Variable Lipoprotein Genes of *Mycoplasma agalactiae* Are Activated In Vivo by Promoter Addition via Site-Specific DNA Inversions

Ravenna Flitman-Tene,1 Sigalit Mudahi-Orenstein,1 Sharon Levisohn,2 and David Yogev1*

*Department of Membrane and Ultrastructure Research, The Hebrew University-Hadassah Medical School, Jerusalem 91120,1 and Mycoplasma Unit, Kimron Veterinary Institute, Beit Dagan 50250,2 Israel

Received 2 December 2002/Returned for modification 25 February 2003/Accepted 26 March 2003

*Mycoplasma agalactiae*, the etiological agent of contagious agalactia of small ruminants, has a family of related genes (*avg* genes) which encode surface lipoprotein antigens that undergo phase variation. A series of 13 *M. agalactiae* clonal isolates, obtained from one chronically infected animal over a period of 7 months, were found to undergo major rearrangement events within the *avg* genomic locus. We show that these rearrangements regulate the phase-variable expression of individual *avg* genes. Northern blot analysis and reverse transcription-PCR showed that only one *avg* gene is transcribed, while the other *avg* genes are transcriptionally silent. Sequence analysis and primer extension experiments with two *M. agalactiae* clonal isolates showed that a specific 182-bp *avg* 5’ upstream region (*avg*-B2) that is present as a single chromosomal copy serves as an active promoter and exhibits a high level of homology with the vsp promoter of the bovine pathogen *Mycoplasma bovis*. PCR analysis showed that each *avg* gene is associated with the *avg*-B2 promoter in a subpopulation of cells that is present in each subclone. Multiple sequence-specific sites for DNA recombination (vis-like), which are presumably recognized by site-specific recombinase, were identified within the conserved *avg* 5’ upstream regions of all *avg* genes and were found to be identical to the recombination sites of the *M. bovis* vsp locus. In addition, a gene encoding a member of the integrase family of tyrosine site-specific recombinases was identified adjacent to the variable *avg* locus. The molecular genetic basis for *avg* phase-variable expression appears to be mediated by site-specific DNA inversions occurring in vivo that allow activation of a silent *avg* gene by promoter addition. A model for the control of *avg* genes is proposed.

*Mycoplasma agalactiae* belongs to the genus *Mycoplasma*, which comprises more than 180 wall-less species that represent the smallest self-replicating life forms on earth and phylogenetically are related to gram-positive eubacteria (28, 29). Most mycoplasmas have been identified as infectious agents of human or other animals (29, 36). *M. agalactiae*, commonly known as the cause of contagious agalactia of sheep and goats, is one of the most devastating mycoplasmal pathogens of small ruminants (3). The disease is expressed clinically as mastitis, arthritis, and keratoconjunctivitis and is of a considerable economic importance worldwide, particularly in the Mediterranean basin (3). The pathogen is present in milk and other body fluids of infected animals and also in animals that do not exhibit clinical signs. These asymptomatic carrier animals play an important role in transmission of the disease (3). The successful persistence of *M. agalactiae* within hosts indicates that frequently the host defense mechanisms fail to eliminate the bacterial pathogen. Recent advances in the molecular biology of mycoplasmas have indicated that high-frequency variation of surface components, particularly lipoproteins, plays an important role in establishing the chronic nature of mycoplasma infections and is an important parameter in the interaction of these small wall-less bacteria with their hosts (15, 39, 40, 42). Several studies in recent years have shown that despite their limited genome capacity, the mycoplasmas are replete with mutation-based systems that provide variation in expression and structure of specific gene products as an adaptive strategy for survival (4, 12, 13, 24, 35, 41, 42).

*M. agalactiae* is phylogenetically closely related to *Mycoplasma bovis*, an important bovine pathogen that causes mastitis, arthritis, and respiratory disease and is capable of producing subacute to acute inflammations of various organs (27). The two species exhibit 99% homology in their 16S rRNA genes (26) but less than 40% homology at the genomic level (2). In previous reports it was shown that *M. bovis* contains an elaborate genetic system in which multiple related but divergent genes encoding variable surface lipoproteins (Vsp's) undergo spontaneous on-off switching, which generates extensive surface antigenic variation (20, 25). The Vsp antigens were shown to be highly immunogenic and to contain adhesive structures in the repetitive domains of the Vsp molecules (32). Site-specific DNA inversions that occur within the vsp locus, as well intrachromosomal recombination between closely related vsp genes, dictate the degree of Vsp diversification within propagating populations by determining which vsp gene is expressed in a given cell (20, 21, 25).

