAT G	76, forward
pmpH1-REV	5′-AAA GAT TCT ATT CAA GCC CAT GG	3035, reverse
pmpI1-FOR	5′-CAG GAT CCC TTA GGT GAA ACC G	73, forward
pmpI1-REV	5′-CAG GGA GTA ACT ATG GCT TTG CC	2610, reverse

[a] Primer locations indicate base position of each primer in serovar D for each *pmp* gene.

serovars of *C. trachomatis*. We chose *pmpE* and *pmpH* because of their provocative RFLP patterns and demonstration of their surface location and *pmpI* because of its apparent nucleotide diversity.

MATERIALS AND METHODS

Strains and isolates. Elementary bodies (EBs) were partially purified (22) by differential centrifugation followed by density gradient centrifugation in Percoll from the following laboratory strains: A/571-B/OT, B/TW-5/OT, Ba/Ap-2/OT, C/TW-3/OT, D/UW-3/Cx, E/UW-5/Cx, F/UW-6/Ur, G/UW-57/Cx, H/UW-4/Cx, I/UW-12/Ur, J/UW-36/Cx, K/UW-53/Cx, L1/440, L2/434, and L3/404.

DNA extraction. DNA was extracted from purified EBs by using previously described methods (27). The only modification was the addition of dithiothreitol (4 mM final concentration) before boiling to assist in disruption of EB outer membrane proteins. DNA from purified EBs was resuspended into 50 μl of Tris buffer (10 mM Tris-Cl [pH 8.5]).

***pmp* PCR amplification.** The *pmp* genes were amplified from *C. trachomatis* serovars A to L3 by using PCR primers (Table 1) based on the published serovar D sequences. We were unable to amplify the *pmpB* gene from any serovar after trying three separate sets of primers. All PCR products were of the expected sizes based on the published serovar D genes (26).

DNA amplification was done with the Expand high-fidelity kit (Roche Molecular Biochemicals, Indianapolis, Ind.). Reaction mixtures were 50 μl and consisted of 5 μl of template DNA, 15 pmol of each primer, 1.5 mM MgCl$_2$, 200 μM concentrations of each deoxynucleoside triphosphate, and 0.75 μl of Expand polymerase in a 1× reaction buffer. Cycling conditions began with an initial 5-min denaturation step at 95°C followed by 35 cycles of denaturation (95°C, 30 s), annealing (58°C, 30 s), and extension (68°C, 4 min). An additional 10-min extension at 68°C was performed at the end of the 35 cycles.

RFLP analysis. PCR products were purified with the QIAquick PCR purification kit (QIAGEN). Products were eluted from QIAquick columns into 50 μl of elution buffer (10 mM Tris [pH 8.5]). Restriction endonuclease reactions were carried out according to the manufacturer's instructions. Reaction mixtures were

TABLE 2. *pmpE*, *pmpH*, and *pmpI* internal sequencing primers

Primer	Sequence (5′–3′)	Nucleotide location[c] and direction
pmpE-560	TAC CTT TGT TGT GAG CGA GAA TC	560, forward
pmpE-570R	AAC AAG ACT GAT TCT CGC TCA C	570, reverse
pmpE-1259	CAA TAA GGA AGC TGA TCA AAC AG	1259, forward
pmpE-1510R	CCW ATA TGA TTM AGT GTT AYA GAG G	1510, reverse
pmpE-2359R	TCT CCT TGR GTA TAG AAA TGG TGA C	2359, reverse
pmpE-2499	GCT CTT GGT ATT TAT TCT AGC CTG	2499, forward
pmpH-443REV	GGC ACA ATA GAC AAA GGA GAA TAC C	443, reverse
pmpH-1026	TCT TGT AGC AGA TGA TTC GGT TGT C	1026, forward
pmpH-1097REV[a]	CCA TTC GTG CTC ATA TTG ATA GTA G	1097, reverse
pmpH-1174REV[a]	ACC AGA GCT TGC TTC ACT CAC	1174, reverse
pmpH-1914[a]	CAC TAT TAA TGA TGC AGA TGC TGC	1914, forward
pmpH-2011[a]	GTA CCT CCT AAT ACC AAT AAC ACT CTG	2011, forward
pmpH-2188REV	CAT GCA GAG ATG CTA CAA ATC CAA C	2188, reverse
pmpH-2619	TCG TCA GTT CGC TAG AGA TTA TGA G	2619, forward
pmpI-435R	GCA GAG ATG GCT CCT CCA GAG	435, reverse
pmpl-627	CGG CGC TAT TTG CTG TAG TAA TC	627, forward
pmpI N2[b]	<u>TCC CCC GGG</u> CGT GAC TTC TCC CAC CCC	1118, forward
pmpI C2[b]	<u>AGA AAG CTT</u> AGG GGC GTG GAT AGT TAC	1438, reverse
pmpI-1659	TTC TTG GAA AGA TTC TGA TGA AGG G	1659, forward
pmpI-1729REV	ACG AGT GTA GAT TGA CGT TCT GG	1729, reverse
pmpI-2301	TGA GGA GAG TTC GGG AGA GAT TC	2301, forward

[a] pmpH-1174REV and pmpH-2011 did not work on all strains due to mismatches in those regions with LGV and ocular strains. Therefore, pmpH-1097REV and pmph-1914 were made as alternate primers.
[b] Underlined bases in pmpI N2 and pmpI C2 indicate restriction sites added to the primer and are not part of the *pmpI* gene.
[c] Primer locations indicate base position of each primer in serovar D for each *pmp* gene.

TABLE 3. RFLP analysis of *pmp* genes from *C. trachomatis* serovars A to L3

pmp locus	Enzyme	No. of expected sites[a]	RFLP groupings (serovars with the same RFLP pattern)
A	*Cfo*I	3	No variation
	*Hin*fI	12	No variation
	*Taq*I	12	(ABBaCDEFGHI and LGV) (JK)
C	*Taq*I	16	(ABCEF and LGV) (Ba) (DGHIJK)
	*Tru*9I	19	(ABBaC and LGV) (EF) (DGH) (IJK)
D	*Cfo*I	10	No variation
	*Hin*fI	17	(B) (C) (EF) (ABaDGHIJK and LGV)
E	*Tru*9I	9	(ABBaC) (DEFGH) (IJK and LGV)
	*Hin*fI	14	(ABBaC) (DEFGH) (IJK) (LGV)
F	*Hin*fI	10	(ABBaC) (DEFGH) (IJK) (LGV)
	*Cfo*I	5	(ABBaC) (DEFGH) (IJK) (LGV)
	*Taq*I	6	(ABa) (BC) (DEFGH) (IJK) (LGV)
G	*Cfo*I	6	No variation
	*Taq*I	11	(BC) (ABaDEFGHIJK) (LGV)
	*Hin*fI	15	(ABBaCI) (DEFGHJK) (LGV)
H	*Cfo*I	7	(ABBaC) (DEFGHIJK) (LGV)
	*Tru*9I	16	(ABBaC) (DEFGHIJK) (LGV)
	*Taq*I	14	(ABBaC) (DEFGH) (IJK) (L1, L2) (L3)
I	*Cfo*I	10	(ABBaCDEFGHIJK) (LGV)
	*Hin*fI	12	(ABC) (Ba) (EF) (DGHIJK) (LGV)
	*Taq*I	6	(ABaC) (B) (DEFGHI and LGV) (JK)

