

Simultaneous Variation of the Immunodominant Outer Membrane Proteins, MSP2 and MSP3, during *Anaplasma marginale* Persistence In Vivo

Kelly A. Brayton,^{1*} Patrick F. M. Meeus,² Anthony F. Barbet,² and Guy H. Palmer¹

Program in Vector Borne Diseases, Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, Washington 99164-7040,¹ and Department of Pathobiology, College of Veterinary Medicine, University of Florida, Gainesville, Florida 32611-0880²

Received 28 March 2003/Returned for modification 18 June 2003/Accepted 11 August 2003

Vector-borne bacterial pathogens persist in the mammalian host by varying surface antigens to evade the existing immune response. To test whether the model of surface coat switching and immune evasion can be extended to a vector-borne bacterial pathogen with multiple immunodominant surface proteins, we examined *Anaplasma marginale*, a rickettsia with two highly immunogenic outer membrane proteins, major surface protein 2 (MSP2) and MSP3. The simultaneous clearance of variants of the two most immunodominant surface proteins of *A. marginale* followed by emergence of unique variants indicates that the switch rates and immune selection for MSP2 and MSP3 are sufficiently similar to explain the cyclic bacteremia observed during infection in the immunocompetent host.

Transmission of vector-borne animal pathogens relies on the ability of the pathogen to persist in the immunocompetent host. This persistence is achieved by varying the immunodominant surface protein(s) which allows for evasion of the variant-specific immune response (reviewed in references 4, 5, 9, 20, and 24). In perhaps the best understood example, the African trypanosome, there is a single surface coat protein, the variable surface glycoprotein (VSG) (4). For most vector-borne pathogens, such as *Plasmodium falciparum*, and bacteria in the genera *Anaplasma*, *Borrelia*, and *Ehrlichia*, there are multiple immunodominant outer membrane proteins (1, 13, 15, 17, 21, 23, 25). In the case of *P. falciparum*, these genes (e.g., *var* and *rif*) are expressed on the surface of the host erythrocyte (13, 23); however, most bacterial pathogens express the immunodominant proteins on the bacterial surface (1, 15, 17, 21, 25). This key difference between organisms with a single immunodominant surface protein and those with multiple immunodominant surface proteins represents a major gap in our knowledge regarding mechanisms of persistence for vector-borne bacterial pathogens—does simultaneous variation in multiple immunodominant outer membrane proteins occur during infection of an immunocompetent host?

To test whether the trypanosome model of surface coat switching and immune evasion can be extended to a vector-borne bacterial pathogen with multiple immunodominant surface proteins, we examined *Anaplasma marginale*, a rickettsia with two highly immunogenic outer membrane proteins, major surface protein 2 (MSP2) and MSP3. *A. marginale* is a tick-transmitted pathogen of domestic and wild ruminants that causes severe anemia during acute infection (14). Phenotypically similar to African trypanosomiasis, *A. marginale* infection is characterized by recurring cycles of bacteremia in which

organisms replicate to a peak of $\geq 10^6$ /ml and are then controlled by a variant-specific immune response (10–12). MSP2, an ~44-kDa outer membrane protein, undergoes variation in the central hypervariable region (HVR) such that antibody present at the first bacteremic peak will not recognize new emergent variants at subsequent peaks (11); thus, MSP2 is a key component in evasion of the host immune response (reviewed in reference 19). Importantly, the predominant antibody response to *A. marginale* is directed not only at MSP2 but also at a second highly variable protein of 65 to 80 kDa, MSP3 (1). These two immunodominant surface proteins are encoded by separate multigene families, each composed of seven to nine pseudogenes and a single expression site (2, 6, 16, 18). These two gene families belong to a superfamily of related genes, but the total sequence identity between MSP2 and MSP3 is low, on average ~35%, with the exact percent identity dependent on which variants are used for comparison. Although the *msp2* and *msp3* gene families share a common mechanism to generate variation in the MSP2 and MSP3 proteins, employing gene conversion using all or part of the central HVRs of donor pseudogenes, the two expression loci are physically separated by more than 60 kb in the genome. We hypothesized, based on the trypanosome model of antigenic variation, that the frequency of recombination and host immune selection for these two independent surface protein gene families was sufficient to allow the emergence of MSP2 and MSP3 escape variants in an immunocompetent animal. This hypothesis was tested by quantitative analysis of expressed variants in two sequential bacteremic peaks in vivo.

The South Idaho strain of *A. marginale* undergoes a restriction during replication and development within the tick, such that only two MSP2 variants are subsequently transmitted to the next calf (22), creating a semiclinal starting point for analysis of new variation. Laboratory-reared adult male ticks of the Reynolds Creek stock of *Dermacentor andersoni* were allowed to feed on calf 855, which was persistently infected with the South Idaho strain of *A. marginale* (7), and this infection

* Corresponding author. Mailing address: Program in Vector Borne Diseases, Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA 99164-7040. Phone: (509) 335-6340. Fax: (509) 335-8529. E-mail: kbrayton@vetmed.wsu.edu.