Recently, a chromosomal region containing a cluster of lipoprotein-encoding genes (designated the *avg* genes) that undergo in vivo rearrangements in naturally infected animals was identified in *M. agalactiae* (9, 10). The *avg* genes exhibit high levels of homology in their 5’ upstream regions and in their N-terminal encoding regions with the vsp genes of *M. bovis*, while the rest of the *Avg* molecule shows considerable sequence divergence. Adjacent to the *vsp* and *avg* loci there is an open reading frame that exists as a single chromosomal copy and is predicted to encode a site-specific tyrosine recombinase (31).

* Corresponding author. Mailing address: Department of Membrane and Ultrastructure Research, The Hebrew University-Hadassah Medical School, Jerusalem, 91120, Israel. Phone: 972-2-6758176. Fax: 972-2-6784010. E-mail: yogev@cc.huji.ac.il.
pears to be mediated by site-specific restriction endonuclease digestion, gel electrophoresis of DNA, and Southern blot hybridization were performed as previously described (19, 20). The conditions used for oligonucleotides were synthesized with a model 380B DNA synthesizer (Applied Biosystems, Foster City, Calif.). Sequences of both strands were confirmed by partial sequencing (21, 22) as well as by digestion and autoradiographic analysis. 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), isopropyl-β-D-thiogalactopyranoside (IPTG), and ampicillin were purchased from Sigma Chemical Co. (St. Louis, Mo.). [γ-32P]ATP, [α-32P]CTP, and α-32P-labeled deoxynucleoside triphosphates were purchased from Amersham (Little Chalfont, United Kingdom).

**MATERIALS AND METHODS**

**Bacterial strains, plasmids, chemicals, and growth conditions.** Type strain *M. agalactiae* PG2 was originally obtained from D. G. E. Edward, Wellcome Research Laboratories, Beckenham, Kent, England. Isolation of *M. agalactiae* clinical isolates has been described elsewhere (10). *M. agalactiae* strains were grown in mycoplasma medium based on heart infusion broth (Difco) containing 0.5% *E. coli* in mycoplasma medium (10). *M. agalactiae* (Stratagene, La Jolla, Calif.) was used as a host. Recombinant clones were constructed in the plasmid vector pKS (36) and subjected to PCR by using primers pf-1 avgF and pf-2 avgF. PCR products were purified and sequenced using High Pure filter columns (Boehringer Mannheim) and directly sequenced. The PCR primers used in this study are listed in Table 1.

**RT-PCR.** Primers rb-1 and rb-2 were designed from the avgB gene sequence to enable synthesis of a PCR product that was 533 bp long and spanned the avgB intergenic region. Mycoplasma RNA from isolates 627#2 and 627#4 was treated with RNase-free DNase I (MBI Fermentas) and then reverse transcribed by using the RevertAid Moloney murine leukemia virus reverse transcriptase (MBI Fermentas). The reverse transcription (RT)-PCR mixture containing 2 μg of RNA and 20 pmol of primer rb-1 was incubated at 70°C for 5 min and then chilled on ice. Four microliters of 5× reaction buffer (MBI Fermentas), each of the four deoxynucleoside triphosphates (MBI Fermentas) at a concentration of 2.5 mM, and 20 U of RNase inhibitor (MBI Fermentas) were added, and the mixture was incubated at 37°C for 5 min. Then 200 U of RevertAid Moloney murine leukemia virus reverse transcriptase (MBI Fermentas) was added, and the reaction mixture was incubated at 42°C for 20 min. The reaction was stopped by heating the mixture at 70°C for 10 min. The resultant cDNA product was subjected to PCR by using primers rb-1 and rb-2 and the following program: 3 min at 95°C and then 25 cycles of 35 s at 95°C, 45 s at 55°C, and 35 s at 72°C. The resulting PCR products were purified by using High Pure filter columns (Boehringer Mannheim) and directly sequenced. The PCR primers used in this study are listed in Table 1.

**RESULTS**

Identification of an avg upstream region displaying homology to the vsp promoter of *M. bovis* and undergoing in vivo rearrangements. Recently, it was shown that a specific vsp upstream region (designated the A2 cassette) serves as an

