Antigenic Variation of *Anaplasma marginale* by Expression of MSP2 Mosaics

ANTHONY F. BARBET,^{1*} ANNA LUNDGREN,¹ JOOYOUNG YI,¹
FRED R. RURANGIRWA,² AND GUY H. PALMER²

*Department of Pathobiology, College of Veterinary Medicine, University of Florida,
Gainesville, Florida 32611-0880,¹ and Program in Vector-Borne Diseases,
Department of Veterinary Microbiology and Pathology, Washington
State University, Pullman, Washington 99164-7030²*

Received 25 May 2000/Returned for modification 28 June 2000/Accepted 9 August 2000

Anaplasma marginale is a tick-borne pathogen, one of several closely related ehrlichial organisms that cause disease in animals and humans. These *Ehrlichia* species have complex life cycles that require, in addition to replication and development within the tick vector, evasion of the immune system in order to persist in the mammalian reservoir host. This complexity requires efficient use of the small ehrlichial genome. *A. marginale* and related ehrlichiae express immunoprotective, variable outer membrane proteins that have similar structures and are encoded by polymorphic multigene families. We show here that the major outer membrane protein of *A. marginale*, MSP2, is encoded on a polycistronic mRNA. The genomic expression site for this mRNA is polymorphic and encodes numerous amino acid sequence variants in bloodstream populations of *A. marginale*. A potential mechanism for persistence is segmental gene conversion of the expression site to link hypervariable *msp2* sequences to the promoter and polycistron.

Ehrlichiae are major causes of tick-borne diseases, including the recently emergent human monocytic and two granulocytic ehrlichioses and the most prevalent tick-borne infection in cattle worldwide, anaplasmosis (6). These pathogens, classified in genogroups I and II of the tribe *Ehrlichieae*, have a complex life cycle characterized by acute and persistent infection in the mammalian host and several replicative and developmental stages within the tick vector (10, 17). Notably, this complexity is achieved using a small genome, only 0.8 to 1.5 Mb in a single chromosome (1, 25). Persistence of *Anaplasma marginale* in cattle, which is fundamental for continued transmission, reflects sequential expression of antigenically variant outer membrane proteins that are encoded by the *msp2* multigene family (21). The outer membrane proteins of different ehrlichial organisms are significantly similar to one another in amino acid sequence, are all encoded by multigene families, and possess one to four variable regions (8, 13, 18–20, 22, 26, 29, 32). In *A. marginale*-infected cattle, three to six MSP2 variants are expressed in each sequential rickettsemic cycle, which recur every 4 to 8 weeks during persistent infection (8, 9, 15, 16). Thus, over the 7-year period in which *A. marginale* has been shown to persist, over 500 variants may be expressed. Although the cyclic emergence and immune control of *A. marginale* is similar to that occurring in African trypanosomiasis, the mechanisms used by the organism to persist in mammalian hosts are unknown. We show here that variation of *msp2* in erythrocyte stages of *A. marginale* proceeds through the formation of different sequence mosaics in a polycistronic expression site.

MATERIALS AND METHODS

Isolation and sequencing of *A. marginale* genomic DNA. Florida and South Idaho strains of *A. marginale* were maintained as liquid nitrogen-cryopreserved stabulates of infected bovine erythrocytes in dimethyl sulfoxide-phosphate-buff-

ered saline that were then used to infect cattle (20). *A. marginale* genomic DNA was isolated from highly rickettsemic (>50% rickettsemia) bovine blood by lysis with sodium dodecyl sulfate and lysozyme, treatment with proteinase K and RNase, phenol-chloroform extraction, and ethanol precipitation (5). In some cases, for PCR amplification and analysis of low rickettsemias, genomic DNA was isolated from 200 μ l of infected blood by using a QIAamp DNA mini kit (Qiagen, Valencia, Calif.). To obtain the sequence of the complete *msp2* coding region, the sequence flanking *orf2* and *msp2* was obtained in both 5' and 3' directions and then confirmed on both strands following PCR amplification of the entire locus (14, 28). Sequencing was performed at the University of Florida DNA Sequencing Core Laboratory (Gainesville, Fla.) using ABI Prism dye terminator cycle sequencing protocols developed by Applied Biosystems (Perkin-Elmer Corp., Foster City, Calif.). The fluorescently labeled extension products were analyzed on an Applied Biosystems model 373 Stretch DNA Sequencer (Perkin-Elmer Corp.). Oligonucleotide primers were designed using OLIGO 5.0 (Molecular Biology Insights, Cascade, Colo.) software and synthesized by Genosys Biotechnologies (The Woodlands, Tex.). Nucleotide sequences were analyzed using the GCG programs (Genetics Computer Group, University of Wisconsin) available through the Biological Computing core facilities of the Interdisciplinary Center for Biotechnology Research at the University of Florida. Sequence alignments were made using PILEUP and GAP, and similarities were displayed using PLOTSIMILARITY. Prokaryotic factor-independent RNA polymerase terminator sequences were predicted using TERMINATOR.