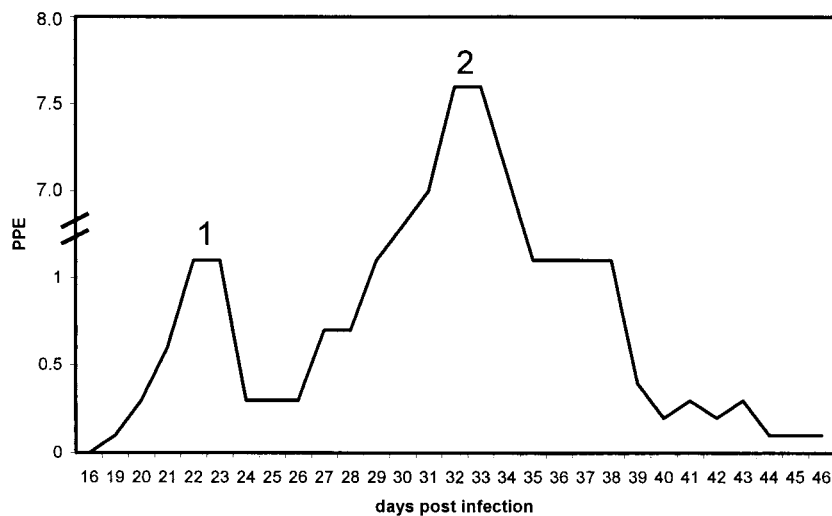


FIG. 1. Bacteremic peaks 1 and 2. The plot illustrates two peaks of *A. marginale* bacteremia with clearance between the two peaks. Percent parasitized erythrocytes (PPE) is shown on the y axis.

was subsequently transmitted to calf 894 by allowing infected *D. andersoni* ticks to feed on this calf for a period of 7 days. Blood samples were collected daily after tick exposure, and the *A. marginale* infection in calf 894 was quantitated as the percentage of infected erythrocytes (Fig. 1). Infected blood was washed three times in phosphate-buffered saline with removal of the buffy coat at each wash, and the PureGene kit (Gentra Systems) was used to extract genomic DNA from infected erythrocytes. DNA from bacteremic peak 1 (day 21 postinfection) and bacteremic peak 2 (day 28 postinfection) was used to examine changes in the *msp2* and *msp3* expression sites.

We identified expressed MSP2 and MSP3 variants during these two successive bacteremic peaks by amplifying the expression site genes using PCR Master (Roche), cloning them into the pCR-4-TA-TOPO cloning system (Invitrogen), and sequencing them. The primers used to amplify the *msp2* expression site have been described previously (6) (*msp2* forward primer [5' TCC TAC CAA GCG TCT TTT CCC C 3'] and *msp2* reverse primer [5' TTA CCA CCG ATA CCA GCA CAA 3']), and the positions of the primers are shown in Fig. 2. It is important to note that these primers will not detect *msp2* pseudogenes, and all sequences in the resulting PCR fragments will correspond to an expressed *msp2* gene. PCR conditions were as follows: (i) denaturation for 5 min at 95°C; (ii) 30 cycles, with 1 cycle consisting of 15 s at 94°C, 30 s at 55°C, and 30 s at 72°C. PCR was done using PCR master mix (Roche). The *msp3* expression site was obtained by seminested amplification, as the cloning efficiency of the large primary amplification product (~4 kb) was poor. The expression site was specifically targeted by using primer AB973 (5' CAG CTT GCA CAC TGG AGG CTA TAG GAC AAG TTA CA 3'), a unique sequence 5' to the expressed *msp3* gene, and primer *msp3* 3' con (5' GCA TCC AAG TTA TTA ATA TCC CTA G 3'), which sits in the 3' *msp3*-specific conserved region of the expressed *msp3* gene. For secondary PCR, primer *msp3* 5' con (5' CTA CAA CAT GAA CTA GCA AAG C 3'), which sits in the 5' *msp3*-specific conserved region was used with the *msp3* 3' con primer (Fig. 2B) (16). PCR conditions for primary PCR

of *msp3* were as follows: (i) denaturation for 5 min at 95°C; (ii) 30 cycles, with 1 cycle consisting of 15 s at 94°C, 10 s at 56°C, and 4 min at 72°C. PCR was done using PCR master mix (Roche). Seminested PCR was performed in an identical manner, except that the extension time was shortened to 1.5 min. Sequences were compiled, analyzed, and aligned using the Vector NTI Suite (Informax, North Bethesda, Md.). From bacteremic peak 1, 28 *msp2* and 23 *msp3* sequences were analyzed; from bacteremic peak 2, 29 *msp2* and 19 *msp3* sequences were analyzed (Fig. 3) (GenBank accession numbers AF540565 to AF540593).