[a] The number of expected sites is based upon published serovar D sequences for each *pmp* gene.
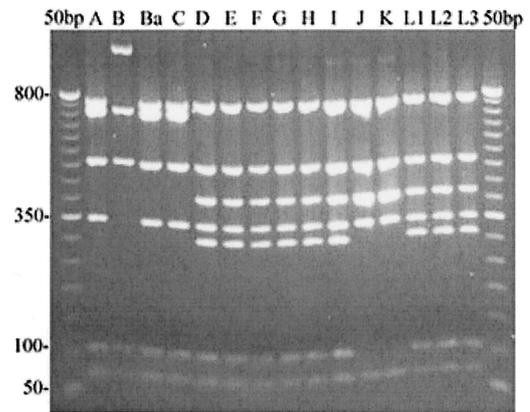

FIG. 1. RFLP analysis suggests *pmpI* is most variable. *pmpI* PCR products amplified from 15 serovars (indicated above their respective lanes) of *C. trachomatis* digested with *Taq*I endonuclease and separated on a 2% agarose gel are shown. Numbers on the left are size standards in base pairs.

10 μl in volume and consisted of 5 μl of *pmp* PCR product and 1 to 5 U of enzyme in a 1× reaction buffer. Computer-generated restriction digests of the *pmp* genes from serovar D (Clone Manager, version 4.01; Scientific and Educational Software) were used to determine which enzymes would be appropriate for RFLP analysis and included *Cfo*I, *Bam*HI, *Hin*fI, *Taq*I, and *Tru*9I (all from Roche Molecular Biochemicals). PCR products were digested with at least two different restriction endonucleases that cleaved from 3 to 19 times within each gene. The restriction products were then electrophoresed through 2 to 4% agarose gels, stained with ethidium bromide, and digitally photographed with a UV transilluminator.

DNA sequencing. PCR products for *pmpE*, *pmpH*, and *pmpI* were each amplified in duplicate and pooled for sequence analysis to insure sequence fidelity. All sequencing reactions were done with Big Dye terminator chemistry and an ABI Prism 310 automated DNA sequencer (Applied Biosystems, La Jolla, Calif.). To obtain full-length sequences from *pmpE*, *pmpH*, and *pmpI*, several internal primers were used in addition to the PCR amplification primers (Table 2). Sequences were collected from both strands in order to resolve ambiguities and ensure consistency. In addition, electropherograms were manually checked for computer miscalls.

Nucleotide and amino acid sequence alignment. Sequences were aligned with the assistance of the program ESEE (5) based on both nucleotide and amino acid residues. In regions of high variability, the alignment was based primarily on the most-conservative amino acid substitutions and secondarily on nucleotide substitutions.

The primers used to amplify *pmpE* extended from nucleotide 80 of the coding region to nucleotide 2881 based on the serovar D *pmpE* gene. Therefore, the first 79 bases and the last 15 bases are not included in the alignment. The sequence alignment for *pmpE* contained 2,811 nucleotide sites (nucleotides plus alignment gaps) and 936 amino acid sites (amino acids plus alignment gaps).

The primers used to amplify *pmpH* extended from nucleotide 76 of the coding region to nucleotide 3051 (terminal stop codon) based on the serovar D *pmpH* gene. Therefore, the first 75 bases are not included in the alignment. The sequence alignment for *pmpH* contained 2,991 nucleotide sites (nucleotides plus alignment gaps) and 996 amino acid sites (amino acids plus alignment gaps).

The primers used to amplify *pmpI* extended from nucleotide 73 of the coding region to nucleotide 2610 based on the serovar D *pmpI* gene. Therefore, the first 72 bases and the last 27 bases are not included in the alignment. The sequence alignment for the *pmpI* gene contained 2,538 nucleotides and 846 amino acids. There were no in/del events in *pmpI*.

Nucleotide and amino acid comparisons. After sequences were aligned, nucleotide and amino acid sequences were compared manually in a pairwise fashion for differences. Contiguous in/dels at the nucleotide and amino acid levels were counted as individual differences, i.e., a deletion of six nucleotides between two pairs was counted as six differences rather than one deletion event, and the resulting two-amino-acid deletion was counted as two deletions rather than as one deletion event.

Phylogenetic analysis. (i) Nucleotide sequences. Because the 5′ and 3′ terminal sequences of the *pmp* genes are primer defined, those regions were excluded from the phylogenetic analyses. Consequently, 2,783 of 2,811 nucleotides were compared for *pmpE*, 2,940 of 2,991 nucleotides were compared for *pmpH*, and 2,493 of 2,538 nucleotides of *pmpI* were compared.

All phylogenetic and evolutionary analyses were conducted with MEGA (molecular evolutionary genetics analysis), version 2.1 (17). Once aligned, the corrected nucleotide divergence of sequences was calculated in MEGA, incorporating the Kimura two-parameter correction for multiple substitutions at a single

TABLE 4. Sequence diversity in the *pmpI* gene among serovars of *C. trachomatis*

Serovar	No. of dissimilarities with serovar[a]:														
	A	B	Ba	C	D	E	F	G	H	I	J	K	L1	L2	L3
A		1	2	2	5	6	5	6	5	5	9	10	6	6	6
B	5		3	1	6	7	6	7	6	6	8	9	5	5	5
Ba	2	7		4	7	8	7	8	7	7	11	12	8	8	8
C	3	4	5		7	8	7	8	7	7	9	10	6	6	6
D	8	13	10	11		7	6	1	0	0	4	5	7	7	7
E	10	15	12	13	14		1	8	7	7	11	12	8	8	8
F	9	14	11	12	13	1		7	6	6	10	11	7	7	7
G	9	14	11	12	1	15	14		1	1	5	6	8	8	8
H	8	13	10	11	0	14	13	1		0	4	5	7	7	7
I	6	11	8	9	2	12	11	3	2		4	5	7	7	7
J	15	16	17	14	9	21	20	10	9	9		1	9	9	9
K	16	17	18	15	10	22	21	11	10	10	1		10	10	10
L1	19	20	21	18	23	23	22	24	23	21	25	27		0	0
L2	18	19	20	17	22	22	21	23	22	20	25	26	1		0
L3	18	19	20	17	22	22	21	23	22	20	25	26	1	0	