<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pa-2</td>
<td>avgA</td>
<td>GATCCGTACACCCCTTCAAG</td>
</tr>
<tr>
<td>pb-1</td>
<td>avgB</td>
<td>GATACCCGGCCTACCCATT</td>
</tr>
<tr>
<td>rb-1</td>
<td>avgB</td>
<td>CGTAAATCCATCCCTGTAAG</td>
</tr>
<tr>
<td>rb-2</td>
<td>avgB</td>
<td>CGCTTAATACCAAGGGG</td>
</tr>
<tr>
<td>pes-1</td>
<td>avgB, catase</td>
<td>AGCTCAACTTCTTCTACT</td>
</tr>
<tr>
<td>pp-1</td>
<td>avg-B, catase</td>
<td>AGTAGAAAGATGTATGCA</td>
</tr>
<tr>
<td>pc-1</td>
<td>avgC</td>
<td>CAGTGTTGTCGGCCGTTTTA</td>
</tr>
<tr>
<td>pc-2</td>
<td>avgC</td>
<td>CCTGTTTGATGCTGTTGGC</td>
</tr>
<tr>
<td>pd-1</td>
<td>avgD</td>
<td>GTGACGCGTAGTAAATCTG</td>
</tr>
<tr>
<td>pe-1</td>
<td>avgE</td>
<td>CGCCCTTTTCTCCTGGGAC</td>
</tr>
<tr>
<td>pf-1</td>
<td>avgF</td>
<td>GGGTATAGCAGGTATGAC</td>
</tr>
<tr>
<td>pf-2</td>
<td>avg-F, catase</td>
<td>GCCCTATATAATAGGCC</td>
</tr>
</tbody>
</table>

* Data from this study.

**TABLE 1. Oligonucleotides used in this study**

sodium dodecyl sulfate, 50 μg of salmon sperm DNA per ml, 6× SSPE (1× SSPE is 0.18 M NaCl plus 10 mM NaPO₄ [pH 7.7]), and 1 mM EDTA (33). Hybridization was performed at 42°C. The membranes were washed twice for 10 min at room temperature and once at 42°C in 6× SSPE–0.1% sodium dodecyl sulfate. The membranes were dried and autoradiographed by using Super RX Fuji X-ray film (Fujif, Tokyo, Japan). A PCR product spanning the highly conserved signal sequence common to all avg genes was amplified by using primers pp-1 and pb-1 (Table 1) and was used as a probe in a Northern blot analysis to monitor the presence of avg mRNA.

*Primer extension.* Primer pc-1 was end labeled by using T4 polynucleotide kinase and [γ-32P]ATP and then purified on a G-50 mini column (Boehringer Mannheim GmbH, Indianapolis, Ind.). The primer extension reaction mixture (final volume, 15 μl) contained 2 μg of total RNA, 100 ng of labeled primer, and 3.9 μl of reaction buffer (0.1 M Tris-HCl [pH 8.3]; 0.14 M KCl, 10 mM MgCl₂, 10 mM dithiothreitol). The reaction mixture was incubated for 10 min at 75°C, followed by 5 min at room temperature. After annealing, a deoxynucleoside triphosphate mixture containing each deoxynucleoside triphosphate at a concentration of 2.5 mM (Boehringer Mannheim) and 5 U of avian myeloblastosis virus reverse transcriptase (Promega) were added. The reaction mixture was incubated at 42°C for 60 min. RNase (66 μg/ml) was then added, and the reaction mixture was incubated for 30 min at 37°C, then the reaction was terminated by heating at 100°C for 10 min. Primer extension products were mixed with loading buffer and resolved on a 6% polyacrylamide sequencing gel. DNA sequence analysis was performed by the dideoxy chain termination method (34) by using a Thermo Sequenase radiolabeled terminator cycle sequencing kit (Amersham).

PCR. PCR products were carried out in 25-μl mixtures containing 10 ng of template DNA, 1 U of ExTaq DNA polymerase in 1× T-ES buffer (MBI Fermentas), 2 mM MgSO₄, each deoxynucleoside triphosphate at a concentration of 0.2 mM, and 125 ng of each primer. PCR amplification was performed by using a Perno- sonal cycler (Biometra, Gottingen, Germany) programmed for 31 cycles as follows: one cycle of 3 min at 95°C, 45 s at 55°C, and 35 s at 72°C, followed by 25 cycles of 35 s at 95°C, 45 s at 55°C, and 30 s at 72°C. The reaction mixtures were then incubated for 10 min at 72°C and allowed to cool to 10°C. The resultant PCR products were purified by using High Pure filter columns (Boehringer Mannheim) and directly sequenced. The PCR primers used in this study are listed in Table 1.

RT-PCR. Primers rb-1 and rb-2 were designed from the avgB gene sequence to enable synthesis of a PCR product that was 533 bp long and spanned the avgB intergenic region. Mycoplasma RNA from isolates 627#2 and 627#4 was treated with RNase-free DNase I (MBI Fermentas) and then reverse transcribed by using the RevertAid Moloney murine leukemia virus reverse transcriptase (MBI Fermentas). The reverse transcription (RT)-PCR mixture containing 2 μg of RNA and 20 pmol of primer rb-1 was incubated at 70°C for 5 min and then chilled on ice. Four microliters of 5× reaction buffer (MBI Fermentas), each of the four deoxynucleoside triphosphates (MBI Fermentas) at a concentration of 2.5 mM, and 20 U of RNase inhibitor (MBI Fermentas) were added, and the mixture was incubated at 37°C for 5 min. Then 200 U of RevertAid Moloney murine leukemia virus reverse transcriptase (MBI Fermentas) was added, and the reaction mixture was incubated at 42°C for 20 min. The reaction was stopped by heating the mixture at 70°C for 10 min. The resultant cDNA product was subjected to PCR by using primers rb-1 and rb-2 and the following program: 3 min at 95°C and then 25 cycles of 35 s at 95°C, 45 s at 56°C, and 35 s at 72°C. The resultant PCR products were visualized after gel electrophoresis and ethidium bromide staining.

