

Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions

Bindu Bhugra,^{1†} LeRoy L. Voelker,² Nianxiang Zou,¹ Hulan Yu² and Kevin Dybvig^{1,2*}

¹Department of Microbiology and ²Department of Comparative Medicine, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA.

Summary

The chromosome of the murine pathogen *Mycoplasma pulmonis* undergoes rearrangements at a high frequency. We show that some of these rearrangements regulate the phase-variable expression of a cluster of genes (the *vsa* locus) that encode the variable V-1 surface antigens. Only one *vsa* gene is associated with an expression site; the other *vsa* genes are transcriptionally silent. The silent genes lack the 5' end region (promoter and ribosome-binding site) that is present in the expressed gene, and DNA rearrangements regulate gene expression by reassorting the 5' end region from an expressed gene with the 3' end region from a previously silent gene. All *vsa* rearrangements identified so far are site-specific DNA inversions that occur between copies of a specific 34 bp sequence that is conserved in each *vsa* gene. Interestingly, DNA inversions within the *vsa* locus apparently occur in concert with inversion of the *hsd1* element, which regulates restriction and modification activity in *M. pulmonis*.

Introduction

Pathogenic mycoplasmas are known for being difficult to eradicate and for causing slowly progressive, chronic diseases of the respiratory tract, genital tract and joints (reviewed by Cassell *et al.*, 1978; Krause and Taylor-Robinson, 1992; Simecka *et al.*, 1992). The chronic nature of these infections indicates that mycoplasmas have the potential to respond and adapt to the rapidly changing conditions encountered in the animal hosts. Although the basis for chronicity is not well understood, it is apparent that several species of mycoplasma undergo high-frequency changes in phenotype associated with

the phase-variable production of key surface antigens (reviewed by Dybvig, 1990; Wise *et al.*, 1992). Such variability may contribute to the ability of these organisms to adapt successfully and evade host defences. One mechanism by which phenotypic diversity is generated in many bacterial systems involves high-frequency DNA rearrangements that reassort coding regions and create populations of bacterial cells that are poised to respond to even the most rapid environmental changes (reviewed by Dybvig, 1993).

Murine respiratory mycoplasmosis (MRM), caused by *Mycoplasma pulmonis*, is one of the most common and important naturally occurring diseases of laboratory rats and mice (Cassell *et al.*, 1985). MRM is easily reproducible and is an excellent model system to study chronic respiratory diseases, particularly in respect to the pathogenic mechanisms of mycoplasmal diseases of humans and animals. Phenotypic switching in *M. pulmonis* may be a factor contributing to the chronicity of MRM. Known examples of high-frequency phenotypic switching in *M. pulmonis* include changes in growth properties, susceptibility to mycoplasma virus P1, generation of restriction-fragment-length polymorphisms (RFLPs), changes in the production of DNA restriction and modification enzymes and variation in the V-1 surface antigen (Bhugra and Dybvig, 1992; Dybvig, 1990; Dybvig and Yu, 1994).

Restriction and modification activity in *M. pulmonis* is regulated by the *hsd1* DNA inversion system (Dybvig and Yu, 1994). The finding of a phase-variable restriction and modification system challenges the notion that the primary function of restriction enzymes is to protect cells from the invasion of foreign DNA (e.g. bacteriophage infection). In the present study, DNA rearrangements that occur in concert with *hsd1* inversion were examined. These rearrangements occur within the locus designated *vsa* (variable surface antigen) and result in the phase-variable expression of the *vsa* genes, which encode the highly repetitive V-1 surface antigens. Accordingly, we propose that *M. pulmonis* has a complex system involving the co-ordinate regulation of phase-variable restriction and modification enzymes and surface proteins.

The experiments described in the present study demonstrate that extensive phenotypic variability is generated via high-frequency DNA rearrangements in the *vsa* locus. A single *vsa* gene is predominately expressed in any given subclone, and subclones that express alternative *vsa* genes have been isolated and characterized. Nucleotide

Received 25 April, 1995; revised 14 July, 1995; accepted 24 July, 1995. †Present address: Department of Molecular Genetics, Microbiology and Biochemistry, 3110 MSB, University of Cincinnati, Cincinnati, Ohio 45267, USA. *For correspondence. E-mail dybvig@orion.cmc.uab.edu; Tel. (205) 934 9327; Fax: (205) 975 4418.

sequence data indicate that the silent *vsa* genes become expressed via a DNA rearrangement that disassociates the 5' end region from the expressed gene and combines it with the 3' end region of the previously silent gene. The DNA rearrangements regulating *vsa* gene expression that have been identified so far have all been site-specific DNA inversions that occur within a specific 34 bp sequence designated the *vrs* (*vsa* recombination sequence) box.

Results

Identification of a DNA inversion system in the *M. pulmonis* *vsa* locus

M. pulmonis strain UAB 6510 is a respiratory pathogen of rats (Davis *et al.*, 1982) that served as the progenitor of the other strains used in this study. A description of the lineage and pertinent properties of these strains is provided in the *Experimental procedures*. In UAB 6510 and in progeny strains KD735 and KD735-15, the *hsd1* locus is in orientation I. In several subclones of UAB 6510 including strains KD7-11, KD7-12, KD735-16 and 6a14, *hsd1* has inverted to orientation II (Dybvig and Yu, 1994). In addition to the *hsd1* inversion, these strains have undergone a second rearrangement, in the *vsa* locus, that combines precursor *HindIII* fragments of 4.7 and 5.8 kb into product fragments of 8.3 and 2.2 kb (Bhugra and Dybvig, 1992; Dybvig and Yu, 1994). To date, rearrangement of the 4.7 and 5.8 kb fragments in the absence of *hsd1* inversion has not been observed.

To characterize rearrangements within the *vsa* locus, the 4.7 kb fragment from KD735-15 and the 8.3 kb fragment from KD735-16 were cloned and their complete nucleotide sequences were determined. These fragments were found to have a 3.6 kb region in common, including the open reading frames (ORFs) referred to in Fig. 1 as *vsaB*-S and *vsaC2*-S. (-S denotes a silent copy.) The relative location and orientation of the 4.7 and 5.8 kb precursor fragments in the KD735-15 chromosome was examined by designing a series of primers that would amplify by the polymerase chain reaction (PCR) the 4.7/5.8 kb junction region. One set of oligonucleotides, o.4294 and o.4218 (see Fig. 1A, top), amplified a 1.6 kb fragment, the nucleotide sequence of which demonstrated that the 4.7 and 5.8 kb fragments were adjacent in the chromosome. Additionally, comparison of the nucleotide sequences of the 4.7 and 5.8 kb precursor fragments with that of the 8.3 kb product fragment revealed that *vsaB*-S was in an inverted orientation in KD735-16 relative to its orientation in KD735-15, suggesting that the 8.3 kb fragment arose by DNA inversion. The majority of the 2.2 kb product fragment in KD735-16, containing *lipA* and *vsaD*-S, was isolated by PCR amplification (using primers o.3535 and o.4294, see Fig. 1A, middle). Its nucleotide sequence revealed that the *vsa* rearrangement in KD735-16 was the 5 kb

DNA inversion depicted in Fig. 1A. This inversion occurred via recombination between copies of a perfectly conserved 49 bp sequence, designated *vrsB* and *vrsD*, that is located at the 5' ends of *vsaB*-S and *vsaD*-S. This sequence contains the *vrs* box that is also conserved among several other ORFs (*vsaA*, *vsaC1*-S, *vsaC2*-S and *vsaE2*-S) within the *vsa* locus (Fig. 1B). One of the genes, *vsaC2*-S, possesses two *vrs* boxes: *vrsC2a* and *vrsC2b*.