[a] Values below the diagonal are nucleotide dissimilarities out of 2,538 nucleotides matched. Values above the diagonal are amino acid differences out of 846 amino acids matched. Within-disease group comparisons can be made for serovars A to C, D to K, and L1 to L3.
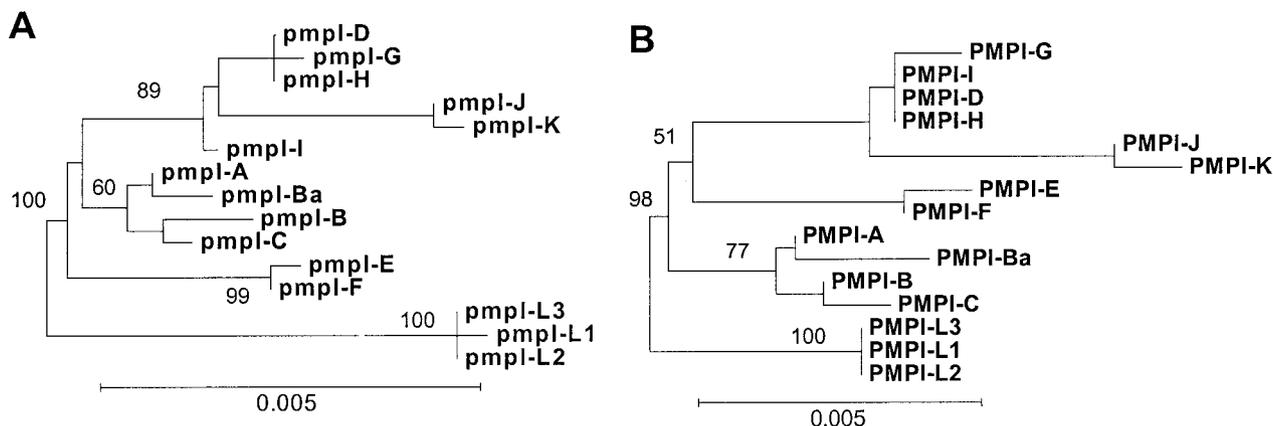
FIG. 2. Evolution of *pmpI* shows weak segregation of serovars compatible with disease groups. Neighbor-joining trees based on the nucleotide (A) and inferred amino acid (B) sequences of *pmpI* are shown. Bootstrap values are shown only at the nodes separating major clades and only when they exceeded 50%. The scale bar equals a corrected sequence divergence of 0.005, which is the same as the scale used in Fig. 4 and 6. However, due to lesser divergence in *pmpI*, the branches and scale bar here are lengthened in order to plainly illustrate the branching patterns.

site. The branching pattern of taxa was estimated from the distance matrix by the neighbor-joining method (24) as implemented in MEGA. A cladistic reconstruction of the phylogenetic relationships among taxa also was performed by using maximum parsimony as implemented in MEGA with close-neighbor interchange. Confidence levels for the branching pattern were estimated by a bootstrap resampling of the data. Bootstrap values for the trees were obtained from a consensus tree based on 1,000 replications. Although *pmp* genes have been identified in *C. trachomatis*, *Chlamydia pneumoniae*, and *Chlamydia psittaci*, no homologues exist for *C. trachomatis pmp* genes. Therefore, an appropriate outgroup could not be identified, and the trees were not rooted.

(ii) Amino acid sequences. Because 5′ and 3′ terminal sequences of the *pmp* genes are primer defined, 927 of 936 amino acids of PmpE, 974 of 996 amino acids of PmpH, and 831 of 846 amino acids of PmpI were included in the analyses.

All phylogenetic and evolutionary analyses were conducted using MEGA, version 2.1 (17). Phylogenetic reconstructions based on the inferred amino acid sequences were performed by using both distance and parsimony methods as implemented in MEGA. As with the nucleotide data, bootstrap values for the trees were obtained from a consensus tree based on 1,000 randomly generated datasets and the trees were not rooted.

RESULTS

RFLP analysis shows differing degrees of variation within the *pmp* gene family. RFLP results are summarized in Table 3. Serovars that are grouped within parentheses in the table had the same RFLP pattern for the particular enzyme listed. For example, digestion of *pmpA* with *Taq*I produced two RFLP patterns, dividing the 15 serovars into two RFLP groups. One group consisted of serovars J and K while the remaining sero-

vars had a different RFLP pattern. Interestingly, *pmpA*, the only *pmp* gene without a predicted leader peptide sequence, exhibited the least amount of variation, with two of the three enzymes demonstrating no differences among serovars. Variation in RFLP patterns was seen for the other *pmp* genes, but the amount of variation differed.

Digestion of *pmpF* with three different enzymes produced nearly the same patterns, and the patterns indicate a segregation of the ocular, urogenital, and LGV serovars based on this gene. *Hin*fI and *Cfo*I digests were identical, producing the same four RFLP groups: one containing the ocular serovars; one containing D through H; one including I, J, and K; and one with the LGV serovars. The *Taq*I digest exhibited some variation within the ocular serovars, producing one pattern for serovars A and Ba and another pattern for serovars B and C.

A very similar arrangement of serovars was seen for *pmpE* digests. Digestion of *pmpE* with *Hin*fI produced the same serovar groupings as seen for the *Hin*fI and *Cfo*I digests of *pmpF*. *Tru*9I produced a similar pattern, the only difference being that LGV serovars also had the same *Tru*9I pattern as serovars I, J, and K.

The most interesting RFLP pattern was seen for *pmpH*. Two of the three enzymes used produced an arrangement of serovars for this gene that mirrors the three disease groups of *C. trachomatis*. Both *Tru*9I and *Cfo*I produced three RFLP groups, one containing the ocular serovars, one containing the urogenital serovars, and one containing the LGV serovars. The *Taq*I digest showed some variation within the urogenital and LGV groups, but for all three enzymes serovars in different disease groups never shared the same RFLP pattern.

RFLP analysis of *pmpI* suggests that it is the most variable *pmp* gene, but sequence analysis shows *pmpI* to be highly conserved. *pmpI* exhibited the most variation of any of the *pmp* genes by RFLP analysis (Fig. 1). Three different enzymes were used to digest *pmpI*, and each enzyme produced a different set of RFLP patterns. Since *pmpI* exhibited interserovar variation, we chose to search for intraserovar variation in *pmpI* by using 13 serovar E clinical isolates. We selected serovar E isolates because the MOMP gene is remarkably conserved among se-
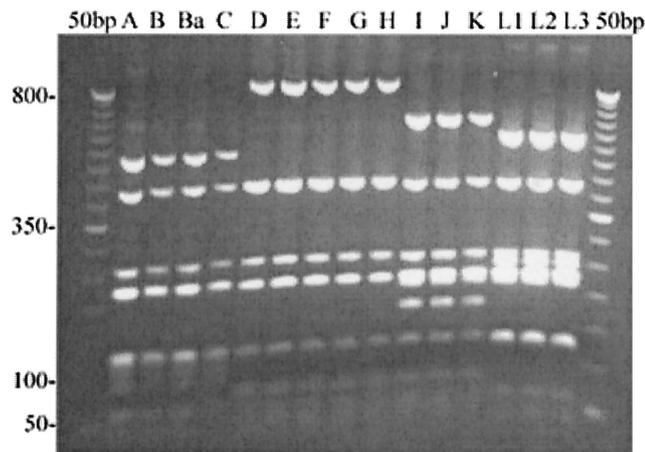
FIG. 3. PmpE digested with *Hin*fI restriction endonuclease shows segregation of ocular, urogenital, and LGV serovars. An 8-μl aliquot of the *pmpE* PCR products was digested with *Hin*fI endonuclease, and restriction fragments were separated on a 2% agarose gel. Serovars are indicated above their respective lanes. Numbers on the left are size standards in base pairs.

rovar E isolates (9, 27). However, no variation was evident in the *pmpI* gene from the 13 clinical isolates by using either *Taq*I or *Cfo*I (data not shown).