DNA sequence analysis. DNA sequence analysis of both strands was performed by the dideoxy chain termination method (34). The T7 promoter sequence and the T3 sequence located on the pKS vector, as well as the amplified sequences, were used as primers. Sequencing was done with a dye terminator cycle sequencing automatic sequencer (model ABI PRISMA 377; Perkin-Elmer, Foster City, Calif.). Sequence data were analyzed by using the AssemblyLIGN and MacVector 6.5.3 software.

**Nucleotide sequence accession numbers.** The nucleotide sequences of the *M. agalactiae* avg genomic region locus determined in this study have been deposited in the GenBank database under accession number AY195887.
identical nucleotides are indicated by shading. The shading of each gene is indicated on the left. The numbers above the sequences indicate nucleotide positions relative to the initiation codon. The positions of the two upstream cassettes (cassette no.1 and cassette no.2) are indicated on the right. The positions of the putative transcriptional start sites (+1) for the avg-B2 and vsp-A2 regions (20) are indicated by arrows. The positions of prokaryotic ρ-dependent consensus sequences (−10) and a ribosome binding site (SD) are indicated by solid lines. The position of the 35-bp putative DNA recombination site (vis) (20, 31) is indicated by a dotted line. The binding site and its orientation for the pex-I oligonucleotide are indicated by a solid line.

FIG. 2. In vivo rearrangements of M. agalactiae avg-B2-bearing genomic fragments in successive isolates from a chronically infected ewe. HindIII-digested genomic DNAs from 13 clinical isolates (lanes 2 to 14) and from the M. agalactiae PG2 type strain (lane 1) were subjected to Southern blot hybridization with an avg-B2-specific oligonucleotide probe (pex-I). The isolates were obtained from a single naturally infected animal, designated animal #627, over a period of 7 months, and M. agalactiae was consistently isolated at each sampling time (at 2- to 4-week intervals) (10). The positions of molecular size markers are indicated on the left.
mosomal copy of an open reading frame, designated mar, that was predicted to encode a site-specific DNA recombinase that belongs to the λ integrase family of tyrosine recombinases (31). Comparison of the nucleotide sequences of the avg loci of the two isolates revealed that major rearrangement events affecting the avg gene configuration occurred in the population during in vivo passage. For example, a 2.2-kb fragment carrying the avgA and avgE genes in isolate 627#3 (Fig. 3A) was found in isolate 627#4 in an inverted orientation (Fig. 3B). Also, the avgB gene that was located upstream of the avgC gene (Fig. 3A) was inverted and repositioned upstream of the mar gene (Fig. 3B). Moreover, a change in the organization of several upstream cassettes was observed when the two isolates were compared. Cassettes F2, D2, C2, B2, A2, and E2, which in isolate 627#3 were located upstream of the avgF, avgD, avgC, avgB, avgA, and avgE genes, respectively (Fig. 3A), in isolate 627#4 were found upstream of the avgB, avgF, avgD, avgC, avgE, and avgA genes, respectively (Fig. 3B). Notably, in addition to the variations in avg gene organization, two changes were observed in the avgB and avgC structural genes. The avgB gene in isolate 627#4 contained a stop codon at nucleotide position 496 which resulted in premature termination of translation. In addition, in isolate 627#3 there were four repeats in the avgC gene instead of the five repeats found in isolate 627#4.

The deduced proteins encoded by the six avg genes, derived from isolate 627#3, were compared, and a schematic representation of the Avg structure is shown in Fig. 4. In addition to the conserved S' noncoding sequence, the Avg N-terminal region consisting of 25 amino acids is also a highly homologous domain; there is 93% amino acid identity in the Avg proteins and 90% homology to the corresponding region of the Vsp proteins of M. bovis (19). This peptide is a typical prokaryotic

FIG. 3. Schematic representation, partial restriction map, genomic organization, and comparison of the M. agalactiae avg locus for two isolates obtained 2 weeks apart from the same chronically infected ewe (animal #627). The thick solid line represents about 10 kb of the avg locus of M. agalactiae isolate 627#3 (A) or isolate 627#4 (B). The positions of HindIII (H), EcoRI (E), and XbaI (X) restriction sites are indicated. The locations and orientations of six avg genes are indicated by large shaded arrows. A solid box 5' to each avg gene represents homologous cassette no. 1. Cassette no. 2 of each avg gene is indicated. The position of the avg-B3 promoter cassette is indicated by a stippled box. The location of the putative tyrosine recombinase gene mar is indicated by a cross-hatched arrow. The open triangles indicate the location of the DNA recombination vir-like sites within conserved cassette no. 1 of each avg gene. The locations of oligonucleotide primers used in this study for PCR amplification across the junction between the avg-B3 promoter and the avg structural genes are indicated by small arrows.