Analysis of *msp2* mRNA. Total RNA was isolated from whole blood obtained during acute rickettsemia with the FL strain of *A. marginale* by extraction with 6 M urea–3M LiCl (27). For reverse transcription-PCR (RT-PCR) analysis, RNA was digested with RQ1 DNase (Promega, Madison, Wis.) and the reactions were conducted with or without reverse transcriptase (Retroscript kit; Ambion, Austin, Tex.) in a total volume of 20 μ l. Then 1.5 μ l of the RT reaction mixture was used in primary PCRs and 2.5 μ l of the primary PCR mixture was used in secondary or nested PCRs. For primer extension analysis, the oligonucleotide primer was 5'-end labeled with ³²P using T4 polynucleotide kinase and extended in sequencing reactions with reverse transcriptase, using as template 50 μ g of RNA isolated from *A. marginale*-infected blood or 0.5 μ g of a 3.9-kb PCR product DNA containing *msp2*, *orf2* to *orf4*, and flanking regions (12). RNase protection assays were conducted with an antisense RNA probe prepared from *msp2* cDNA clone AR11 (7), which contains coding sequence for both *msp2* and *orf2* (see below). AR11 plasmid DNA was linearized with *Xho*I and labeled with [³²P]UTP using the Maxiscript kit (Ambion) and SP6 RNA polymerase, and the 722-nucleotide probe was purified by gel isolation. The labeled probe was hybridized with different amounts of *A. marginale* RNA or carrier yeast RNA and then subjected to digestion with RNase A–RNase T₁ as described by the manufacturer (RPAIII kit; Ambion). Protected probe fragments were analyzed by electrophoresis in an 8 M urea–5% acrylamide gel with ³²P-labeled RNA Century Marker Plus standards (Ambion).

Southern blotting of *A. marginale* genomic DNA. Probes for Southern blots were derived as follows: *msp2* probe, insert DNA from *msp2* cDNA clone AR3

* Corresponding author. Mailing address: Department of Pathobiology, College of Veterinary Medicine, University of Florida, P. O. Box 110880, Gainesville, FL 32611-0880. Phone: (352) 392-4700, ext. 5819. Fax: (352) 392-9704. E-mail: barbeta@mail.vetmed.ufl.edu.

(7) that includes nucleotides 18 to 555 of *msp2*; *orf2* probe, PCR amplification product from *msp2* cDNA clone AR9 (7) with oligonucleotide primers AB688 and AB689 that contains nucleotides 28 to 275 of *orf2*; *orf3* probe, PCR amplification product from *msp2* cDNA clone AR9 with primers AB690 and AB691 and with nucleotides 618 to 877 of *orf3*; *orf4* probe, PCR amplification product from Florida strain *A. marginale* genomic DNA with primers AB783 and AB747 that contains nucleotides 467 to 896 of *orf4*. DNA probes were labeled with fluorescein-dUTP, hybridized, washed under high-stringency conditions ($0.1 \times$ SSC [$1 \times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate] plus 0.1% sodium dodecyl sulfate at 60°C) and detected by chemiluminescence (Illuminator chemiluminescent detection system; Stratagene, La Jolla, Calif.). Oligonucleotide probes were synthesized with a 5' fluorescein end label (Genosys Biotechnologies), hybridized, washed, and detected similarly, using final wash temperatures of 5°C below their respective T_m values. Molecular size standards were Illuminator nonradioactive markers (Stratagene).