Contrary to the RFLP results, sequences from 15 serovars showed that *pmpI* is highly conserved, with no more than 1% sequence dissimilarity across all serovars, including the LGV strains (Table 4). There also were no insertion or deletion (in/del) events in *pmpI* among the serovars. The sequence for *pmpI* from serovar D/UW3 was identical to the sequence deposited at the *Chlamydia* genome website for *pmpI* from serovar D (http://chlamydia-www.berkeley.edu:4231/).

Phylogenetic analyses of the nucleotide and amino acid sequences of *pmpI* did not produce highly stable trees (Fig. 2), likely due to the lack of diversity among the serovars. Neighbor-joining and parsimony trees based upon the nucleotide sequence of *pmpI* were nearly identical; Fig. 2A shows the neighbor-joining tree. Branch lengths are proportional to distances between taxa, but because there is so little diversity in *pmpI*, the scale bar and branch lengths in Fig. 2 are inflated (compared to trees drawn for *pmpE* and *pmpH*) in order to clearly illustrate branching patterns. Evolutionary analysis of the inferred amino acid sequences of PmpI produced a slightly different tree (Fig. 2B). In this tree, serovars diverge into three clades consistent with disease groups, whereas in the nucleotide tree, serovars E and F diverge prior to a clade composed of a mixture of ocular and urogenital serovars.

Sequence analysis of *pmpE* indicates substantial genetic evolution in the ocular serovars for this protein. The RFLP pattern for *pmpE* (Table 3) with *Hin*fI (Fig. 3) was identical to those seen for *pmpF*, showing a separation of ocular, urogenital, and LGV serovars but with additional differences in serovars I, J, and K that separate them from the rest of the urogenital serovars. Since PmpE has been shown to be surface exposed (28), we obtained sequence data for *pmpE* from the 15 serovars.

The nucleotide and inferred amino acid sequences of *pmpE* were nearly identical among serovars within a disease group,

with less than 0.5% dissimilarity, but they were more different when comparing serovars across disease groups (Table 5). Unlike PmpI, five in/del events have occurred in PmpE, all occurring in the N-terminal portion of the protein, between amino acids 147 and 533 (based on serovar D/UW3). The events range from one to four amino acids in length, and in three of the five events, the ocular serovars contain an in/del relative to the other 11 serovars.

Overall, sexually transmitted disease (STD) serovars differed from LGV serovars by 1.6 to 2.5% at the amino acid level and 1.3 to 2.5% at the nucleotide level (Table 5). There was also an increased amount of divergence seen in serovars I, J, and K compared to the other STD serovars, again consistent with the RFLP patterns. At the sequence level, the most remarkable thing about *pmpE* was the amount of diversity in the ocular serovars compared to the two other groups. Ocular serovars exhibited three times more variation in *pmpE* than the other two disease groups: 5.1 to 6.5% at the nucleotide level and 6.1 to 7.0% at the amino acid level (Table 5). The sequence for *pmpE* from serovar D/UW3 was identical to the sequence deposited at the *Chlamydia* genome website for *pmpE* from serovar D (http://chlamydia-www.berkeley.edu:4231/).

Neighbor-joining and parsimony analyses of the nucleotide sequences of *pmpE* produced the same tree (Fig. 4A). Branch lengths are proportional to nucleotide divergence, and while the scale bar represents the same amount of divergence as seen in Fig. 2, note that the actual bar is smaller due to the increased amount of divergence in *pmpE* compared to *pmpI*. The evolutionary analysis of *pmpE* nucleotide sequences was consistent with the RFLP patterns seen in Fig. 3 and Table 3, with serovars I, J, and K forming a distinct group separate from the rest of the urogenital serovars. The evolutionary tree obtained using the amino acid sequences of PmpE (Fig. 4B) was the same as the nucleotide tree. Most remarkable, however, is the long branch separating the ocular serovars from the remaining serovars, suggesting either an increased rate of evolution in the

TABLE 5. Sequence diversity in the *pmpE* gene among serovars of *C. trachomatis*

| Sero-var | No. of dissimilarities with serovar[a]: | | | | | | | | | | | | | | |
	A	B	Ba	C	D	E	F	G	H	I	J	K	L1	L2	L3
A		4	4	4	64	63	63	64	64	58	57	57	66	66	63
B	6		0	2	63	62	62	63	63	57	56	56	63	63	60
Ba	4	2		2	63	62	62	63	63	57	56	56	63	63	60
C	6	4	4		63	62	62	63	63	57	56	56	63	63	60
D	179	181	179	180		1	1	0	0	22	21	21	22	22	19
E	175	177	175	176	4		0	1	1	21	20	20	21	21	18
F	175	177	175	176	4	0		1	1	21	20	20	21	21	18
G	179	181	179	180	0	4	4		0	22	21	21	22	22	19
H	179	181	179	180	0	4	4	0		22	21	21	22	22	19
I	145	147	145	146	69	65	65	69	69		1	1	19	19	16
J	144	146	144	145	69	65	65	69	69	2		0	18	18	15
K	144	146	144	145	68	64	64	68	68	1	1		18	18	15
L1	163	163	161	162	57	53	53	57	57	39	39	38		0	3
L2	164	164	162	163	58	54	54	58	58	40	40	39	1		3
L3	161	161	159	160	52	48	48	52	52	37	37	36	8	9	