DNA inversions regulate *vsa* gene expression

Northern hybridization experiments demonstrated that only one of the *vsa* genes is predominately transcribed in each subclone. In KD735-15 and KD735-16, *vsaA* is transcribed (Fig. 2). These results were anticipated because the nucleotide sequence revealed that the *vsaA* gene from these strains had a clearly defined 5' end region consisting of a putative ribosome-binding site (Shine-Dalgarno sequence) located upstream of an ATG initiation codon and encoding a predicted gene product containing a consensus signal peptide sequence. Each of these features was absent in other *vsa* genes (designated with the suffix -S). Southern hybridization analysis demonstrated that the 5' end region of the expressed *vsaA* gene is present in only a single copy in the chromosome (data not shown). Additional Northern hybridization experiments revealed that some strains predominately expressed *vsaB* instead of *vsaA* (Fig. 2). The strains expressing *vsaB* (KD7-11, KD7-12 and 6a14) have previously been shown to have undergone a DNA rearrangement similar to that of KD735-16 (Bhugra and Dybvig, 1992). Namely, in each of these strains, the 4.7 and 5.8 kb precursor fragments had recombined to generate a 8.3 kb fragment. In a derivative of 6a14, designated strain 6a14/4, *vsaA* but not *vsaB* was once again predominately transcribed. As in the other examples of *vsa* phase variation that are described above, the change in *vsa* expression in 6a14 versus 6a14/4 was accompanied by an inversion of *hsd1*. To examine the basis for the phase-variable expression of *vsaA* and *vsaB*, strains 6a14 and 6a14/4 were chosen for further study.

A DNA fragment containing the 5'-termini and upstream flanking regions of *vsaA* and *vsaB* was isolated by PCR amplification (using primers o.4083 and o.6859; see Fig. 1A). Amplification of DNA from strains 6a14, 6a14/4, KD735-16 and KD7-11 resulted in PCR products of identical size: 1.3 kb. The nucleotide sequences of the 6a14 and 6a14/4 PCR products were determined, revealing that the orientation of a region encompassing the 5' end region of *vsaA* was inverted in 6a14 compared with 6a14/4 and KD735-16. Therefore, a single 5' end region dictates expression of *vsaA* and *vsaB*. This 5' end region (expression site) is associated with *vsaA* in KD735-15, KD735-16 and 6a14/4 but is incorporated into *vsaB* in 6a14, converting *vsaB*

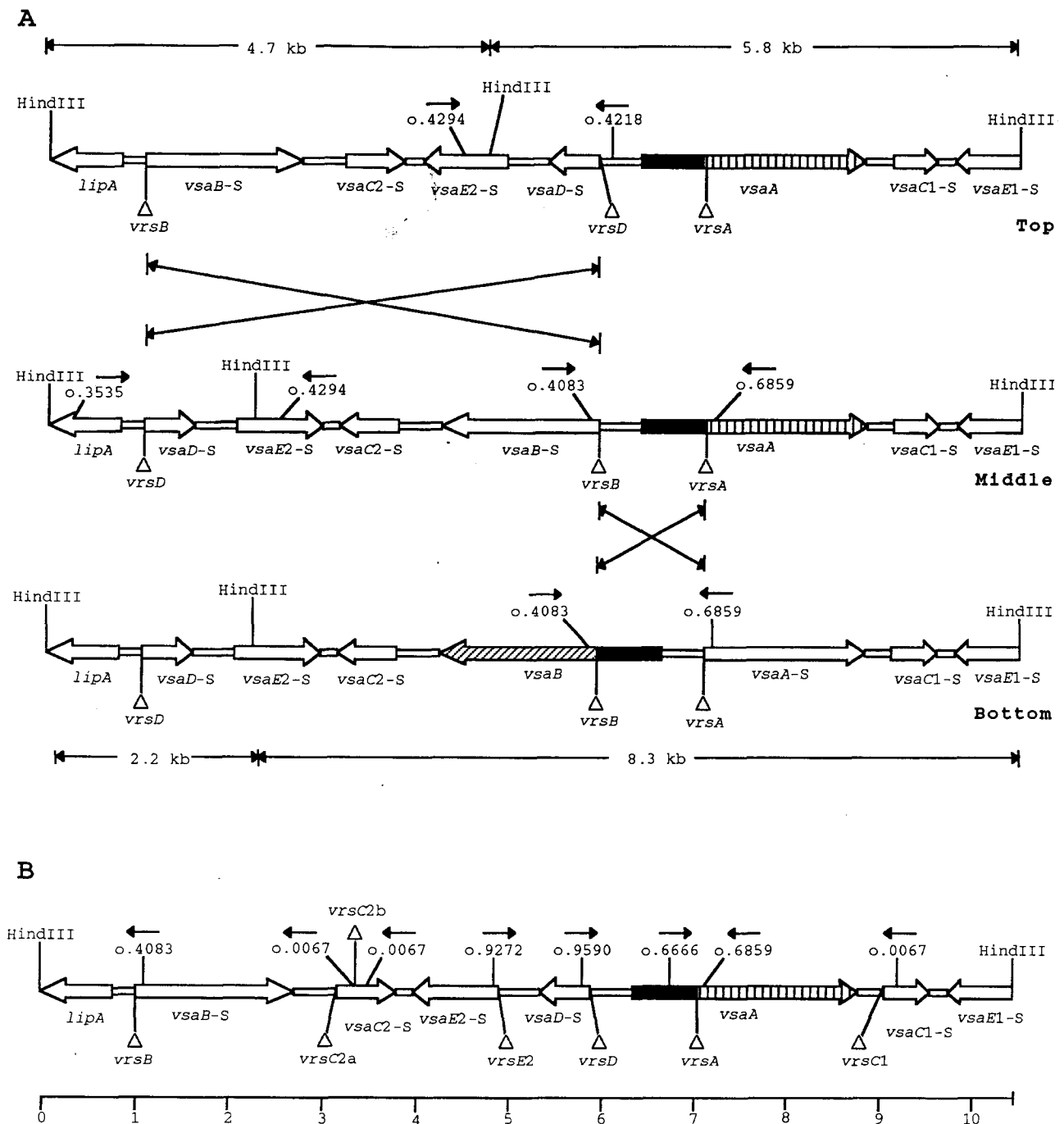


Fig. 1. Organization of the *M. pulmonis* *vsa* locus and delineation of DNA inversion sites. Major ORFs are shown with the arrows indicating the direction of transcription. Silent copies of *vsa* genes are designated with the suffix -S. Triangles indicate locations of *vrs* boxes. The region shaded in black contains the first 712 nucleotides of the *vsa* gene that is expressed, and the striped regions denote the repetitive 3' part of the expressed gene's coding region. The location of the target sites for selected oligonucleotides are shown, with the arrowhead indicating the orientation of primer binding. The *vsaE1* and *lipA* genes are partials; the 5' end of *vsaE1* and the 3' end of *lipA* have not been isolated.

A. Schematic diagram illustrating DNA inversions in the *vsa* locus. Depicted is the organization of the *vsa* locus in strains KD735 and KD735-15 (top), strains KD735-16 and 6a14/4 (middle) and strains 6a14, KD7-11 and KD7-12 (bottom). For comparison, the *Hind*III fragments shown here as being 8.3, 4.7 and 5.8 kb were described previously as being 9.5, 4.9 and 5.7 kb, respectively, by Bhugra and Dybvig (1992).

B. Map of the *vsa* locus from KD735-15 showing the locations of all *vrs* boxes. The target sites for primers that were used for PCR analysis to determine which *vrs* boxes are recombinogenic (see text) are shown. The line at the bottom indicates nucleotide positions, in kilobases. The nucleotide sequence of the KD735-15 *vsa* locus was deposited in the GenBank database with the accession number U23947.

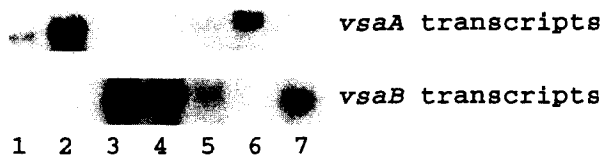


Fig. 2. Northern hybridization analysis of *M. pulmonis* RNA. Northern blots of *M. pulmonis* RNA were probed with oligonucleotides specific for the repeat regions of *vsaA* (top) or *vsaB* (bottom). The *vsaA*-specific probe was 5'-GTAGTTGGAGGAGTCATTG-3', and the *vsaB*-specific probe was 5'-CCATCTCTCCTGCATTTG-CATC-3'. Analysed in each lane is RNA from the indicated strain: lane 1, KD735-15; lane 2, KD735-16; lane 3, KD7-11; lane 4, KD7-12; lane 5, 6a14; lane 6, 6a14/4; and lane 7, 6a14/6. Transcripts from both *vsaA* and *vsaB* were approx. 2.4 kb. The intensity of the bands in lanes 1 and 5 is weak because less RNA was loaded in these lanes, as evidenced by the intensity of the ethidium bromide staining pattern.

from a silent to an expressed gene (i.e. *vsaB*-S converted to *vsaB*). Inversion of sequences that regulate the expression of *vsaA* and *vsaB* occurred between the *vrsA* and *vrsB* boxes (see Fig. 1A).