FIG. 4. Structural features and comparison of avg genes and Avg products derived from isolate 627#3. The structures of each avg gene and predicted Avg protein are indicated schematically by aligned rectangles. A highly conserved cassette no. 1 and the cassette no. 2 upstream of each avg gene are indicated by the first two boxes. The avg-B3 promoter cassette is indicated by a gray box. The third box (solid box labeled S) represents a highly homologous 75-bp DNA sequence encoding a conserved prolipoprotein signal peptide. A 6-amino-acid sequence common to all Avgs is indicated by the fourth box (shaded box). In-frame reiterated coding sequences within the AvgB (Rb1 to Rb2), AvgC (Rc1 to Rc4), and AvgE (Rg1 to Rg4) proteins are represented by cross-hatched boxes. A 56-amino-acid region present once in the AvgA protein and twice in the AvgB protein or a 134-amino-acid region present in both the AvgD and AvgF proteins are cross-hatched differently. The numbers on the right indicate the length of each polypeptide chain.
lipoprotein signal peptide, which begins with a sequence consisting of three positively charged Lys residues that is followed by a core of 20 hydrophobic amino acids and ends with the tetrapeptide Ala-Ala-Lys-Cys. The remaining portions of the mature Avg proteins show sequence divergence. Notably, unlike the abundance of repetitive domains throughout the Vsp molecules of *M. bovis* (19), only three Avg proteins (AvgB, AvgC, and AvgE) contained reiterated coding sequences consisting of 186, 86, and 91 amino acids (RB, RC, and RE, respectively) (Fig. 4). The RB, RC, and RE sequences were unique to the corresponding proteins and were not found in other Avg proteins.

**avg upstream region B2 serves as an active promoter.** The high level of homology of the avg-B2 cassette to the vsp-A2 promoter and the finding that in clinical isolates it was located upstream of different avg genes (avgB and avgC [Fig. 3A and B, respectively]) raised the possibility that the avg-B2 cassette is, in fact, an active promoter and that acquisition of this element by other members of the avg family activates the recipient gene. We therefore examined whether the avg-B2 recipient genes, namely, the avgB gene in isolate 627#3 (Fig. 3A) and the avgC gene in isolate 627#4 (Fig. 3B), were transcribed.

Total cellular RNAs of the two *M. agalactiae* isolates were extracted and subjected to Northern blot analysis by using two oligonucleotides complementary to a unique sequence of the avgB or avgC structural genes (pb-1 and pc-1, respectively). A single transcript was detected by the pb-1 probe only in isolate 627#3 (Fig. 5A, lane 1), while a single transcript was detected by the pc-1 probe only in isolate 627#4 (Fig. 5B, lane 2). To determine whether additional avg-related mRNA might be present in the two isolates, a Northern blot analysis was performed by using a PCR product spanning the highly conserved

![Image](image_url)
avg upstream region and the N-terminal region present in all known avg genes as a probe. No additional avg mRNAs were detected, suggesting that in the two clonal isolates tested the avg-B2 recipient gene is transcribed, while the other avg genes are transcriptionally silent (data not shown).

The presence of avg mRNA corresponding to the avg-B2 recipient gene was also monitored by RT-PCR. Primers rtb-1 and rtb-2 were designed from the avgB gene sequence to enable synthesis of a 533-bp PCR product. RT-PCR amplification of the avgB gene in isolate 627#3 and in isolate 627#4 indicated that the avgB gene was transcribed only in isolate 627#3 (Fig. 5C, lane 1). No product was generated in isolate 627#4 (Fig. 5C, lane 2). Importantly, control reaction mixtures containing total cellular RNA that had been treated with DNase without reverse transcriptase yielded no PCR product (data not shown), while total genomic DNA of the both isolates yielded the expected 533-bp PCR product (Fig. 5C, lanes 3 and 4, respectively).