[a] Values below the diagonal are nucleotide dissimilarities out of 2,783 nucleotides matched. Values above the diagonal are amino acid differences out of 927 amino acids matched. Within-disease group comparisons can be made for serovars A to C, D to K, and L1 to L3.
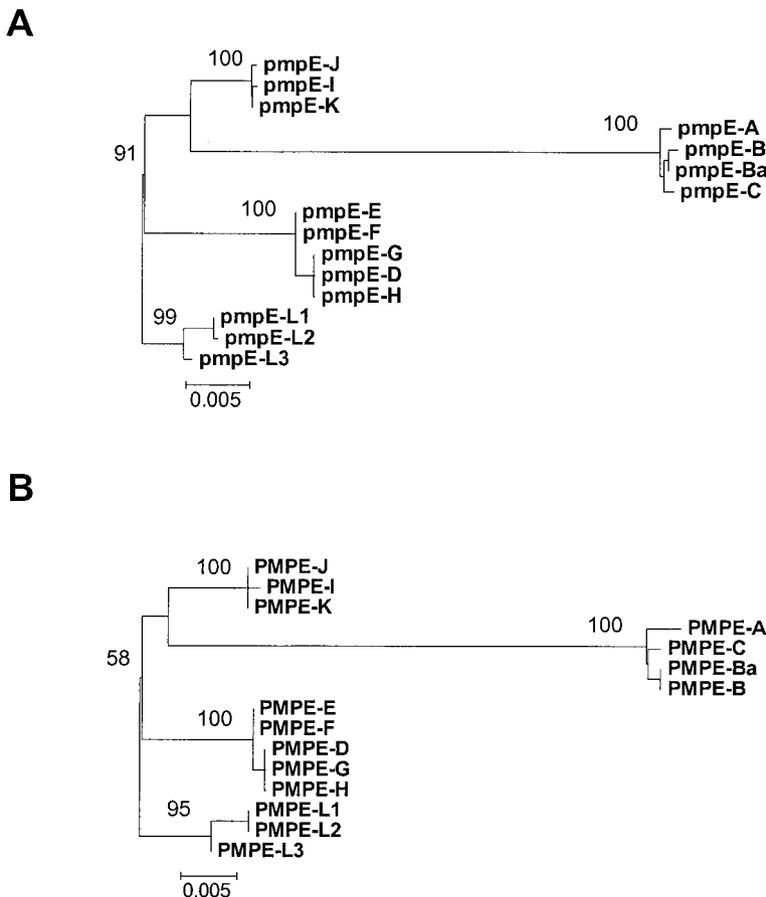
FIG. 4. Sequence analysis of *pmpE* shows an extraordinary amount of evolution in ocular serovars. Phylogenetic trees based on the nucleotide (A) and inferred amino acid (B) sequences of *pmpE*. Bootstrap values are shown only at the nodes separating major clades and only when they exceeded 50%. The scale bar equals a corrected sequence divergence of 0.005, which is the same as the scale used in Fig. 2 and 6. However, because there is significantly more divergence in *pmpE* and *pmpH* than in *pmpI*, the actual bars here and in Fig. 6 are similar in length compared to the elongated bar used in Fig. 2.

ocular serovars or a deceleration in the nonocular strains for this protein.
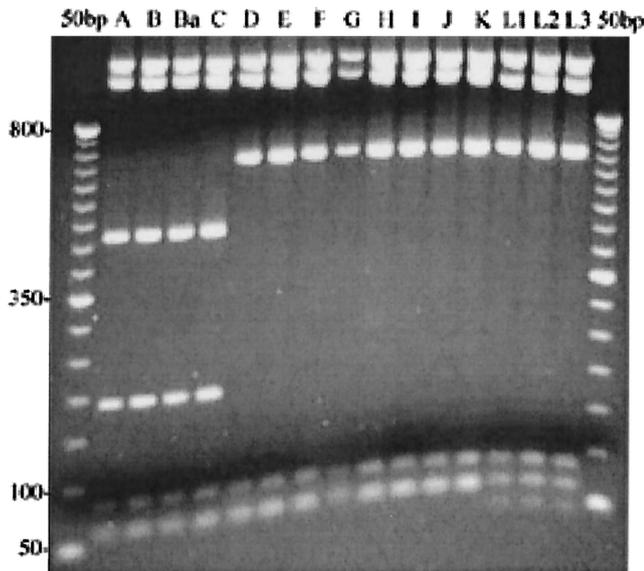
RFLP and sequencing analysis illustrates *pmpH* evolution parallels the separation of *C. trachomatis* serovars into three disease groups. Digestion of *pmpH* (Table 3) with *Cfo*I (Fig. 5) and *Tru*9I divided the serovars into three groups compatible with the three disease groups of *C. trachomatis*. Because no other gene in *C. trachomatis* has exhibited such a divergence pattern, we obtained sequences of *pmpH* for 15 serovars.

The nucleotide and inferred amino acid sequences of *pmpH* were nearly identical among serovars within a disease group but quite different for serovars compared across disease groups. For instance, ocular serovars differed from LGV serovars by approximately 8% and from urogenital serovars by about 6% (Table 6). However, compared to each other, ocular serovars differed by no more than 0.1%. Similarly, urogenital serovars differed from each other by no more than 0.7%, and LGV serovars differed from each other by no more than 0.2% (Table 6). The sequence for *pmpH* from serovar D/UW3 was identical to the sequence deposited at the *Chlamydia* genome website for *pmpH* from serovar D (http://chlamydia-www.berkeley.edu :4231/).

There were six in/del events in *pmpH*, all of which occur in

the first ~1,800 bp of the gene. The most notable was a large 36-bp deletion in the LGV serovars compared to serovars A through H around position 480 in the gene (Table 7). Serovars I, J, and K also have a 6-bp deletion in that same region compared to serovars A through H. The other in/dels found in the gene were small, only 3 or 6 bases (1 to 2 amino acids) in length.

Neighbor-joining and parsimony analyses of the nucleotide sequences of *pmpH* produced the same tree (Fig. 6A). Branch lengths are proportional to nucleotide divergence, and the scale bar measures the same amount of divergence as the scale bars in Fig. 2 and 4, although the amount of diversity in *pmpH* is more similar to that seen in *pmpE* than to that seen in *pmpI*. The evolutionary analysis of *pmpH* nucleotide sequences was consistent with the RFLP patterns seen in Fig. 5 and Table 3. The evolutionary tree obtained by using the amino acid sequences of PmpH (Fig. 6B) was slightly different than the nucleotide tree. In the nucleotide tree, serovars E and F diverge prior to the DGHIJK cluster, but in the amino acid tree, E and F are sister taxa to DGH, with IJK diverging first. In addition, there was some rotation within the ocular clade when comparing the neighbor-joining tree to the parsimony tree based on amino acid se-

FIG. 5. RFLP analysis of *pmpH* yields three groups consistent with the three disease groups of *C. trachomatis*. An 8-μl aliquot of the *pmpH* PCR products was digested with *Cfo*I endonuclease, and restriction fragments were separated on a 4% Metaphor agarose gel. In the LGV strains, the loss of the 54-bp fragment is difficult to see because it comigrates with the 58-bp fragment in the other serovars, and the new 15-bp band in the LGV strains is not visible. However, the new 39-bp fragment in the LGV strains is faintly visible below the 58-bp fragment. Serovars are indicated above their respective lanes. Numbers on the left are size standards in base pairs.

quences, but no single pattern was well supported by bootstrap analysis. Most essential, in all four analyses of *pmpH*, the 15 serovars reliably formed three clades that were congruent with the three disease groups of *C. trachomatis*: ocular disease, urogenital disease, and LGV.