Characterization of *vrs* box-mediated DNA inversion

An alignment of the nucleotide sequences of the *vrs* boxes is provided in Fig. 3. The 34 bp sequence comprising the smallest of the *vrs* boxes, *vrsE2*, is regarded as the minimal requirement for a *vrs* box. As shown in Fig. 3, sequences immediately upstream of the various *vrs* boxes are not conserved. Single nucleotide differences in the sequences of the *vrs* boxes provide information on the location of strand exchange during DNA inversion. The 18th nucleotide of *vrsA* is G instead of the usual A. Because this G nucleotide was retained by *vrsA* regardless of whether the *vsaA* gene is expressed (as in KD735-15) or silent (as in 6a14), strand exchange leading to DNA inversion must occur within the first 18 nucleotides of the *vrs* box. There is a T instead of the usual C at position 15 of the *vrsC2a* sequence. Because PCR analysis (described below) failed to reveal a product consistent with recombination at *vrsC2a*, it is possible that the C at position 15 is critical for DNA inversion.

Each of the *vrs* boxes that has been identified so far is

recombinogenic, with the possible exception of *vrsC2a*. PCR analysis was performed on 35 different subclones (including all subclones described in the *Experimental procedures*) containing a variety of RFLPs and phenotypic differences. Primers were designed such that they would PCR amplify the *vrs* box region of the particular *vsa* gene that is in the expression locus (see Fig. 1B). From each subclone, PCR products were obtained that corresponded to amplification across *vrsA* (using primer pair o.6666 and o.6859), *vrsB* (using o.6666 and o.4083), *vrsC1* and *vrsC2* (using o.6666 and o.0067), *vrsD* (using o.6666 and o.9590), and *vrsE2* (using o.6666 and o.9272). These data indicated that each *vsa* gene is associated with the expression site in a subpopulation of cells that is present in each subclone. In the case of the *vsaC1* and *vsaC2* genes, a single PCR product was obtained despite the fact that primer o.0067 binds to *vsaC2* at two sites and *vsaC1* at one site. The nucleotide sequence of the PCR product was consistent with amplification of an expressed copy of *vsaC1* and/or an expressed copy of *vsaC2* that arose as a result of recombination at *vrsC2b*. The nucleotide sequence downstream of *vrsC1* is identical to the sequence that is downstream of *vrsC2b*. Therefore, PCR products corresponding to these regions were indistinguishable at the nucleotide level. A PCR product consistent with amplification of an expressed copy of *vsaC2* that arose as a result of recombination at *vrsC2a* was not obtained. Perhaps, *vrsC2a* is not recombinogenic, or it may undergo DNA inversion at a lower frequency than do the other *vrs* boxes.

Other mechanisms of variation in the *vsa* locus: size variation resulting from gain or loss of tandem repeats and gene duplication

The *vsa* genes, except for *vsaE1* and *vsaE2*, contain a highly repetitive domain located 3' to the *vrs* boxes. Each repetitive domain consists of tandem repeats with the length of the repeat unit always being a multiple of three. The repeat units of *vsaA*, *vsaB*, *vsaC1*, *vsaC2* and *vsaD* contain 51, 39, 36, 36 and 33 nucleotides, respectively.

<i>vrs</i> Box	Strain	Nucleotide Sequence
<i>vrsA</i>	(KD735-15)	AAAAACATCAAATAATGAACAAGGTGGAAATAATTCGCTTCAAATGGTTCAA
<i>vrsA</i>	(6a14)
<i>vrsB</i>	(KD735-15)	..TT..C.....A.....A.....A..A..GA...
<i>vrsB</i>	(KD735-16)	..T.....A.....A.....A..A..GA...
<i>vrsB</i>	(6a14)A.....A.....A..A..GA...
<i>vrsC1</i>	(KD735-16)	..TT..T.....A.....AA.....A..CA...
<i>vrsC2a</i>	(KD735-16)	G...T.....T..A.....A.....A..GAG...
<i>vrsC2b</i>	(KD735-16)	TCCTC.....A.....AA.....A..CA...
<i>vrsD</i>	(KD735-15)	..T.....A.....A.....A..GA...
<i>vrsD</i>	(KD735-16)	..TT..G.....A.....A.....A..GA...
<i>vrsE2</i>	(KD735-15)	G...T.....A.....A.....AAA..CC..AAGGTC..
<i>VsaA</i> product		...LysThrSerAsnAsnGluGlnGlyGlyAsnAsnSerGlySerAsnGly...
<i>VsaB</i> product		...LysThrSerAsnAsnGluGlnSerGlyAsnAsnSerGlySerLysAsp...

Fig. 3. Comparison of the nucleotide sequences of the *vrs* boxes and the encoded amino acid sequences. The large box is drawn around the 34 nucleotides that are considered to be the minimum *vrs* box sequence, but sequence identity sometimes continues 3' of this minimum. Sequences immediately upstream of each *vrs* box are also shown. Dots in the sequence indicate identity with the *vrsA* sequence shown on the top line. References in the text to *vrs* box nucleotide positions consider the first C in the *vrsA* sequence as nucleotide position number 1.

Constant region

MKNSNMKKKFLFFLGGSTSFSLSIAGATLIAC*QKQDDGAIKKDLESKIDVLTFFPKSVDETAQWTVSRLLSEYRTAKSS
 DSNLTFDAFVKSGLGSKLGSYELTFKEKGADKIEYTYKKSNAKAVSSTKVLEDTKIGQSDAQKVNAWLDRVKSLVAN
 SSDLKGKPKPSEVIEIETSDVKFKETRDGQDVEFDVPTNWGTVITKSSANDAANDAEGTLKVSQVQSSGDFRFRSKDLNL
 TGLKTSNNEO--Variable region

Vrs box start

Variable region

A--GGNNSGSNSGSKPM [TPPTTGGSGSGSTDTK (P/A) Q]₃₂TPPTTGGSGGSIN*
 VrsA/Repeat I

B--SGNNSGSKDENPMPNKP IGEKPMTDKSNGSDGDK [GKNAGDANAG (E/D) M]₃₉DEGNTGDTETNYI*
 VrsB Repeat II Repeat I

C1--SGNNSSKDONSTPPKTDOGSNGSSGNEMKNPLASSGDKNV [TP (P/S) TTGD (A/E) NTSQ]₁₀TPSTTGV*
 VrsC1 Repeat III/Repeat I Repeat IV

C2--SGNNSSKDONSTPPKTDOGSNGSSGNEMKNPLASSGDKNV [TP (P/S) TTGD (A/E) NTSQ]₇TPSTTGV*
 VrsC2b Repeat III/Repeat I Repeat IV

D--SGNNSGSKDENSTPPKTDOGSSSQTKPQKPMTEKPMTDOGSNGSGNEMK [NPPAGGTMKEE]₁₁NPPAV*
 VrsD Repeat III Repeat II/Repeat I Repeat IV

E2--SGNNSGKTEGOAEGSVSNSGKTEGOTEGNGSDTGQTNQVDLQMLLDSIESVEISSKFETKSEIVLNSSTGIK
 VrsE2/Repeat E Repeat E

LKKADNSDFSI PNGVKLLLSFQKEKNEKNNKFSLGISLMLSSSNSEGVSKDLIISKFESNQDDNA [LKSISR
 LEETTKKIDLVSQELDLKDQEEIVLVKQSELI SNNLKSLLFKLLENFKNAAKDVKLTKEQETLLNKESISLEE
 FDSLKGLTQKKSQVFTKIELTKEQEDKFKNVIQDAKVKLEDLTKKLEKIKSEKENIEKKLDP I IKELKERL
 KINE] LDSTKSFVVIYK*

Fig. 4. Deduced amino acid sequences of the Vsa proteins. The first 242 amino acids of the Vsa protein defines the N-terminal constant domain that is common to all Vsa proteins and terminates within residues encoded by the vrs box. The Vsa signal peptide sequence is underlined, with the Cys residue to which fatty acid is predicted to be attached indicated by C*. Brackets denote the tandem repeat domains (in the case of VsaE2, brackets denote the large α -helical domain) of the variable C-terminal end of each Vsa protein with the subscript denoting the number of repeat units. Amino acids in parenthesis refer to heterogeneity in the repeat region. For example, the 16th amino acid in the VsaA repeat region is sometimes proline (P) and sometimes alanine (A), depending on the particular repeat unit. Charged amino acids are indicated by a + or - sign above the amino acid. Variable-region sequences that tend to be conserved (Vrs box amino acids, Repeats I-IV and Repeat E) are also underlined.