Determination of *msp2* expression site structure in progenitor and progeny populations of *A. marginale*. Bovine no. 196 was infected with *A. marginale* (Florida strain). At 40 days after the initial infection, the animal experienced a peak of acute rickettsemia of 6% (the percentage of erythrocytes containing *A. marginale*), which subsequently resolved, and the animal became an asymptomatic carrier of infection. At 107 days after initial infection, the animal was splenectomized, which caused recrudescence of a microscopically detectable rickettsemia. The animal was exsanguinated 19 days after splenectomy with a rickettsemia of 95% to obtain the Florida-relapse strain.

Synthetic oligonucleotides. The synthetic oligonucleotides used in this study were (5'-3') AB192 (CTATCCTTGAAGCTAATCTTG), AB198 (AAGGCAACCTAACACCCCAAC), AB688 (GGACTGCTTGCCTTCACGCTGTT), AB689 (TGAGCTGGGGAAAAGACGCTTGG), AB690 (CGGCGGCGTGGAGTTCCTTGA), AB691 (TGCCTGCTTCGACGCCAAGGT), AB747 (ATACAACCCGACCACAAAATCC), AB750 (GGATTTTGTGGTCGGGGTTTGTAT), AB752 (CACCGTTGATGAAGTTTGC), AB764 (GCGTTCGGCAGGCATTTTGG), AB765 (GGAACAACCCCAATACCATC), AB766 (GTATGTCGATTTCGCGGAAGAGCCTGTTGT), AB767 (ACGCGCTTGAATAATCGTT), AB783 (AGTATCACATTGGGGAGGTTT), AB784 (GGAGGGAAGCCGAAGTTG), AB847 (ACAGACACACTGCTTCTGTGTAGGGGAACAAAGAC), AB871 (GTACCACTGGCAGTGGTGCCATTGAT), and AB872 (ATTGTTGGTGTGCCACTGGTGCGGTTAAC).

Nucleotide sequence accession numbers. The sequences reported have been assigned GenBank accession numbers AF200925 to AF200927.

RESULTS

MSP2 is encoded on a polycistronic mRNA transcript. Examination of over 250 *msp2* transcripts expressed during acute or persistent *A. marginale* rickettsemia revealed a single, central hypervariable region that encodes the approximately 100-amino-acid antigenically variable domain (8, 9). The hypervariable *msp2* region, characterized by nucleotide insertions, deletions, and substitutions, is flanked by highly conserved 5' and 3' ends. This structure is shared by the MSP2 homolog (Blastp E value of 10^{-102}) in a causative agent of human granulocytic ehrlichiosis (13, 18, 32). Using cDNA clones derived from *A. marginale* during acute rickettsemia, we identified transcripts that extended 5' to the ATG initiation codon for MSP2 (7). In the longest cDNA clone, AR9, this 5' sequence extended 726 bp beyond the *msp2* ATG and contained two additional open reading frames (ORFs). The immediately 5'-proximal ORF (*orf2*) was separated from the *msp2* ATG by only 12 bp and was separated from *orf3* (Fig. 1) by 23 bp. This structure is typical of prokaryotic polycistronic mRNAs, where the ORFs are generally separated by <30 bp. To obtain the structure of a complete transcriptional unit, we sequenced genomic DNA (14, 28) and identified a single locus containing *msp2* linked to *orf2* and *orf3* and a fourth ORF (*orf4*) immediately 5' to *orf3* (Fig. 1). Numbering from *msp2*, *orf2* encodes a polypeptide with a molecular weight of 13,219 that is predicted by the PSORT (<http://psort.nibb.ac.jp>) algorithm to be an outer membrane protein. *orf2* did not resemble any sequence in the databases. *orf3* and *orf4* also encoded predicted outer membrane proteins with molecular weights of, respectively, 31,118 and 29,061; both were significantly similar (Blastp E values of 2×10^{-19} and 1×10^{-12} , respectively) to the outer membrane protein OMP1b of *Ehrlichia chaffeensis*, the causative agent of human monocytic ehrlichiosis (19). There were