TABLE 6. Sequence diversity in the *pmpH* gene among serovars of *C. trachomatis*

Sero-var	No. of dissimilarities with serovar[a]:														
	A	B	Ba	C	D	E	F	G	H	I	J	K	L1	L2	L3
A		1	1	1	47	47	48	47	47	46	47	47	79	78	80
B	4		0	0	47	46	47	46	46	46	47	47	79	78	80
Ba	2	2		0	47	46	47	46	46	46	47	47	79	78	80
C	2	2	0		47	46	47	46	46	46	47	47	79	78	80
D	168	168	166	166		2	3	0	0	3	4	4	58	57	59
E	169	169	167	167	11		1	2	2	5	6	6	58	57	59
F	170	170	168	168	12	1		3	3	6	7	7	59	58	60
G	168	168	166	166	0	11	12		0	3	4	4	58	57	59
H	168	168	166	166	0	11	12	0		3	4	4	58	57	59
I	169	170	168	168	8	19	20	8	8		1	1	55	54	56
J	171	172	170	170	10	21	22	10	10	2		0	56	55	57
K	170	171	169	169	9	20	21	9	9	1	1		56	55	57
L1	247	248	246	246	215	210	211	215	215	207	209	208		2	1
L2	245	246	244	244	211	206	207	211	211	203	205	204	6		3
L3	247	248	246	246	215	210	211	215	215	207	209	208	2	6	

[a] Values below the diagonal are nucleotide dissimilarities out of 2,991 nucleotide sites matched. Values above the diagonal are amino acid differences out of 996 amino acid sites matched. Within-disease group comparisons can be made for serovars A to C, D to K, and L1 to L3.

DISCUSSION

The genomes of obligate intracellular organisms have been streamlined via a reductive mode of evolution that is thought to rid the genome of unnecessary coding sequences (1). For example, while free-living bacteria, such as *Escherichia coli*, have multiple copies of the rRNA gene operon (*E. coli* has seven [4]), obligate intracellular and parasitic bacteria sometimes have only 1 to 2 copies (1, 2, 8, 10). Therefore, the discovery of a superfamily of large, protein-encoding genes with unknown function in the *C. trachomatis* genome is of considerable interest.

The *pmp* genes are large, ranging from 2,600 to 5,300 bp. An exhaustive sequencing survey of nucleotide diversity within these large genes for the 15 serovars of *C. trachomatis* would result in about 500,000 bp of nucleotide sequence data. Therefore, we chose instead to first screen for diversity by using RFLP analysis and to then further investigate three of the potentially interesting genes with sequence analysis.

The original goal of our study was to investigate the *pmp* genes for use as molecular epidemiologic markers. We obtained RFLP analysis of eight genes and sequenced three that looked the most interesting by RFLP analysis. Based on our results, however, it is unlikely that there is significant variation among strains of the same serovar. *pmpI* was highly conserved, and a recent report suggests that *pmpG* is also highly conserved (W. J. Jackson, R. B. Taylor, J. H. Tian, K. Johnson, X. Ding, N. Chang, and H. H. Yang, Abstr. 102nd Gen. Meet. Am. Soc. Microbiol., abstr. E-53, 2002). *pmpE* exhibited up to 6.5% nucleotide and amino acid dissimilarity among serovars in different disease groups, but between serovars within a disease group, *pmpE* differed by less than 0.5%. A similar result was obtained for *pmpH*, which showed 8% nucleotide dissimilarity among serovars in different disease groups, but nearly identical sequence for serovars within a disease group. Thus, it is unlikely there will be substantial intraserovar variation in any of the *pmp* genes.

Evolutionary analysis of *pmpE*, *pmpH*, and *pmpI* produced trees that are significantly different from the published MOMP/*omp1* tree (9, 27). In the MOMP evolutionary tree, ocular, urogenital, and LGV serovars are intermixed and form three clades that are not congruent with disease groups: the B serogroup (B, Ba, D, L1, and L2), the C serogroup (A, C, H, I, J, K, and L3), and the intermediate group (serovars F and G). Analysis of *pmpE*, *pmpH*, and *pmpI* indicated that these genes have evolved such that the serovars of *C. trachomatis* diverge into clades that are more consistent with disease groups. However, the evolutionary pattern of each *pmp* gene is somewhat unique.

We were surprised to find that *pmpI* is highly conserved. The paucity of genetic variability resulted in an ambiguous phylogenetic pattern (Fig. 2). In evolutionary terms, highly conserved genes either code for important proteins for which primary amino acid sequence as well as tertiary structure must be conserved or are only recently evolved and have not had time to accumulate mutational variation (18). The high degree of sequence conservation suggests three possibilities for PmpI: (i) PmpI may play a general role in pathogenesis, (ii) PmpI may not be expressed, or (iii) PmpI may not be exposed on the cell surface and thus not under selective pressure by host

TABLE 7. Sequence alignment of *pmpH* in variable region around bp 500

Serovar	Sequence (from bp 476)[a]
A	CGCCTGCTCTAGATCCATCCCCTACCGCTTCAAGCTCTTCATCTCCCACAGTCAGTGATGCTCGCCAGGGATCTATTTTTTCTATAGAGACCAGTTTAGA
B	........C...
Ba	........C...
C	........C...
D	T..A..A.C..CA...G.T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
E	T..A..A.C..CA...G.T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
F	T..A..A.C..CA...G.T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
G	T..A..A.C..CA...G.T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
H	T..A..A.C..CA...G.T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
I	T..A..A.C..C------T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
J	T..A..A.C..C------T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
K	T..A..A.C..C------T...G.T............T.....A.....T..........GA.A..G.................G......T.....G..
L1	.....-------------------------------..A.....T..........GA.A..G.................G......T.....G..
L2	.....-------------------------------..A.....T..........GA.A..G.................G......T.....G..
L3	.....-------------------------------..A.....T..........GA.A..G.................G......T.....G..

[a] *pmpH* sequences for all serovars were compared to that of serovar A on the top line. ., identity with serovar A; −, deletion compared to serovar A.

defenses. Consistent with the latter two suggestions, PmpI has not been demonstrated yet to be in the outer membrane of *C. trachomatis* (28).

PmpE showed a remarkable and quite unexpected evolutionary pattern. The amount of diversity in PmpE in the ocular strains compared to the STD and LGV strains was three times the amount of diversity between STD and LGV strains, and the majority of nucleotide substitutions were nonsynonymous. The amount of genetic change taking place in PmpE within the ocular group suggests that PmpE may be under different selective pressure in the ocular strains versus the STD and LGV strains. Two possible hypotheses regarding the source of this difference are (i) pressure for variability in the ocular strains, for instance, due to selection at the cell surface and (ii) pressure towards conservation in the nonocular strains, for instance, due to constraints on the amino acid structure of the protein. However, until the function of PmpE is known, it is

difficult to speculate on the source of the difference. Furthermore, while one study has demonstrated PmpE to be surface exposed in serovar L2 (28), a different study failed to find this protein on the surface of serovar L2 (21). Therefore, further study of PmpE is necessary.