Therefore, the predicted vsa gene products contain tandem repeats of 17, 13, 12, 12 and 11 amino acids, respectively (Fig. 4). The number of repeat units per molecule ranges from 7 to 39. Because the size of the repeat region in vsaC and vsaD is small (< 150 nucleotides), the complete nucleotide sequence of each repeat unit could be determined. In contrast, the repeat regions of vsaA and vsaB comprise about 1.6 kb. The actual nucleotide sequence of the vsaA and vsaB repeats is only available for a few copies of the repeat units that are present at either end of the repeat region. The lengths of the vsaA and vsaB repeat regions were determined by restriction enzyme mapping of the cloned genes and by sizing PCR products obtained by amplification of the entire repeat regions using genomic templates (see the *Experimental procedures*), as well as by estimating the size of the mRNA transcripts (Fig. 2). All approaches yielded the same conclusion: there are about 32 copies of the tandem repeat in vsaA and about 39 copies of the tandem repeat in vsaB.

Genetic variation in the vsa locus is not limited to DNA inversion and can involve gain or loss in the number of repeat units per gene. One of the vsa HindIII fragments in strain KD735-15-53 is about 5.6 kb, instead of the parental 5.8 kb fragment of KD735-15 (Bhugra and Dybvig, 1992). PCR amplification was used to examine various regions of the 5.6 kb fragment to identify the specific area that differed from the 5.8 kb fragment. Primers that amplified across the entire vsaD repeat region (see the *Experimental procedures*) resulted in a PCR product that was 180 bp shorter than the product obtained using a KD735 DNA template. The nucleotide sequence of the PCR product revealed that vsaD from KD735-15-53 had only six copies of the tandem repeat unit compared with 11 copies of the repeat unit in most of the other subclones (e.g. KD735-15). Therefore, changes in the number of repeat units introduce heterogeneity in the size of some vsa genes.

Another type of variation that may occur in the vsa locus is gene duplication. The region of the locus in KD735-15

that encompasses nucleotides 3372–4686 (containing *vsaC2-S* and *vsaE2-S*) is almost 100% identical with the region encompassing nucleotides 9094–10515 (containing *vsaC1-S* and *vsaE1-S*). Except for a difference in the number of tandem repeats within the two *vsaC* genes, these two regions differ at only two nucleotides and almost certainly arose via gene duplication.

Structural features of the Vsa proteins

The amino acid sequences of the predicted *vsa* gene products are shown in Fig. 4. Because there is a single expression locus for the various *vsa* genes, all Vsa proteins have the same N-terminal end. This constant region encompasses 242 amino acids and begins with a consensus signal peptide sequence that is characteristic of prokaryotic lipoproteins (Hayashi and Wu, 1990). The variable regions of the Vsa proteins begin with amino acid 243 and display striking differences. Most of the Vsa proteins contain a tandem repetitive domain. The tandem repeat units of the majority of these proteins are proline-rich, but the VsaB repeat lacks proline. The VsaE proteins lack tandem repeats, and computer analysis of their predicted amino acids indicate the presence of a large α -helical region (bracketed region in Fig. 4). The predicted sizes and pI values of the various Vsa proteins are as follows: VsaA (80 kDa, pI=8.5); VsaB (80 kDa, pI=4.3); VsaC1 (43 kDa, pI=4.5); VsaC2 (40 kDa, pI=4.8); VsaD (45 kDa, pI=5.0); and VsaE2 (61 kDa, pI=5.8). When compared with the GenBank/EMBL databases, none of the *vsa* gene products had significant similarity to known peptides.

Preceding the tandem repeat regions, various blocks of amino acids within the variable regions are conserved among several of the Vsa proteins. In particular, strong conservation of sequences is apparent in the variable regions of VsaB, VsaC1 and VsaC2 and VsaD. These conserved regions are designated in Fig. 4 as Repeats I–IV. The nucleotides encoding these amino acids are also highly conserved, suggesting that the variable regions of the *vsa* genes may have evolved from a single ancestral sequence. In support of this conclusion, it is noted that the nucleotides of the tandem repeat regions of *vsaA* and *vsaC* are also strikingly similar, resulting in similar repeat units in the gene products (beginning with TPPTTG and ending in Q). Although VsaE2 contains two copies of a repeat sequence (Fig. 4, Repeat E), this repeat is not related to sequences in the other Vsa proteins.

The *vsa* genes encode the *M. pulmonis* V-1 antigens

Synthetic peptides corresponding to the repeat regions of the predicted *vsaA* and *vsaB* gene products were commercially prepared and used to generate rabbit polyclonal antibody that was specific for the Vsa proteins (see the

Experimental procedures). The constant region of the *vsa* expression locus (encoded by the region shaded in black in Fig. 1) contains three TGA codons, which encode tryptophan in mycoplasmas but are stop codons in other eubacteria. Therefore, *Escherichia coli* cells containing the cloned *vsa* genes were not examined for the production of Vsa proteins. Rather, the antisera were used to identify the Vsa proteins produced in *M. pulmonis*.

Western blot analysis of *M. pulmonis* proteins separated by SDS-PAGE and reacted with VsaA- and VsaB-specific polyclonal antibody revealed a ladder pattern (Fig. 5). As anticipated from the nucleotide sequence and Northern blot data, proteins from strains KD735-15, KD735-16 and 6a14/4 reacted more strongly with the VsaA-specific antibody (Fig. 5C), whereas strain 6a14 proteins reacted more strongly with the VsaB-specific antibody (Fig. 5B). However, each strain reacted to some extent with both VsaA-specific and VsaB-specific antibody, indicating the strains are heterogeneous and contain subpopulations of cells that produce different Vsa proteins. The estimated size of the peptides in the predominant, top band of the Vsa ladder pattern (150 to 220 kDa) is considerably larger than the predicted size of the VsaA and VsaB proteins based on nucleotide sequence analysis (80 kDa), suggesting either that the Vsa proteins retain secondary structure in the presence of SDS or that they undergo significant post-translational modification.

The Vsa proteins and the previously described V-1 antigens (Watson *et al.*, 1988; 1989) are probably one and the same. Figure 5A shows the ladder pattern obtained by reacting Western blots of *M. pulmonis* proteins with a monoclonal antibody specific for V-1. Reaction of strains KD735-15, KD735-16 and 6a14/4 with the monoclonal antibody yielded a VsaA-like ladder pattern, and reaction of strain 6a14 yielded a VsaB-like ladder pattern. The monoclonal antibody evidently reacts with a conserved epitope present in both the VsaA and the VsaB proteins. This epitope may be located within the constant N-terminal domain of the Vsa proteins. As was described above for reactions using Vsa-specific polyclonal antibody, reactions using the V-1-specific antibody indicated that multiple forms of V-1 (both the VsaA and the VsaB proteins) are produced in each strain, presumably because of the presence of heterogeneous cell populations.

The V-1-specific antibody reacted more strongly with proteins of apparently lower molecular weight than did the Vsa-specific polyclonal antibody (see the bottom halves of the blots shown in Fig. 5). The epitope recognized by the monoclonal antibody is evidently present on all rungs of the V-1 ladder pattern. Presumably, the polyclonal antibody, which is specific for the repeat regions of the Vsa proteins, reacts weakly with the bottom of the Vsa protein ladder because few copies of the repeat unit are present in these molecules.