The role of the avg-B2 cassette as an active promoter was further examined by mapping the avg transcription start site for each of the avg-B2-recipient genes (avgB and avgC) by using pes-1 as a primer with total cellular RNA obtained from isolates 627#3 and 627#4. A potential transcriptional start site was identified within the avg-B2 cassette for the avgB gene and for the avgC gene (Fig. 5D and E). This site was located 205 and 189 bp upstream of the initiation codons of the avgB and avgC genes, respectively (Fig. 1 and 5). The start site was preceded at the appropriate spacing by the sequence TAT AAT, corresponding to the prokaryotic σ70-dependent −10 consensus sequence. It should be noted that additional primer extension products were observed, which may be attributed to the presence of subpopulations in each isolate. Collectively, the data indicate that the avg-B2 cassette serves as an active promoter that allows transcription initiation of the recipient silent gene.

Each avg gene is associated with the avg-B2 promoter in a subpopulation of cells that is present in each clonal isolate. In a previous study, it was shown that each of the M. bovis vsp genes has in its upstream cassette no. 1 a putative DNA binding site (vis), which was identified as the recombination site for DNA inversions (20, 31). Each of the M. agalactiae avg genes that was analyzed in this study has in its upstream region an analogous vis-like site (with only one base difference compared to the vis site of M. bovis) for DNA recombination (Fig. 1). This finding, together with the finding that in each of the two isolates analyzed the avg-B2 promoter was found upstream of a distinct avg gene, suggests that a propagating population of M. agalactiae (even a propagating population of a clonal isolate) contains, in fact, an array of cells that exhibit different avg configurations in which the avg-B2 promoter can be located upstream of any avg gene. To verify this, primers were designed to PCR amplify any avg structural gene associated with the B2 active promoter. Five primers, representing unique sequences of five avg structural genes (avgA, avgB, avgC, avgD, and avgE; primers pa-2, pb-1, pc-2, pd-1, and pe-1, respectively) and a primer representing a unique sequence of the avg-B2 promoter (primer pp-1), were used. Genomic DNAs for PCRs were obtained from the M. agalactiae PG2 type strain, from M. agalactiae clonal isolates 627#3 and 627#4 (Fig. 3A and B, respectively), and from an additional isolate, isolate 627#5 isolated from animal #627 (Fig. 2, lane 5). PCR amplification of populations of DNAs from each of these strains (Fig. 6) yielded products that corresponded to amplification across the junction between the avg-B2 promoter and avgA (when primers pp-1 and pa-2 were used) (lane 1), avgB (when primers pp-1 and pb-1 were used) (lane 2), avgC (when primers pp-1 and pd-2 were used) (lane 3), avgD (when primers pp-1 and pd-1 were used) (lane 4), and avgE (when primers pp-1 and pe-1 were used) (lane 5). Sequence analysis of the resultant PCR products confirmed the presence of the avg-B2 promoter upstream of the avgA, avgB, avgC, avgD, and avgE, structural genes in the various M. agalactiae populations tested. These data demonstrate that each clonal isolate of M. agalactiae contains subpopulations of cells that have a mixture of avg configurations in which each avg gene is associated with the avg-B2 promoter.

DNA inversion model for phase-variable expression of the avg genes. Comparison of the nucleotide sequences at the 5’ and 3’ ends of the inverted fragments in isolates 627#3 and 627#4 (Fig. 3) revealed that the inversion events occurred within the 35-bp vis-like site present in cassette no. 1 (Fig. 1). These findings and the proximity of a gene encoding a site-specific tyrosine recombinase (Mar) to the avg locus (Fig. 3) (31) suggest that site-specific DNA inversion is a possible mechanism for control of avg genes. It should be noted that vis site-mediated DNA inversion can occur between vis copies that are in opposite orientations in the chromosome but not be-