Although we did not sequence *pmpF* for this study, the RFLP patterns for this gene are very similar to those seen for *pmpE*, suggesting a possible divergence of serovars consistent with disease group. In all three digests of *pmpF*, serovars in different disease groups do not share the same RFLP pattern. In addition, *pmpE* and *pmpF* are separated by only 3 bp, and they are located in the same cluster of *pmp* genes as *pmpG* and *pmpH* (26). Therefore, sequencing *pmpF* may demonstrate that evolution of this gene is comparable to patterns seen in *pmpE* and/or *pmpH*.

PmpH exhibited up to 8% nucleotide and amino acid dissimilarity and several in/del events among serovars when com-
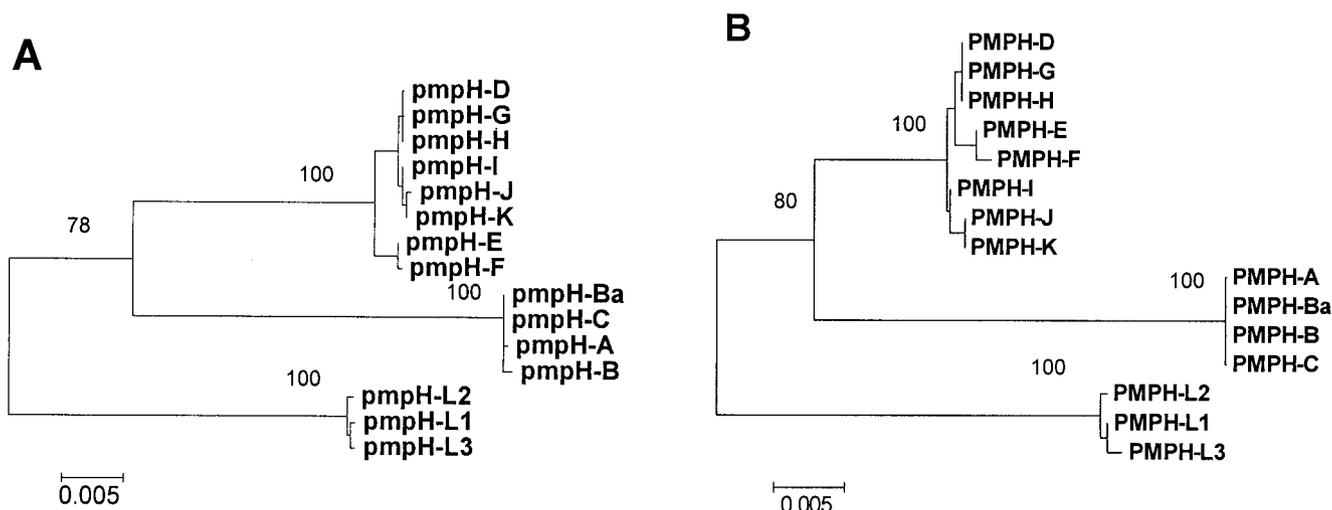

FIG. 6. Evolution of *pmpH* indicates strong divergence of serovars compatible with disease groups. Phylogenetic trees based on the nucleotide (A) and inferred amino acid (B) sequences of *pmpH*. Bootstrap values are shown only at the nodes separating major clades and only when they exceeded 50%. Based on *pmpH*, the 15 serovars clearly evolve into three distinct disease groups, with strong bootstrap support (~100%) for the three clades. The scale bar equals a corrected sequence divergence of 0.005, which is the same as the scale used in Fig. 2 and 4. However, because there is significantly more divergence in *pmpE* and *pmpH* than in *pmpI*, the actual bars here and in Fig. 4 are similar in length compared to the elongated bar used in Fig. 2.

paring sequences among disease groups. However, within a disease group, the serovars were nearly identical in sequence (Table 6). Phylogenetic analysis of *pmpH* nucleotide and inferred amino acid sequences (Fig. 6) indicates that, based on this gene, the serovars of *C. trachomatis* diverge distinctly into three clades that are compatible with disease groups: an ocular-trachoma clade (serovars A, B, Ba, and C), a urogenital-STD clade (serovars D, E, F, G, H, I, J, and K), and an LGV clade (serovars L1, L2, and L3). No other *C. trachomatis* gene studied to date shows a pattern of evolution that parallels disease. The bootstrap support for this pattern was very strong for both the nucleotide and amino acid trees, with values hovering around 100%. In addition, two studies have shown PmpH to be one of the proteins located in the outer membrane of *C. trachomatis* serovar L2 (21, 28). Therefore, based on its location on the cell surface and its evolutionary pattern consistent with disease groups, we hypothesize that PmpH may be essentially involved in the mechanisms of pathogenesis, perhaps as a ligand on the cell surface involved in attachment or invasion.

The three *pmp* genes investigated here each show a different pattern of evolution. The evolution of PmpH is most provocative, however, especially with recent evidence that the *Chlamydia* spp. *pmp* family closely resembles the autotransporter or type V secretion family of proteins (3, 7, 14). Proteins within this family are known toxins, adhesins, and mediators of intracellular motility, including *Bordetella* pertactin, FHA, and TcfA; *E. coli* AIDA-I and TibA; *Haemophilus influenzae* Hap, Hia, and Hsf; *Rickettsia* spp. rOmpA and rOmpB; *Moraxella catarrhalis* UspA1 and UspA2 h; and *Shigella flexneri* IcsA (reviewed in reference 15). Therefore, we hypothesize that *C. trachomatis* PmpH may play an important role in pathogenesis based on its evolutionary pattern, together with its outer membrane location and similarity to type V secretion proteins. It is important to further investigate the function of PmpH and those of the other *pmp* proteins to determine if they play a role in pathogenesis. It is possible that PmpH may be a ligand responsible for attachment or invasion and would thus be an attractive vaccine candidate.