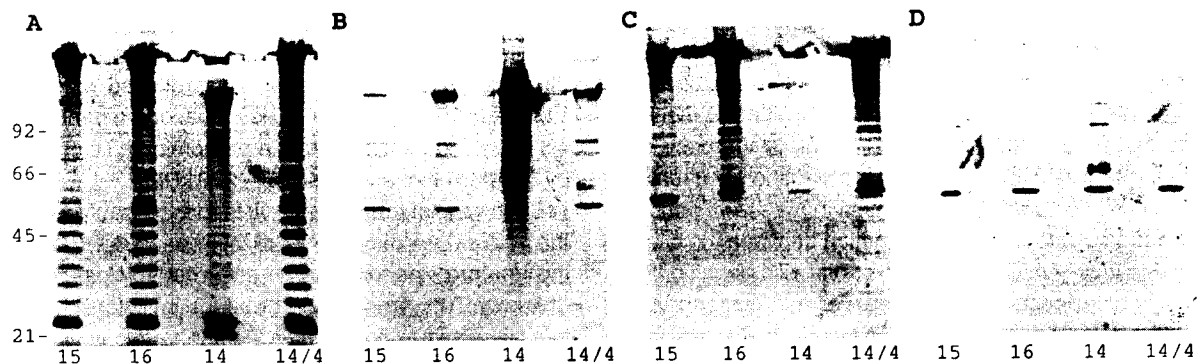


Fig. 5. Western blot analysis of the *M. pulmonis* proteins. Immunoblots of total protein from *M. pulmonis* strains were reacted with the V-1-specific monoclonal antibody 7.1-2 (A), rabbit antisera specific for the VsaB repeat unit (B), rabbit antisera to the VsaA repeat unit (C), or to biotinylated, anti-rabbit IgG conjugate as a control with no primary antibody in the protocol (D). The protein samples contained in lanes from left to right are abbreviated as: 15, KD735-15; 16, KD735-16; 14, 6a14; and 14/4, 6a14/4. Compared with molecular size standards, the V-1 ladder pattern ranges in size from 20 to about 220 kDa (bottom to top).

The lipA gene

Adjacent to the *vsa* locus is an ORF designated *lipA* (encoding a putative lipoprotein) that is preceded by a Shine-Dalgarno sequence upstream of an ATG start codon (see Fig. 1). As was the case with the Vsa proteins, the predicted amino acid sequence of the *lipA* gene product begins with a consensus signal peptide sequence that is characteristic of prokaryotic lipoproteins. However, the specific amino acid sequences of the signal peptides of LipA and the Vsa are not closely related. When compared with the GenBank/EMBL databases, the LipA amino acid sequence was found to share some similarity (25% amino acid identity) to the ClfA protein of *Staphylococcus aureus* (McDevitt *et al.*, 1994). ClfA is thought to be a receptor for binding of fibrinogen, but the degree of amino acid similarity between LipA and ClfA is insufficient to warrant speculation regarding LipA function.

Discussion

DNA inversion and *vsa* gene expression

The portion of the *vsa* locus that has been characterized so far has 10515 nucleotides, but the entire locus has not yet been cloned. The 5' end of *vsaE1* has not been isolated, and it is possible that future studies will identify additional *vsa* genes. The characterized portion of the *vsa* locus contains a single *vsa* expression site, consisting of a full-length *vsa* gene. The expressed gene contains a copy of the 34 bp *vrs* box beginning at nucleotide 713 of the *vsa* coding region. Also present in the *vsa* locus are five (six if *vsaE1* is included) silent *vsa* genes. Each of the silent *vsa* genes lacks the first 712 nucleotides of the *vsa* coding region and contains a *vrs* box at the beginning

of the gene's ORF. The phase-variable expression of the *vsa* genes is regulated by site-specific DNA inversions that occur between copies of the *vrs* boxes and join a previously silent gene with the expression locus. Each of the *vsa* genes has been found by PCR analysis to be sometimes in association with the expression locus, indicating that the full spectrum of *vsa* genes is expressed in the heterogeneous population of cells that are present in a culture of *M. pulmonis*.

DNA inversions within the *vsa* locus occur at high frequencies, as evidenced by the ease at which strains containing inversions were isolated. For example, KD735-16 was 1 of 34 colonies examined from KD735, 6a14/4 was 1 of 105 colonies examined from KD735-15, and 6a14/4 was one of only six colonies examined from 6a14 (Dybvig and Yu, 1994). Western blot analysis supports the conclusion that cultures of *M. pulmonis* are highly heterogeneous; both VsaA and VsaB proteins were detected in all strains that were analysed. Also, some heterogeneity in mRNA transcripts was detected by Northern blot analysis, as evidenced by the presence of a low level of *vsaA* transcripts in RNA isolated from 6a14 cells (see Fig. 2).

Some DNA inversions within the *vsa* locus, such as the one that generated strain KD735-16, occur between silent genes and do not result in a change in *vsa* gene expression. However, DNA inversions that are phenotypically silent do contribute to the overall variability of the system because it is not possible to express each of the *vsa* genes by a single DNA inversion. For example, a single DNA inversion in KD735 cells can result in expression of *vsaE2* or *vsaD*, but it cannot result in expression of either *vsaB*, *vsaC2* or *vsaC1* because these genes are in a direct orientation relative to the expressed gene. The phenotypically silent DNA inversion that generated KD735-16 cells is important in that it poises the system to express

vsaB or *vsaC2* by undergoing another DNA inversion, such as the one that gave rise to strain 6a14.

Many pathogens vary their surface structures by undergoing gene rearrangements, but strikingly different mechanisms can be used to achieve this goal. In the case of large gene families, the mechanism of variation is usually not DNA inversion. Pilin gene variation in *Neisseria gonorrhoeae* reflects extensive recombination between the pilus expression locus (*pilE*) and silent, partial pilin gene (*pilS*) sequences. Non-reciprocal gene conversion (Segal *et al.*, 1986), transformation-mediated recombination (Facijs and Meyer, 1993) and repair-driven recombination (Hill *et al.*, 1990) are some of the mechanisms proposed for gonococcal pilus variation. Through recombination between linear plasmids, a formerly silent variable major protein (*vmp*) gene in *Borrelia hermsii* replaces another *vmp* gene at the telomeric expression locus downstream of a promoter causing an antigenic switch (Barbour, 1990; Kitten and Barbour, 1990). Although there are numerous examples of site-specific DNA inversion systems that regulate the phase-variable production of surface organelles in bacteria (Dybvig, 1993; Plasterk and van de Putte, 1984), most of these systems are relatively simple and regulate the expression of one or two genes and not gene families as large as the *vsa* locus. When compared with other DNA inversion systems, the complexity of the *vsa* system resembles most closely that of the shufflon system found in enteric bacterial plasmids within the IncI incompatibility group (Komano *et al.*, 1987).

Characteristics of the predicted *Vsa* proteins

V-1 is a major *M. pulmonis* surface antigen that consists of a family of proteins that have repetitive subunit structures (Dybvig, 1990; Watson *et al.*, 1989). The banding patterns of the *Vsa* proteins on Western blots reveals a ladder pattern that is remarkably similar to that of the V-1 antigens, suggesting that the V-1 antigens are encoded by the *vsa* genes. In an independent effort to clone the gene(s) encoding V-1, a gene essentially identical to *vsaA* was cloned from *M. pulmonis* strain CT (Davis *et al.*, 1985) by using an oligonucleotide probe that encoded a partial amino acid sequence derived from purified V-1 (W. L. Simmons, C. Zuhua, J. I. Glass, J. Simecka, H. L. Watson, and G. H. Cassell, manuscript in preparation).

The N-terminal end of the predicted *Vsa* proteins contains a consensus signal peptide sequence that is characteristic of prokaryotic lipoproteins (Hayashi and Wu, 1990; Wu, 1987). The signal peptide sequence in such proteins terminates with a C residue that serves as the acylation site, which should be the site of membrane anchorage of a mature, processed lipoprotein. Therefore, it is likely that the *Vsa* proteins are acylated, as are many other membrane proteins in mycoplasmas (reviewed by Wieslander

et al., 1992). The presence of fatty acid and possibly other post-translational modifications (or an unusual secondary structure) may explain why the predicted sizes of the *VsaA* and *VsaB* proteins (80 kDa) based on genetic analysis is an underestimate of the predominant sizes obtained by Western blot analysis (150–220 kDa).