FIG. 1. Structure and variability of the *msp2* polycistronic expression site. (A) Diagram of the four ORFs comprising the polycistronic expression site for *msp2*, showing the molecular weights (M.W.) of encoded proteins, primary (1°) and secondary (2°) RT-PCR products generated from *msp2* mRNA, location of probe used in the RNase protection assay (RPA probe, see Fig. 2), and *FspI* cleavage sites (see Fig. 4 and 5). P, predicted promoter region, 72 to 108 bp 5' to *orf4*; T, predicted prokaryotic terminator sequence, GTAGACCAGC...TAGT CGTCAC, 149 to 200 bp 3' to *msp2*. (B) PLOTSIMILARITY profile of nucleotide sequence variability between *msp2* expression sites of Florida (F) and South Idaho (I-1) strains of *A. marginale*. A similarity score of 1.0 indicates identical sequence in a sliding window of 10 nucleotides, and a decreasing score from 1.0 to 0.0 indicates increasing variation. (C) PLOTSIMILARITY profile of expression site variability in the South Idaho strain of *A. marginale* examined at two time points 10 days apart in acute rickettsemia (I-1 and I-2).

no other significant ORFs within 130 bp 5' or 240 bp 3' to these four ORFs in the genome. Thus, the structure in Fig. 1 appeared to represent the entire coding region for a polycistronic *msp2* transcript.

To confirm that the genomic region shown in Fig. 1 encoded a polycistronic transcript, RNA obtained from *A. marginale* during acute rickettsemia was analyzed by RT-PCR, RNase protection, and primer extension (Fig. 2). RNA encoding MSP2 and also containing *orf2* to *orf4* was present as a polycistronic transcript that could be amplified by RT-PCR. No amplified products were present in control reactions without RT. The sequenced RT-PCR product of 2 kbp (Fig. 2) contained the regions of *msp2* and *orf2* to *orf4* expected from the genomic locus. An RNase protection assay was performed to

FIG. 2. Analysis of *msp2* transcript structure by RT-PCR, primer extension analysis, and RNase protection assay (RPA). For RT-PCR, total DNase-treated RNA of Florida strain *A. marginale* was reverse transcribed into DNA using oligonucleotide primer AB198, which anneals to the 3' end of *msp2*. The cDNA was amplified in a primary PCR with primers AB765 and AB766, which generated a product of 3.2 kbp. The primary PCR product was amplified in a secondary or nested PCR with primer combinations AB192 and AB764, AB689 and AB764, and AB192 and AB688 to generate products of 2.0, 1.5, and 0.7 kbp, respectively. See Fig. 1 for the locations of RT-PCR products. S, molecular size standards; +, with reverse transcriptase; -, negative control reactions without reverse transcriptase. For primer extension, oligonucleotide primer AB784, which anneals 153 nucleotides 3' to the ATG initiation codon of *orf4*, was radiolabeled with ^{32}P and extended in sequencing reactions using reverse transcriptase and either total RNA of *A. marginale* or denatured, PCR-amplified genomic DNA containing *orf2* to *orf4*, *msp2*, and flanking regions as templates. The order of sequencing reactions T, G, C, and A, is shown above the DNA lanes and was the same in the RNA lanes. A strong stop was detected in RNA at the A (underlined) in sequence TGCAACCCACACACCCATAAGG, with evidence also for a minority of transcripts continuing to the next base, C (italic). This corresponds to the coding-strand sequence CCTTATGGGTGTGTGGGTGCA (see the text). For the RNase protection assay, a ^{32}P -labeled antisense RNA probe of 722 nucleotides (317 nucleotides of the *msp2* gene, 307 nucleotides of *orf2* and intergenic spacer, 98 nucleotides of plasmid vector) was allowed to hybridize to various amounts (10, 3, and 0.1 μg) of total RNA of *A. marginale* and carrier yeast RNA or to yeast RNA alone (lane 0) and then unprotected single-stranded probe was digested with RNase and analyzed by denaturing polyacrylamide gel electrophoresis. C, probe plus yeast RNA, not digested with RNase. The positions of molecular size standards are shown on the left. A band of 624 bp containing the *A. marginale* sequences within the probe (*msp2* and *orf2*) was the predominant fragment protected.

investigate whether *msp2* was transcribed primarily as a polycistronic mRNA. An antisense RNA probe of 624 bases (317 bases at the 5' end of *msp2* and 307 bases of the intergenic sequence and *orf2* [Fig. 1]) was protected from RNase digestion by hybridization with total *A. marginale* RNA (Fig. 2), suggesting that the majority of *msp2* transcripts were polycistronic and linked to *orf2* when expressed during acute rickettsiaemia. The transcription start site (+1) was identified by primer extension analysis (Fig. 2) as 74bp 5' to the ATG initiation codon of *orf4* within the sequence:

-35
ATTTAAATTGCAAAATATATTGCCAGCTCTCTTGACT
-10 +1
TAATTCTTTTATTAGTATACCTTATGGGTGTGTGGG

This sequence contained elements that may be -35 and -10 promoter regions (underlined). The consensus *Escherichia coli* -35 region is TTGACA and -10 is TATAAT, with an optimal spacing between them of 17 ± 1 bp (11). The three of six identities at -10 were the most highly conserved first, second, and sixth positions of the *E. coli* consensus sequence. There was also an AT-rich region upstream of -35 in the *A. marginale* sequence typical of prokaryotic promoters (11). Interestingly, a second region was found (double underline) with a similar level of sequence identity to the consensus *E. coli* -35 and -10 sequences, separated from the first by only 6 bp. The significance of this potential dual-promoter structure in transcriptional control is unknown.

This *msp2* expression site encodes hypervariable sequences expressed in vivo. To analyze the variability of MSP2 encoded by the polycistronic expression site, we sequenced the expression site in two genetically and biologically distinct *A. marginale* strains obtained during acute rickettsiaemia (Fig. 1B) and at two different time points in a single acute infection (Fig. 1C). Alignment of complete expression site sequences revealed minor substitutions in *orf2* and *orf3* and no changes in the predicted promoter region (*orf2* gene product, 98% amino acid identity between the Florida [F] and South Idaho [I-1] strains; *orf3* gene product, 99% identity). There were more substitutions in *orf4*, and these were distributed throughout the *orf4* sequence (96% identity between the F and I-1 strains). The greatest differences were in *msp2* and included substitutions, insertions, and deletions (92% identity between F and I-1). Essentially similar results were obtained comparing expression site structure at two sequential time points in a single acute infection (Fig. 1C). In *msp2*, most of the expression site variability was localized to the region encoding the known MSP2 hypervariable region (8, 9) (Fig. 1B and C). We sequenced this hypervariable region in multiple, independent DNA clones of the genomic expression site and compared the encoded MSP2 peptide sequences (Fig. 3). Each population of *A. marginale*, whether taken from animals with acute rickettsiaemia with the Florida or South Idaho strains or from different time points in a single infection, was extremely polymorphic in the part of the expression site encoding the MSP2 hypervariable region.

Do the expressed *msp2* RNA transcripts reflect the sequence of this genomic locus? The major variant type in *msp2* cDNA

FIG. 4. Structure of *msp2* and *orf2* to *orf4* in genomic DNA of Florida and South Idaho strains of *A. marginale*. Southern blots of Florida (F) or South Idaho (I) genomic DNA digested with the restriction enzyme *FspI* and hybridized with probes specific for *msp2*, *orf2*, *orf3*, or *orf4* are shown. *FspI* cleaves 41 nucleotides 5' to *orf4* and 268 nucleotides 3' to *msp2* to release a fragment of 3.76 kbp containing the complete polycistronic *msp2* expression site sequence (see Fig. 1) from both Florida and South Idaho genomic DNAs. Molecular size standards are shown in the left lane of each blot. Multiple *msp2*-related sequences are detected in genomic DNA of both strains; only *msp2* sequences located in the expression site are contiguous with *orf2* to *orf4*.

ny) derived after >3 months of persistent bovine infection. Oligonucleotide probes to *msp2* hypervariable region sequences present only in the expression site of the progeny strain detect an extra band in the progeny population. Probes to the hypervariable region sequence present in the expression site in both the progenitor and progeny strains detect an expression site band in both populations, as expected (Fig. 5). Non-expression-site-associated bands remain unchanged in both the progenitor and progeny strains; these may be templates for segmental gene conversion of the expression site.

DISCUSSION

A mechanism involving recombination into a polycistronic expression site, especially one containing other outer membrane protein genes, is unusual and represents efficient use of a small genome to generate diversity. Since *orf3* and *orf4* are distantly related members of the *msp2* gene family, they, too, could vary by recombination with partially homologous sequences elsewhere in the genome. Comparison of *orf3* and *orf4* in different *A. marginale* populations shows that variation in these genes occurs at a lower rate than in *msp2*. The large number of diverse sequences found in the expression site of a single *A. marginale* population at a single time is consistent with the extensive variation of *msp2* during the multiple ricketsemia cycles that comprise a single infection.