![FIG. 6. Association of the avg-B2 promoter with avg genes in a subpopulation of cells present in each subclone: PCR amplification of M. agalactiae isolates. Total genomic DNAs of the M. agalactiae PG2 type strain (A), clonal isolate 627#3 (B), clonal isolate 627#4 (C), and clonal isolate 627#5 (D) were subjected to PCR amplification across the junction between the avg-B2 promoter and five avg structural genes. The following primer pairs were used: for the avgA gene, pp-1 and pa-2 (lane 1); for the avgB gene, pp-1 and pb-1 (lane 2); for the avgC gene, pp-1 and pc-2 (lane 3); for the avgD gene, pp-1 and pd-1 (lane 4); and for avgE gene, pp-1 and pe-1 (lane 5). The size of each PCR product (in base pairs) is indicated above the band.](http://iai.asm.org/content/38/10/3826/F6)
between copies that are oriented in the same direction. Comparison of the avg configurations of isolates 627#3 and 627#4 indicated that at least four distinct DNA inversions are required to recombine the avgC gene with the avg-B2 promoter and to generate the avg configuration identified in isolate 627#4 (Fig. 7). The locations and directions of the four putative inversion events are unknown. One DNA inversion could occur between the vis sites of the avgA and avgE genes (Fig. 7A), giving rise to a 2.2-kb DNA inversion and to the avg configuration illustrated in Fig. 7B. A second inversion is postulated to occur between the vis sites of the avgF and avgC genes, leading to exchange of the F2 and C2 cassettes between the two genes and to orientation of the avgC gene in the chromosome in the direction opposite the direction of the avg-B2 promoter, generating the avg configuration shown in Fig. 7C. A third DNA inversion of a 5.3-kb fragment could then occur between the vis sites of avgC and avgB, generating the avg configuration shown in Fig. 7D. A fourth inversion of 2.3 kb between the vis sites of avgF and avgD could then yield the configuration identified in isolate 627#4 (Fig. 7E). The existence of the avg configuration shown in Fig. 7C was verified by PCR amplification of the region between the mar gene and the avgC gene, as well as the region between the F2 cassette and the avgC structural gene (Fig. 7F). These regions represent specific nucleotide sequences present only in that configuration. Three PCR primers spanning this region (Fig. 7C), designated pmar-1, pf-2, and pc-2, were used. A single 660-bp PCR product and a single 385-bp PCR product were obtained by using primers pmar-1 and pc-2 and primers pf-2 and pc-2, respectively (Fig. 7F). Sequence analysis of the two PCR amplicons confirmed the presence of the mar gene upstream of the avgC gene and the location of the F2 cassette upstream of the avgC gene (Fig. 7C). To verify the configuration shown in Fig. 7D, two primers (pmar-1 and pf-1) were used to amplify a region between the
mar and avgF genes. Notably, the distance between the mar and avgF genes in isolate 627#3 is 0.9 kb (Fig. 7A), whereas in the intermediate configuration (Fig. 7D) the distance is 2.3 kb due to the presence of the avgB gene between the mar and avgF genes. A 2.3-kb amplicon was obtained by using these primers, and sequence analysis confirmed the existence of the specific genomic junction present only in that configuration (Fig. 7G, lane 2).

**DISCUSSION**

In a previous study, it was shown that in a series of *M. agalactiae* clinical isolates obtained from a chronically infected animal, a chromosomal region containing a cluster of related avg genes underwent in vivo rearrangement (10). In the present study, we obtained evidence that the genomic rearrangements represent site-specific DNA inversions that occur in vivo within the avg locus and modulate avg gene expression. Phase-variable expression of the avg genes appears to involve at least three elements. The first element, a sequence-specific site (vis-like site), is necessary for site-specific DNA inversion. This site was identified in a 35-bp sequence present 37 bp upstream of each avg gene and resembles the recombination site (vis) identified upstream of each vsp gene in *M. bovis* (Fig. 1) (20). Since the clustered avg genes are not all oriented similarly, any two oppositely oriented vis-like copies could potentially serve as a recognition site for site-specific DNA inversions. We have shown that in the vsp inversion system of *M. bovis* even a short distance (550 bp) between two adjacent oppositely oriented vis sites is sufficient for a site-specific DNA inversion to occur (20). Thus, the presence of multiple vis-like sites within the *M. agalactiae* avg locus allows a wide spectrum of DNA inversions to occur. It should be noted that in numerous examples of bacterial site-specific DNA inversion systems that regulate phase variation of surface antigens, only two recombination sites and only one or two genes are involved (18). In contrast, the inversion systems of *M. bovis* and *M. agalactiae* are more complex and utilize multiple recombination sites to control expression of a large gene family (20, 31).

The second element is a promoter region. Only one avg upstream region (avg-B2), present as a single chromosomal copy, is recognized by RNA polymerase to initiate avg transcription. As a result, only one avg gene is transcribed, while the other avg genes are transcriptionally silent. Similarly, a single promoter was identified in the vsp system of *M. bovis*. However, despite the high level of nucleotide sequence similarity between the two active promoters (72%), they differ in their transcriptional start sites (Fig. 1).