Each rung of the *Vsa* protein ladder probably represents the gain or loss of one copy of the repeat unit. Exactly why the *Vsa* proteins form ladder patterns on Western blots is unknown. The presence of protease inhibitors during protein isolation has no effect on the appearance of the ladder. It is possible that post-translational modifications generate an array of *Vsa* proteins of varied size that are present on the mycoplasma cell surface. Alternatively, individual cells may produce homogeneous molecules of *Vsa* protein, but the total population of cells present in a culture may contain subpopulations with varied numbers of repeat units in the expressed *vsa* gene. For example, in strain KD735-15-53, the *vsaD* gene was shown to have undergone a deletion that reduced the number of tandem repeats. Presumably, gain or loss of repeat units is a common occurrence in this locus and increases the cell-surface repertoire of the *Vsa* proteins. Numerous examples of surface antigens that contain tandem repetitive domains and undergo size variation through changes in repeat number have been described (reviewed by Dybvig, 1993; Fischetti *et al.*, 1991; Kemp *et al.*, 1987; Wise *et al.*, 1992; Wren, 1991). Slipped-strand mispairing that results from DNA replication errors is usually referred to as the mechanism for this size variation (Levinson and Gutman, 1987). In *M. pulmonis*, the combination of phase and size variation among the *Vsa* proteins has the potential to create an enormous degree of surface antigenic diversity that may be a vital aspect of disease pathogenesis.

Surface antigens having tandem repetitive domains seem to be particularly prevalent in mycoplasmas and their phylogenetic relatives such as streptococci and clostridia (Fischetti *et al.*, 1991; Wren, 1991). Tandem repetitive domains of significant size are present in the Vlp protein family of *Mycoplasma hyorhinis* (Yogev *et al.*, 1991) and in proteins from the human pathogens *Mycoplasma pneumoniae* (Dallo *et al.*, 1990), *Mycoplasma hominis* (Ladefoged *et al.*, 1995) and *Ureaplasma urealyticum* (Zheng *et al.*, 1995). Proteins with tandem repetitive domains are frequently involved with cell–cell interactions, and in many cases, the repeating units are thought to be ligand-binding domains. Each *Vsa* protein presumably has a specific function. Variation in V-1 has been correlated with lung lesion severity using the MRM model (Talkington *et al.*, 1989), colony size (Dybvig *et al.*, 1989), susceptibility to the P1 virus (Dybvig *et al.*, 1988) and haemadsorption (Watson *et al.*, 1993). The phenotypic consequences of variation in V-1 may result in part from the remarkable diversity in the predicted sizes (ranging

from 40 to 80 kDa) and pls (ranging from 4.3 to 8.5) of the Vsa proteins.

DNA inversions in the *vsa* and *hsd1* loci: a connection?

A surprising aspect of chromosomal rearrangements in *M. pulmonis* is that multiple DNA inversions apparently occur in concert. Three DNA inversions, one in *hsd1* and two in *vsa*, occurred to generate KD7-11 and 6a14 from their parent strains. Two DNA inversions, one in *hsd1* and one in *vsa*, occurred to generate 6a14/4 and KD735-16 from their parent strains. These results are surprising because multiple DNA inversions occurring more or less simultaneously should be extremely rare. All strains identified so far that have the *vsa* locus configured as shown in Fig. 1A (top) also have *hsd1* in orientation I. Progeny strains in which *hsd1* has inverted to orientation II have all undergone at least one *vsa* inversion. The converse is also true. Inversion within the *vsa* locus to alter its configuration from that shown in Fig. 1A (top) to an alternative such as shown in Fig. 1A (middle) or Fig. 1A (bottom) has only been observed in subclones that have also undergone *hsd1* inversion. Therefore, an apparent correlation exists between *hsd1* inversion and *vsa* inversion. Why would it be advantageous for the cell to co-ordinately regulate these two inversion systems? One speculative possibility is that the Vsa proteins are a component of a gene transfer process (conjugation?) that is either stimulated or impeded by restriction and modification activity.

Most site-specific DNA inversions are catalysed by members of either the invertase family (e.g. the Hin enzyme of *Salmonella*) or the bacteriophage λ integrase family of site-specific recombinases (Dybvig, 1993). The inverted repeats within which *hsd1* inversion occurs are unrelated to any sequences in or around the *vsa* genes, and there are no *vrs* box-like sequences within *hsd1*. Therefore, it is unlikely that the same recombinase would promote inversion of both the *hsd1* and the *vsa* loci. Perhaps, different site-specific recombinases act concurrently on these loci, or one of the inversion systems may stimulate recombination at the other. For example, changes in the methylation pattern of chromosomal DNA as a result of *hsd1* inversion might regulate the expression of a site-specific recombinase that recognizes the *vrs* boxes, similar to how changes in DNA methylation regulate the expression of *pap* in *E. coli* (Braaten *et al.*, 1994). Alternatively, the restriction enzyme activity induced by *hsd1* inversion may create double-stranded breaks in the chromosome that stimulate recombination activity and lead to reassortment of *vsa* sequences similarly to the manner in which mating type switching in yeast is initiated by the HO endonuclease (Dybvig, 1993; Dybvig and Yu, 1994).

Experimental procedures

Mycoplasmas

All *M. pulmonis* strains used in this study are subclones of strain UAB 6510, a pathogenic isolate that causes MRM in rats (Cassell and Davis, 1978). *M. pulmonis* was grown in mycoplasma medium at 37°C as described (Dybvig and Cassell, 1987). Strains KD7-11 and KD7-12 are previously described mutants that were isolated by using resistance to mycoplasma virus P1 as a selectable phenotype (Dybvig *et al.*, 1988). All other strains used in this study were isolated as random colonies derived by using filter cloning techniques and screened for the presence of RFLPs based on the identification of variable *HindIII* fragments that were present in DNA from some subclones but not others. The derivation and RFLP analysis of these strains are documented by Bhugra and Dybvig (1992) and Dybvig and Yu (1994) and are described here in brief. In the parent strain UAB 6510, *hsd1* is in orientation I and the *vsa* locus is configured as shown in Fig. 1A (top). KD7-11 and KD7-12 are derived from UAB 6510 and have *hsd1* inverted to orientation II and the *vsa* locus configured as shown in Fig. 1A (bottom). Strains KD735 and KD735-15 are derived from UAB 6510 and have properties indistinguishable from UAB 6510. Strain KD735-15-53 is derived from KD735-15 and has properties indistinguishable from UAB 6510 except for a 200 bp deletion corresponding to the loss of five *vsaD* repeat units. Strain KD735-16 is derived from KD735 and has *hsd1* inverted to orientation II and the *vsa* locus configured as shown in Fig. 1A (middle). Strains 6a14 and 6a14/6 are derived from KD735-15 and have *hsd1* inverted to orientation II and the *vsa* locus configured as shown in Fig. 1A (bottom). Strain 6a14/4 is derived from 6a14 and has *hsd1* inverted back to orientation I and the *vsa* locus configured as shown in Fig. 1A (middle).

DNA manipulations

Chromosomal DNA was isolated from 25 ml cultures of *M. pulmonis* as described previously (Dybvig and Alderete, 1988). Restriction fragments of *M. pulmonis* DNA were cloned into plasmid pUC18 and maintained in *E. coli* strain DH5 α as described (Bhugra and Dybvig, 1992). PCR products were cloned into the pCR-II cloning vector (Invitrogen). Plasmid DNA was isolated from *E. coli* using the alkaline lysis method and further purified by CsCl-ethidium bromide density gradient centrifugation. Restriction enzymes and T4 DNA ligase were used according to the specifications of the supplier (Gibco/BRL Life Technologies Inc.). Radiolabelling of DNA with ³²P by nick translation and Southern blotting techniques were as described (Sambrook *et al.*, 1989). Normal (high) stringency conditions were used for all hybridization experiments. The oligonucleotide primers used for PCR amplification are described in Table 1. Conditions for PCR amplification were as described (Dybvig and Woodard, 1992). To estimate the number of tandem repeat units in *vsaA* and *vsaB* and to examine size variation in *vsaD*, PCR amplification of the entire repeat regions of these genes was accomplished using primers o.6666 and o.9762 for *vsaA*, o.3909 and o.4393 for *vsaB*, and o.9349 and o.4218 for *vsaD* (Table 1).