Recombination of portions of silent genes into a monocistronic expression site has been described for other prokaryotes, including the tick-borne *Borrelia burgdorferi* (30). However, *B. burgdorferi* uses plasmid-encoded genes to generate this diversity, an option unavailable to *A. marginale* (1). In some respects, *A. marginale msp2* variation resembles the for-

mation of mosaic genes in the polycistronic expression sites encoding the variable surface glycoprotein (VSG) of African trypanosomes. Complete genes and otherwise inactive pseudogenes are used to generate these mosaics in chronic trypanosomiasis (4, 23). The sequence mosaics encompass most regions of the VSG and are not targeted to a single hypervariable region as in *A. marginale msp2*. The trypanosome, because of its larger genome size, also has sufficient VSG-encoding genes to produce considerable variation without needing combinatorial mechanisms.

Rickettsiae, including the ehrlichial pathogens, are thought to be closely related to the original bacterial endosymbionts of eukaryotic cells that evolved into mitochondria by loss of genes

FIG. 5. Comparison of genes encoding the *msp2* hypervariable region in Florida (F) and Florida relapse (F-rel) strains of *A. marginale*. Hypervariable-region sequences were obtained from independent clones of the *msp2* expression site (ES) in F (10 clones) and F-rel (11 clones) genomic DNA. From this DNA sequence information, different oligonucleotide probes to the hypervariable region were synthesized to use in Southern blotting. These probes contained sequence either unique to one set of expression site clones (the AB871 and AB872 sequences were both observed in 2 of 11 clones of the F-rel expression site and 0 of 10 clones of the F expression site) or were present in different proportions (the AB847 sequence was present in 5 of 10 expression site clones in the F strain and 1 of 11 expression site clones in the F-rel strain). *A. marginale* genomic DNA of the F or F-rel strains was digested with *FspI* to release the 3.76-kbp expression site (arrow), separated by gel electrophoresis, and probed with the indicated fluorescein-labeled synthetic oligonucleotides. With probe AB872, an extra (expression site) band of 3.76 kbp is detected in F-rel compared to F genomic DNA probed with the same sequence. A 4.4-kbp band is detected in both F and F-rel DNA, which may contain a template copy for conversion of the expression site. The expression site band is less intense than the 4.4-kbp band because the AB872 sequence is present in the expression site of only some organisms in the F-rel population of *A. marginale* (approximately 18% [2 of 11] of the total would be predicted from the fraction of F-rel expression site clones containing the AB872 sequence). This figure is approximate because of the small number of clones analyzed and the possibility of representation bias following amplification and cloning of the expression site. A similar result was obtained with a different F-rel-specific probe, AB871, except that there are two potential template copies of the AB871 sequence on 4.4- and 3.5-kbp fragments. The AB847 blot is an example of the converse situation: the AB847 sequence is well represented in the expression site of the F population of *A. marginale* but has nearly disappeared from the expression site in F-rel *A. marginale*. Potential template copies for the AB847 expression site sequence are on 11- and 7.2-kbp fragments.

whose function has been replaced by the host genome (3). *Rickettsia prowazekii* contains a high proportion of noncoding DNA (24%), despite a size of only 1.1 Mb. Rickettsial genes are postulated to be in the process of elimination from the genome by a series of steps comprising mutation to pseudogene, to unrecognizable sequence, to small fragments, to extinction (3). The data reported herein suggest that some pseudogene sequences may not be on their way to extinction but play a critical role in adaptive mechanisms of small-genome pathogens.

In conclusion, the MSP2 outer membrane protein is encoded on a polycistronic RNA transcript in erythrocyte stages of *A. marginale*. In addition to *msp2*, three other genes are present on this transcript that also are predicted to encode outer membrane proteins. All but one of these genes are significantly similar to outer membrane protein genes present in other ehrlichial pathogens, including those that infect humans. There is extensive polymorphism in the genomic expression site for this *msp2* polycistronic transcript and evidence for gene rearrangement during persistent infection. Availability of the expression site sequence should allow the determination of the MSP2 variants expressed from this site during cyclical transmission of *A. marginale* between ticks and cattle. These data provide a possible molecular basis for persistence of *A. marginale* and may permit identification of similar mechanisms of variation in other ehrlichial pathogens. Common mechanisms of variable surface protein expression may be used by diverse pathogens to evade the host immune response and cause persistent infections.