The third element is a site-specific DNA recombinase. DNA site-specific recombinases are found in the chromosomes of many bacteria, where they contribute to a variety of biological processes, including the segregation of chromosomes and plasmids, the integration and excision of lysogenic bacteriophage genomes in the host, the transposition of transposable elements, and the lateral transfer of genetic material (11, 16, 18). In many bacteria, site-specific recombinases catalyze the site-specific DNA inversions that act as efficient genetic switches to allow genes encoding surface antigens to undergo high-frequency on-off switching, thereby generating antigenic variation (1, 4, 6, 7, 11, 18, 23). Two families of site-specific recombinases have been described, the resolvase invertase family and the λ integrase family (8, 11, 38). Recently, Ron et al. described in three pathogenic mycoplasmas, *M. agalactiae*, *M. bovis*, and *Mycoplasma pulmonis*, an open reading frame predicted to encode a site-specific recombinase that belongs to the λ integrase family of tyrosine recombinases (31). This gene is present as a single chromosomal copy and is adjacent to three distinct variable loci, the avg gene family in *M. agalactiae*, the vsp gene family in *M. bovis*, and the vsa gene family in *M. pulmonis* (31). More recently, Sitaraman et al. showed that the *M. pulmonis* recombinase, HvsR, alone is sufficient to catalyze site-specific DNA inversions of vrs-bearing fragments in *E. coli* (37). These findings appear to indicate that site-specific DNA inversions within the vsa locus occur in mycoplasmas without involvement of additional accessory proteins. It is, therefore, very likely that the recombinases Mar and Mbr, encoded by genes adjacent to the avg locus of *M. agalactiae* and the vsp locus of *M. bovis*, respectively, alone catalyze the high-frequency site-specific DNA inversions observed within these loci (20). It should be noted, however, that in several instances, site-specific DNA inversions involve auxiliary cellular factors that function as architectural components of the recombination complex to ensure proper recombination between correctly oriented sites (16, 17). The best-characterized examples of such systems are those of type 1 fimbriae from *E. coli* (1) and the flagellum of *Salmonella enterica* serovar Typhimurium (43). Global regulatory factors such H-NS (histone-like proteins), IHF (integration host factor), and Lrp (leucine response protein) play an important role in stimulating inversion events (17). For example, it has been suggested that IHF acts in concert with Lrp by bending the DNA to form two loops, thereby aligning the recombination sites so that strand exchange can occur (5, 17, 30). In avg and vsp inversion systems, site-specific DNA inversion of an approximately 5.3-kb genomic fragment (Fig. 7) and site-specific DNA inversion of an 10-kb genomic fragment (20), respectively, were identified. It seems likely that to bring the two corresponding vis sites into a position that allows strand exchange by the tyrosine recombinase to occur, an accessory factor is needed. Although such an accessory factor has not yet been identified in the mycoplasmas, this possibility cannot be excluded. The use of a member of the λ integrase family of tyrosine recombinases in three *Mycoplasma* species to achieve high-frequency antigenic variation of surface lipoproteins by juxtaposition of a regulatory element with a silent lipoprotein gene suggests that the DNA inversion systems were acquired from a common evolutionary ancestor.

Just before submission of this paper, another group of researchers reported that site-specific DNA inversions within the vmpa locus of *M. agalactiae* mediate surface diversity (14). Comparison of the nucleotide sequences of the vmpa genes showed that the locus described by these workers is, in fact, the avg locus analyzed in our study. It should be noted however, that the two groups of researchers analyzed different strains. Glew et al. analyzed the vmpa locus from the PG2 type strain, while in this study we characterized the avg system from clinical isolates from an infected animal and further demonstrated that the site-specific DNA inversions occurred in vivo in the natural animal host. In addition, identification of the active promoter within the vmpa locus was based on homology with its coun-
terpart from the vsp locus of M. bovis, implying that the transcriptional start site in the two species are identical. We mapped the transcriptional start site and showed that it differs from that in M. bovis. In the vsp promoter, it was identified 192 bp upstream of the Vsp initiation codon (20), whereas in the avg promoter (in isolate 627#3) it was mapped 205 bp upstream of the Avg initiation codon (Fig. 5).

Another important aspect that emerged from our study was the finding, obtained by PCR analysis, that each of the avg genes is sometimes associated with the B2 promoter, indicating that in propagating populations of M. agalactiae a wide spectrum of Avg phenotypes is maintained (Fig. 6). Some DNA inversions within the avg locus, such as the one that generates the configuration shown in Fig. 7B, occur between two silent genes (avgA and avgE) and do not change avg gene expression. Activation of a silent avg gene is possible when the silent avg gene is in its inverted orientation with respect to the B2 promoter. Such a configuration permits the silent gene to receive the B2 promoter via a single inversion event. Since the avg genes are located in the avg locus in different orientations, it is not possible to express each of the avg genes by a single inversion. Thus, phenotypically silent inversions within the avg locus play an important role and contribute significantly to generation of antigenic variation on the mycoplasma cell surface.

Collectively, the two studies provide compelling independent evidence demonstrating that M. agalactiae possesses a highly variable genomic locus that undergoes high-frequency site-specific DNA inversion that generates surface antigenic variation.

ACKNOWLEDGMENTS

This study was supported in part by research grant IS-2540-95R from BARD, The United States-Israel Binational Agricultural Research and Development Fund, and by grant 2215-1-99 from the Israel Ministry of Science to D.Y. and S.L. in collaboration with Abed Athmama, Research and Development Centre Hamesholash. We gratefully acknowledge the collaboration of the late Eitan Rapoport and his invaluable contribution to the clinical studies on which this research is based. R.F.-T. and S.M.-O. contributed equally to this work.

REFERENCES