Table 1. Oligonucleotide primer sequences used for PCR analysis.^a

Name	Orientation	Target site(s)	Sequence
o.0067	-	3365-3344	TTACATTTTGTCTCC(A/G)CTTGC
o.0067	-	3507-3488	TTACATTTTGTCTCC(A/G)CTTGC
o.0067	-	9231-9210	TTACATTTTGTCTCC(A/G)CTTGC
o.3535	+	263-281	CATCCTTTGCAAGAGCTAC
o.3909	+	1076-1095	AACCCTATGCCTAATAAACC
o.4083	-	1097-1078	TTGGTTATTAGGCATAGGG
o.4218	-	5949-5928	TAAACCAACAAAAAATGGTTC
o.4294	+	4425-4444	GCTCTTGAGATACAAGATCT
o.4393	-	2959-2940	AAATAGTTAGCTTGATTGGC
o.6666	+	6769-6792	GGTCAATCAGATGCTCAAAAAGTA
o.6859	-	7148-7130	GTAGTTGGAGGAGTCATTG
o.9272	+	4853-4875	GATACACTTCCTTCAGCTTGACC
o.9349	+	5084-5107	ATGTAAACATAGTCAACAAAAGTC
o.9590	+	5811-5831	TGGTTTTGTTTGACTACTTCC
o.9762	-	8997-8978	GAAAGTAAACAAACATAGG

a. Names of primer sequences are indicated along with the primer binding sites on the *vsA* sequence, with the nucleotide numbering corresponding to the schematic diagram depicted in Fig. 1B and in GenBank accession number U23947. The orientation of the primers on the coding (+) strand or the complementary (-) strand are also indicated. Primer o.0067 has three binding sites.

DNA sequence analysis

DNA sequencing was performed on both strands of double-stranded plasmid templates by the dideoxy nucleotide chain termination method with the Sequenase 2.0 Kit (US Biochemical Corporation). In some cases, PCR products were directly sequenced, without cloning, using the CircumVent thermal cycle DNA sequencing kit (New England Biolabs). DNA oligonucleotide primers for sequencing and PCR were supplied by the Oligonucleotide Synthesis Core Facility of the University of Alabama at Birmingham. Computer analysis was performed by using the Genetics Computer Group (GCG) programs (University of Wisconsin, Madison) on a VAX-VMS computer at the University of Alabama at Birmingham Cancer Center. The GCG program MOTIFS was used to identify consensus lipoprotein signal peptide sequences. The program FASTA was used to search the GenBank/EMBL databases for sequences related to *M. pulmonis* proteins, and the program GAP was used to calculate percentage identity between aligned sequences.

Northern hybridization

Total RNA from various *M. pulmonis* strains was isolated using the Ultraspec RNA isolation system (Biotecx Laboratories, Inc.), and 2.5 µg RNA from each strain was electrophoresed on a 1.2% agarose-formaldehyde gel and blotted to Zeta-Probe membrane (Bio-Rad Laboratories). Northern blots were probed with oligonucleotides specific for the *vsA* (5'-GTAGTTGGAGGAGTCATTG-3') or *vsB* (5'-CCATCTCTCCTGCATTTGCATC-3'). Probes were labelled with ³²P using T4 polynucleotide kinase (Pharmacia) for 30 min at 37°C (Sambrook *et al.*, 1989). Hybridization and washing conditions were according to the protocol recommended for Zeta-Probe membrane.

Synthetic peptides and generation of polyclonal antisera

Synthetic peptides were designed to the deduced amino acid sequences encoded by the repeat regions of *vsA* and *vsB*.

The *VsaA* synthetic peptide was TGSGSGSTDTKPQTP-PTTGS and the *VsaB* synthetic peptide was GKNNAGDANA-GEMGK. The peptides were synthesized commercially (QCB, Inc.) by a process that couples them to a high-density Multiple Antigenic Peptide (MAP) resin (Tam, 1988). The MAP system consists of a small immunogenically inert core molecule of radially branching lysine dendrites. The peptide antigens are synthesized directly on the branched-lysine core that is attached to a solid-phase peptide synthesis support. The result is a highly immunogenic MAP-coupled peptide molecule that eliminates the need for a carrier protein to elicit an antibody response. A standard immunization protocol was used for polyclonal antibody production. The serum was raised by subcutaneous injection of 0.2 mg peptide into a rabbit in the presence of Freund's complete adjuvant. A booster shot of the antigen was given to the animals at day 14, and in the presence of incomplete Freund's adjuvant at days 35 and 55. Serum was collected from the animals at days 80 and 85. A pre-immune serum sample was also taken from the animal prior to injection with the antigen. The rabbit pre-immune sera did not produce any reaction with immunoblots of *M. pulmonis* proteins.

Protein analysis

Organisms were harvested by centrifugation for 15 min (15 000 × *g*) at 4°C and washed twice in phosphate-buffered saline (PBS). The final pellet was suspended in PBS and the protein concentration determined by the Bio-Rad method (Bio-Rad). Proteins were resolved by one-dimensional SDS-PAGE as described by Watson *et al.* (1988) and immunoblotted by the method of Towbin (1979). In brief, 8-10 µg of protein was dissolved in SDS-dissociation buffer and loaded per well of a SDS-PAGE gel. Samples were concentrated on 4% stacking gels and separated on 10% resolving gels. Proteins were transferred from SDS-PAGE gels to nitrocellulose sheets (Bio-Rad), and non-specific binding sites were blocked by overnight incubation in blocking buffer, as described (Watson *et al.*, 1988). All immunological reactions were performed at room temperature. The immunoblots were reacted

to VsaA-antisera at 1:500, VsaB-antisera at 1:1500 or the V-1-specific 7.1-2 monoclonal antibody (described by Watson *et al.*, 1989) at 1:500 dilution in blocking buffer. The blots were subsequently reacted to biotinylated secondary antibody conjugate (goat anti-mouse IgG conjugate for the V-1 monoclonal antibody at 1:500 dilution or goat anti-rabbit IgG conjugate for the Vsa-specific antisera at 1:4000 dilution). Bound conjugate was detected by reaction with 4-chloro-1-naphthol (HRP color development reagent, Bio-Rad).

Acknowledgements

We thank Jerry Simecka and Padma Patel for help with Northern blots and Portia Caldwell for help with media preparation and RFLP analysis. We thank Gail Cassell for her continued support and guidance. This research was supported by US Public Health Service grants AI31144 and GM51126 (to K. D.) from the National Institutes of Health. L. L. V. was supported by training grant T32 AI07041 from the PHS. The Genetics Computer Group (GCG) programs were made available to us through the University of Alabama at Birmingham (UAB) Center for AIDS Research, supported by PHS Grant P30 AI27767. Synthesis of DNA primers was supported by NCI grant CA13148 to the UAB Comprehensive Cancer Center.