ACKNOWLEDGMENTS

We thank Carlos R. Sulsona and Carla Robertson for excellent technical assistance.

This investigation was supported by USDA grant 95-372204-2348 and NIH grants AI45580 and AI44005.

REFERENCES

- Alleman, A. R., S. M. Kamper, N. Visesakul, and A. F. Barbet. 1993. Analysis of the *Anaplasma marginale* genome by pulsed-field electrophoresis. *J. Gen. Microbiol.* **139**:2439-2444.
- Alleman, A. R., G. H. Palmer, T. C. McGuire, T. F. McElwain, L. E. Perryman, and A. F. Barbet. 1997. *Anaplasma marginale* major surface protein 3 is encoded by a polymorphic, multigene family. *Infect. Immun.* **65**:156-163.
- Andersson, S. G., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. Alsmark, R. M. Podowski, A. K. Naslund, A. S. Eriksson, H. H. Winkler, and C. G. Kurland. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**:133-140.
- Barbet, A. F., and S. M. Kamper. 1993. The importance of mosaic genes to trypanosome survival. *Parasitol. Today* **9**:63-66.
- Barbet, A. F., G. H. Palmer, P. J. Myler, and T. C. McGuire. 1987. Characterization of an immunoprotective protein complex of *Anaplasma marginale* by cloning and expression of the gene coding for polypeptide Am105L. *Infect. Immun.* **55**:2428-2435.
- Dumler, J. S., and J. S. Bakken. 1998. Human ehrlichioses: newly recognized infections transmitted by ticks. *Annu. Rev. Med.* **49**:201-213.
- Eid, G., D. M. French, A. M. Lundgren, A. F. Barbet, T. F. McElwain, and G. H. Palmer. 1996. Expression of major surface protein 2 antigenic variants during acute *Anaplasma marginale* rickettsemia. *Infect. Immun.* **64**:836-841.
- French, D. M., W. C. Brown, and G. H. Palmer. 1999. Emergence of *Anaplasma marginale* antigenic variants during persistent rickettsemia. *Infect. Immun.* **67**:5834-5840.
- French, D. M., T. F. McElwain, T. C. McGuire, and G. H. Palmer. 1998. Expression of *Anaplasma marginale* major surface protein 2 variants during persistent cyclic rickettsemia. *Infect. Immun.* **66**:1200-1207. (Erratum, **66**:2400.)
- Ge, N. L., K. M. Kocan, E. F. Blouin, and G. L. Murphy. 1996. Developmental studies of *Anaplasma marginale* (Rickettsiales: Anaplasmataceae) in male *Dermacentor andersoni* (Acari: Ixodidae) infected as adults by using nonradioactive in situ hybridization and microscopy. *J. Med. Entomol.* **33**:911-920.
- Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**:2237-2255.
- Hollingshead, S. K., V. A. Fischetti, and J. R. Scott. 1987. Size variation in group A streptococcal M protein is generated by homologous recombination between intragenic repeats. *Mol. Gen. Genet.* **207**:196-203.
- Ijdo, J. W., W. Sun, Y. Zhang, L. A. Magnarelli, and E. Fikrig. 1998. Cloning of the gene encoding the 44-kilodalton antigen of the agent of human granulocytic ehrlichiosis and characterization of the humoral response. *Infect. Immun.* **66**:3264-3269.
- Johnston, S. L., M. Strausbauch, G. Sarkar, and P. J. Wettstein. 1995. A novel method for sequencing members of multi-gene families. *Nucleic Acids Res.* **23**:3074-3075.
- Kieser, S. T., I. S. Eriks, and G. H. Palmer. 1990. Cyclic rickettsemia during persistent *Anaplasma marginale* infection of cattle. *Infect. Immun.* **58**:1117-1119.
- Knowles, D., D. E. Torioni, G. Palmer, T. McGuire, D. Stiller, and T. McElwain. 1996. Antibody against an *Anaplasma marginale* MSP5 epitope common to tick and erythrocyte stages identifies persistently infected cattle. *J. Clin. Microbiol.* **34**:2225-2230.