REFERENCES

1. **Andersson, S. G. E., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. Alsmark, R. F. Podowski, A. K. Naslund, A.-S. Eriksson, H. H. Winkler, and C. G. Kurland.** 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396:**133–140.
2. **Andersson, S. G. E., D. R. Stothard, P. A. Fuerst, and C. G. Kurland.** 1999. Molecular phylogeny and rearrangement of rRNA genes in *Rickettsia* species. Mol. Biol. Evol. **16:**987–995.
3. **Birkelund, S., G. Christiansen, B. Vandahl, and A. S. H. Pedersen.** 2002. Are the Pmp proteins parallel beta-helices?, p. 551–554. *In* J. Schachter, G. Christiansen, I. N. Clarke, M. R. Hammerschlag, B. Kaltenboeck, C.-C. Kuo, R. G. Rank, G. L. Ridgway, P. Saikku, W. E. Stamm, R. S. Stephens, J. T. Summersgill, P. Timms, and P. B. Wyrick (ed.), Chlamydial infections: proceedings of the Tenth International Symposium on Human Chlamydial Infections. International Chlamydia Symposium, San Francisco, Calif.
4. **Blattner, F. R., G. Plunkett III, C. A. Bloch, et al.** 1997. The complete genome sequence of *Escherichia coli* K12. Science **277:**1453–1462.
5. **Cabot, E. L., and A. T. Beckenbach.** 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci. **5:**233–234.
6. **Cevenini, R., M. Donati, E. Brocchi, F. DeSimone, and M. LaPlaca.** 1991. Partial characterization of an 89-kDa highly immunoreactive protein from *Chlamydia psittaci* A/22 causing ovine abortion. FEMS Microbiol. Lett. **81:**111–116.
7. **Christiansen, G., A.-S. Pedersen, K. Hjerno, B. Vandahl, and S. Birkelund.** 2000. Potential relevance of *Chlamydia pneumoniae* surface proteins to an effective vaccine. J. Infect. Dis. **181:**S528–S537.
8. **Clark, J.** 1990. Ph.D. thesis. The Ohio State University, Columbus.
9. **Dean, D., and K. Millman.** 1997. Molecular and mutation trends analyses of *omp1* alleles for serovar E of *Chlamydia trachomatis*. Implications for the immunopathogenesis of disease. J. Clin. Investig. **99:**475–483.
10. **Fraser, C. M., J. D. Gocayne, O. White, et al.** 1995. The minimal gene complement of *Mycoplasma genitalium*. Science **270:**397–403.
11. **Giannikopoulou, P., K. Bini, O. D. Simitsek, V. Pallini, and E. Vretou.** 1997. Two-dimensional electrophoretic analysis of the protein family at 90 kDa of abortifacient *Chlamydia psittaci*. Electrophoresis **18:**2104–2108.
12. **Griffiths, P. C., H. L. Phillips, M. Dawson, and M. J. Clarkson.** 1992. Antigenic and morphological differentiation of placental and intestinal isolates of *Chlamydia psittaci* of ovine origin. Vet. Microbiol. **30:**165–177.
13. **Grimwood, J., and R. S. Stephens.** 1999. Computational analysis of the polymorphic protein superfamily of *Chlamydia trachomatis* and *Chlamydia pneumoniae*. Microb. Comp. Genomics **4:**187–201.
14. **Henderson, I. R., F. Navarro-Garcia, and J. P. Nataro.** 1998. The great escape: structure and function of the autotransporter proteins. Trends Microbiol. **6:**370–378.
15. **Henderson, I. R., and J. P. Nataro.** 2001. Virulence functions of autotransporter proteins. Infect. Immun. **69:**1231–1243.
16. **Kalman, S., W. Mitchell, R. Marathe, C. Lammel, J. Fan, W. Hyman, L. Olinger, J. Grimwood, R. W. Davis, and R. S. Stephens.** 1999. Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. Nat. Genet. **21:**385–389.
17. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:**1244–1245.
18. **Li, W.-H., and D. Gauer.** 1991. Fundamentals of molecular evolution. Sinauer Associates, Sunderland, Mass.
19. **Lindquist, E. A., and R. S. Stephens.** 1998. Transcriptional activity of a sequence variable protein family in Chlamydia trachomatis, p. 259–262. *In* R. S. Stephens, G. I. Byrne, G. Christiansen, I. N. Clarke, J. T. Grayston, R. G. Rank, G. L. Ridgway, P. Saikku, J. Schachter, and W. E. Stamm (ed.), Chlamydial infections: proceedings of the Ninth International Symposium on Human Chlamydial Infections. International Chlamydia Symposium, San Francisco, Calif.
20. **Longbottom, D., M. Russell, G. E. Jones, F. A. Lainson, and A. J. Herring.** 1996. Identification of a multigene family coding for the 90kDa proteins of the ovine abortion subtype of *Chlamydia psittaci*. FEMS Microbiol. Lett. **142:**277–281.
21. **Mygind, P. H., G. Christiansen, P. Roepstorff, and S. Birkelund.** 2000. Membrane proteins PmpG and PmpH are major constituents of *Chlamydia trachomatis* L2 outer membrane complex. FEMS Microbiol. Lett. **186:**163–169.
22. **Newhall, W. J., B. Batteiger, and R. B. Jones.** 1982. Analysis of the human serological response to proteins of *Chlamydia trachomatis*. Infect. Immun. **38:**1181–1189.
23. **Rockey, D. D., J. Lenart, and R. S. Stephens.** 2000. Genome sequencing and our understanding of chlamydiae. Infect. Immun. **68:**5473–5479.
24. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
25. **Souriau, A., J. Salinas, C. De Sa, K. Layachi, and A. Rodolakis.** 1994. Identification of subspecies- and serotype 1-specific epitopes on the 80- to 90-kilodalton protein region of Chlamydia psittaci that may be useful for diagnosis of chlamydial induced abortion. Amer. J. Vet. Res. **55:**510–514.
26. **Stephens, R. S., S. Kalman, C. Lammel, J. Fan, R. Marathe, L. Aravind, W. Mitchell, L. Olinger, R. L. Tatusov, Q. Zhao, E. V. Koonin, and R. W. Davis.** 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science **282:**754–759.
27. **Stothard, D. R., G. Boguslawski, and R. B. Jones.** 1998. Phylogenetic analysis of the *Chlamydia trachomatis* major outer membrane protein and examination of potential pathogenic determinants. Infect. Immun. **66:**3618–3625.
28. **Tanzer, R. J., and T. P. Hatch.** 2001. Characterization of outer membrane proteins in *Chlamydia trachomatis* LGV serovar L2. J. Bacteriol. **183:**2686–2690.
29. **Tanzer, R. J., D. Longbottom, and T. P. Hatch.** 2001. Identification of polymorphic membrane proteins of *Chlamydia psittaci* 6BC. Infect. Immun. **69:**2428–2434.
30. **Wang, S. P., C.-C. Kuo, and J. T. Grayston.** 1973. A simplified method for immunological typing of trachoma-inclusion conjunctivitis-lymphogranuloma venereum organisms. Infect. Immun. **8:**356–360.
31. **Wang, S. P., C.-C. Kuo, R. C. Barnes, R. S. Stephens, and J. T. Grayston.** 1985. Immunotyping of *Chlamydia trachomatis* with monoclonal antibodies. J. Infect. Dis. **152:**791–800.
32. **Wang, S. P., and J. T. Grayston.** 1991. Three new serovars of *Chlamydia trachomatis*: Da, Ia, and L2a. J. Infect. Dis. **163:**403–405.