References

- Barbour, A.G. (1990) Antigenic variation of a relapsing fever *Borrelia* species. *Annu Rev Microbiol* **44**: 155–171.
- Bhugra, B., and Dybvig, K. (1992) High-frequency rearrangements in the chromosome of *Mycoplasma pulmonis* correlate with phenotypic switching. *Mol Microbiol* **6**: 1149–1154.
- Braaten, B.A., Nou, X., Kaltenbach, L.S., and Low, D.A. (1994) Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli*. *Cell* **76**: 577–588.
- Cassell, G.H., and Davis, J.K. (1978) Protective effect of vaccination against *Mycoplasma pulmonis* respiratory disease in rats. *Infect Immun* **21**: 69–75.
- Cassell, G.H., Davis, J.K., Wilborn, W.H., and Wise, K.S. (1978) Pathobiology of mycoplasmas. In *Microbiology*. Schlessinger, D. (ed.). Washington, DC: American Society of Microbiology, pp. 399–403.
- Cassell, G.H., Clyde, W.A., and Davis, J.K. (1985) Mycoplasmal respiratory infections. In *The Mycoplasmas*, Vol. IV. Razin, S., and Barile, M.F. (eds). Orlando: Academic Press, pp. 65–106.
- Dallo, S.F., Chavoya, A., and Baseman, J.B. (1990) Characterization of the gene for a 30-kilodalton adhesin-related protein of *Mycoplasma pneumoniae*. *Infect Immun* **58**: 4163–4165.
- Davis, J.K., Thorp, R.B., Maddox, P.A., Brown, M.B., and Cassell, G.H. (1982) Murine respiratory mycoplasmosis in F344 and LEW rats: evolution of lesions and lung lymphoid cell populations. *Infect Immun* **36**: 720–729.
- Davis, J.K., Parker, R.F., White, H., Dziedzic, D., Taylor, G., Davidson, M.K., Cox, N.R., and Cassell, G.H. (1985) Strain differences in susceptibility to murine respiratory mycoplasmosis in C57BL/6 and C3HeN mice. *Infect Immun* **50**: 647–654.
- Dybvig, K. (1990) Mycoplasmal genetics. *Annu Rev Microbiol* **44**: 81–104.
- Dybvig, K. (1993) DNA rearrangements and phenotypic switching in prokaryotes. *Mol Microbiol* **10**: 465–471.
- Dybvig, K., and Alderete, J. (1988) Transformation of *Mycoplasma pulmonis* and *Mycoplasma hyorhinis*: transposition of Tn916 and formation of cointegrate structures. *Plasmid* **20**: 33–41.
- Dybvig, K., and Cassell, G.H. (1987) Transposition of Gram-positive transposon Tn916 in *Acholeplasma laidlawii* and *Mycoplasma pulmonis*. *Science* **235**: 1392–1394.
- Dybvig, K., and Woodard, A. (1992) Cloning and DNA sequence of a mycoplasmal *recA* gene. *J Bacteriol* **174**: 778–784.
- Dybvig, K., and Yu, H. (1994) Regulation of a restriction and modification system via DNA inversion. *Mol Microbiol* **12**: 547–560.
- Dybvig, K., Alderete, J., Watson, H.L., and Cassell, G.H. (1988) Adsorption of mycoplasma virus P1 to host cells. *J Bacteriol* **170**: 4373–4375.
- Dybvig, K., Simecka, J.W., Watson, H.L., and Cassell, G.H. (1989) High-frequency variation in *Mycoplasma pulmonis* colony size. *J. Bacteriol.* **171**: 5165–5168.
- Facijs, D., and Meyer, T.F. (1993) A novel determinant (*comA*) essential for natural transformation competence in *Neisseria gonorrhoeae* and the effect of a *comA* defect on pilin variation. *Mol Microbiol* **10**: 699–712.
- Fischetti, V.A., Pancholi, V., and Schneewind, O. (1991) Common characteristics of the surface proteins from Gram-positive cocci. In *Genetics and Molecular Biology of Streptococci, Lactococci, and Enterococci*. Dunny, G.M., Cleary, P.P., and McKay, L.L. (eds). Washington, DC: American Society for Microbiology, pp. 290–294.
- Hayashi, S., and Wu, H.C. (1990) Lipoproteins in bacteria. *J Bioenerg Biomembr* **22**: 451–471.
- Hill, S.A., Morrison, S.G., and Swanson, J. (1990) The role of direct oligonucleotide repeats in gonococcal pilin gene variation. *Mol Microbiol* **4**: 1341–1352.
- Kemp, D.J., Coppell, R.L., and Anders, R.F. (1987) Repetitive proteins and genes of malaria. *Annu Rev Microbiol* **41**: 181–208.
- Kitten, T., and Barbour, A.G. (1990) Juxtaposition of expressed and variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. *Proc Natl Acad Sci USA* **87**: 6077–6081.
- Komano, T., Kim, S.R., and Nisioka, T. (1987) Distribution of shufflon among IncI plasmids. *J Bacteriol* **169**: 5317–5319.
- Krause, D.C., and Taylor-Robinson, D. (1992) Mycoplasmas which infect humans. In *Mycoplasmas: Molecular Biology and Pathogenesis*. Maniloff, J., McElhaney, R.N., Finch, L.R., and Baseman, J.B. (eds). Washington, DC: American Society for Microbiology, pp. 417–444.
- Ladefoged, S.A., Birkelund, S., Hauge, S., Brock, B., Jensen, L.T., and Christiansen, G. (1995) A 135-kilodalton surface antigen of *Mycoplasma hominis* PG21 contains multiple directly repeated sequences. *Infect Immun* **63**: 212–223.
- Levinson, G., and Gutman, G.A. (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol* **4**: 203–221.
- McDevitt, D., Francois, P., Vaudaux, P., and Foster, T.J. (1994) Molecular characterization of the clumping factor

- (fibrinogen receptor) of *Staphylococcus aureus*. *Mol Microbiol* **11**: 237–248.
- Plasterk, R.H.A., and van de Putte, P. (1984) Genetic switches by DNA inversions in prokaryotes. *Biochim Biophys Acta* **782**: 111–119.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Segal, E., Hagblom, P., Seifert, H.S., and So, M. (1986) Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. *Proc Natl Acad Sci USA* **83**: 2177–2181.
- Simecka, J.W., Davis, J.K., Davidson, M.K., Ross, S.E., Stadtlander, C.K.-H., and Cassell, G.H. (1992) Mycoplasma diseases of animals. In *Mycoplasmas: Molecular Biology and Pathogenesis*. Maniloff, J., McElhaney, R.N., Finch, L.R., and Baseman, J.B. (eds). Washington, DC: American Society for Microbiology, pp. 391–415.
- Talkington, D.F., Fallon, M.T., Watson, H.L., Thorp, R.K., and Cassell, G.H. (1989) *Mycoplasma pulmonis* V-1 antigen surface protein variation: occurrence *in vivo* and association with lung lesions. *Microbial Pathogen* **7**: 429–436.
- Tam, J.P. (1988) Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proc Natl Acad Sci USA* **85**: 5409–5413.
- Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci USA* **76**: 4350–4354.
- Watson, H.L., McDaniel, L.S., Blalock, D.K., Fallon, M.T., and Cassell, G.H. (1988) Heterogeneity among strains and a high rate of variation within strains of a major surface antigen of *Mycoplasma pulmonis*. *Infect Immun* **56**: 1358–1363.
- Watson, H.L., Dybvig, K., Blalock, D.K., and Cassell, G.H. (1989) Subunit structure of the variable V-1 antigen of *Mycoplasma pulmonis*. *Infect Immun* **57**: 1684–1690.
- Watson, H.L., Zheng, X., and Cassell, G.H. (1993) Structural variations and phenotypic switching of mycoplasmal antigens. *Clin Inf Dis* **17**(Suppl 1): S183–S186.
- Wieslander, A., Boyer, M.J., and Wroblewski, H. (1992) Membrane protein structure. In *Mycoplasmas: Molecular Biology and Pathogenesis*. Maniloff, J., McElhaney, R.N., Finch, L.R., and Baseman, J.B. (eds). Washington, DC: American Society for Microbiology, pp. 93–112.
- Wise, K.S., Yogev, D., and Rosengarten, R. (1992) Antigenic variation. In *Mycoplasmas: Molecular Biology and Pathogenesis*. Maniloff, J., McElhaney, R.N., Finch, L.R., and Baseman, J.B. (eds). Washington, DC: American Society for Microbiology, pp. 473–490.
- Wren, B.W. (1991) A family of clostridial and streptococcal ligand-binding proteins with conserved C-terminal repeat sequences. *Mol Microbiol* **5**: 797–803.
- Wu, H.C. (1987) Post-translational modification and processing of membrane proteins in bacteria. In *Bacterial Outer Membranes as Model Systems*. Inouye, M. (ed.). New York: Wiley, pp. 37–71.
- Yogev, D., Rosengarten, R., Watson-McKown, R., and Wise, K.S. (1991) Molecular basis of *Mycoplasma* surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5' regulatory sequences. *EMBO J* **10**: 4069–4079.
- Zheng, X., Teng, L.-J., Watson, H., Glass, J.I., Blanchard, A., and Cassell, G.H. (1995) Small repeating units within the *Ureaplasma urealyticum* MB antigen gene encode serovar specificity and are associated with antigen size variation. *Infect Immun* **63**: 891–898.