- Kocan, K. M., D. Stiller, W. L. Goff, P. L. Claypool, W. Edwards, S. A. Ewing, T. C. McGuire, J. A. Hair, and S. J. Barron. 1992. Development of *Anaplasma marginale* in male *Dermacentor andersoni* transferred from parasitic to susceptible cattle. *Am. J. Vet. Res.* **53**:499-507.
- Murphy, C. L., J. R. Storey, J. Recchia, L. A. Doros-Richert, C. Gingrich-Baker, K. Munroe, J. S. Bakken, R. T. Coughlin, and G. A. Beltz. 1998. Major antigenic proteins of the agent of human granulocytic ehrlichiosis are encoded by members of a multigene family. *Infect. Immun.* **66**:3711-3718.
- Ohashi, N., N. Zhi, Y. Zhang, and Y. Rikihisa. 1998. Immunodominant major outer membrane proteins of *Ehrlichia chaffeensis* are encoded by a polymorphic multigene family. *Infect. Immun.* **66**:132-139.
- Palmer, G. H., G. Eid, A. F. Barbet, T. C. McGuire, and T. F. McElwain. 1994. The immunoprotective *Anaplasma marginale* major surface protein 2 is encoded by a polymorphic multigene family. *Infect. Immun.* **62**:3808-3816.
- Palmer, G. H., F. R. Rurangirwa, K. M. Kocan, and W. C. Brown. 1999. Molecular basis for vaccine development against the ehrlichial pathogen *Anaplasma marginale*. *Parasitol. Today* **15**:281-286.
- Reddy, G. R., C. R. Sulsona, A. F. Barbet, S. M. Mahan, M. J. Burridge, and A. R. Alleman. 1998. Molecular characterization of a 28 kDa surface antigen gene family of the tribe *Ehrlichiae*. *Biochem. Biophys. Res. Commun.* **247**:636-643.
- Roth, C., F. Bringaud, R. Layden, T. Baltz, and H. Eisen. 1989. Active late-appearing variable surface antigen genes in *Trypanosoma equiperdum* are constructed entirely from pseudogenes. *Proc. Natl. Acad. Sci. USA* **86**:9375-9379.
- Rurangirwa, F. R., D. Stiller, D. M. French, and G. H. Palmer. 1999. Restriction of major surface protein 2 (MSP2) variants during tick transmission of the ehrlichia *Anaplasma marginale*. *Proc. Natl. Acad. Sci. USA* **96**:3171-3176.
- Rydkina, E., V. Roux, and D. Raoult. 1999. Determination of the genome size of *Ehrlichia* spp., using pulsed field gel electrophoresis. *FEMS Microbiol. Lett.* **176**:73-78.
- Sulsona, C. R., S. M. Mahan, and A. F. Barbet. 1999. The map1 gene of *Cowdria ruminantium* is a member of a multigene family containing both conserved and variable genes. *Biochem. Biophys. Res. Commun.* **257**:300-305.
- Van Der Ploeg, L. H. T., A. Y. C. Liu, P. A. M. Michels, T. De Lange, P. Borst, H. K. Majumder, H. Weber, H. G. Veeneman, and J. H. Van Boom. 1982. RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes. *Nucleic Acids Res.* **10**:3591-3604.
- Weber, K. L., M. E. Bolander, and G. Sarkar. 1998. Rapid acquisition of unknown DNA sequence adjacent to a known segment by multiplex restriction site PCR. *BioTechniques* **25**:415-419.
- Yu, X., J. W. McBride, X. Zhang, and D. H. Walker. 2000. Characterization of the complete transcriptionally active *Ehrlichia chaffeensis* 28 kDa outer membrane protein multigene family. *Gene* **248**:59-68.
- Zhang, J., J. M. Hardham, A. G. Barbour, and S. J. Norris. 1997. Antigenic variation in Lyme disease *Borreliae* by promiscuous recombination of VMP-like sequence cassettes. *Cell* **89**:275-285.
- Zhi, N., N. Ohashi, and Y. Rikihisa. 1999. Multiple p44 genes encoding major outer membrane proteins are expressed in the human granulocytic ehrlichiosis agent. *J. Biol. Chem.* **274**:17828-17836.
- Zhi, N., N. Ohashi, Y. Rikihisa, H. W. Horowitz, G. P. Wormser, and K. Hechemy. 1998. Cloning and expression of the 44-kilodalton major outer membrane protein gene of the human granulocytic ehrlichiosis agent and application of the recombinant protein to serodiagnosis. *J. Clin. Microbiol.* **36**:1666-1673.