From bacteremic peak 1, four variants of *msp2* were obtained which were identical in the 5' part of the HVR, a position that has been previously described as the block 1 region of the HVR (7, 8). The particular block 1 marker sequence present in these clones is named NAV, and the region corresponding to the *msp2* NAV oligonucleotide probe is underlined in Fig. 3A. There was only a single variant (1-23 in Fig. 3B) of *msp3* detected in bacteremic peak 1. From bacteremic peak 2, 7 days later, 9 *msp2* and 15 *msp3* variants were obtained (Fig. 3). None of the sequences obtained from bacteremic peak 2 were identical to the sequences obtained from bacteremic peak 1 for either gene. Two different block 1 marker sequences were present for *msp2* at bacteremic peak 2, TTV and NAI; however, the block 1 marker TTV was the dominant sequence in the population and was present in 86% (25 of 29) of the sequenced clones (Fig. 3A). Sequence similarity analysis showed that variants within a gene family found in a bacteremic peak were related but that variants found in different bacteremic peaks showed a lower degree of similarity to each other, attributable to the segmental gene conversion mechanism described for both gene families (7, 16). For example, expression site clone 2-311C from peak 2 corresponds to the previously reported TTV pseudogene (GenBank accession no. AF402279). Subsequent to recombination of this pseudogene into the expression site, this sequence has undergone two short sequential changes to give rise to the predominant MSP2 variant for this time point, represented by clone 2-10E

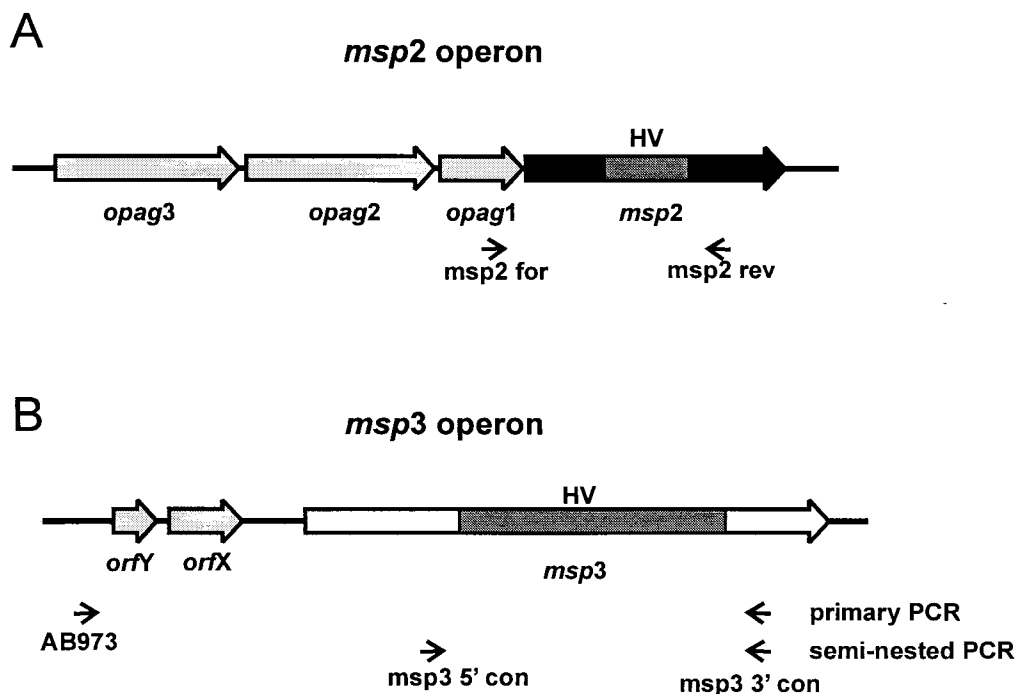


FIG. 2. Schematic representations of the *msp2* and *msp3* expression sites. (A) The *msp2* expression site contains four open reading frames with *msp2* at the 3' end (3). Primers used to clone *msp2* are shown: *msp2* forward (*msp2* for) sits in *opag1* and is specific for the single-copy operon; *msp2* reverse (*msp2* rev) sits in the 3' conserved region of *msp2*. (B) The *msp3* expression site contains three open reading frames. Primers used to clone *msp3* are shown: AB973 is a unique sequence tag 5' to the open reading frames; *msp3* 3' con and *msp3* 5' con sit in the *msp3*-specific conserved regions flanking the hypervariable region. *Msp2* sequences (black bars), *msp3* sequences (white bars), hypervariable regions (HV) (dark gray bars), and operon-associated genes (light gray bars) are shown.

(Fig. 3A). This sequence (2-10E) then served as a template for two further independent recombination events to produce the expression site clones 2-8D and 2-8E. Consequently, 2-311C, 2-10E, 2-8D, and 2-8E are closely related compared to the variants from bacteremic peak 1.

Sequencing of *msp2* and *msp3* clones provided an initial assessment of the variants present in each bacteremia peak and demonstrated a difference in the variants present in different peaks. To quantify the types of variants present at each time point, a cloning reaction mixture of PCR-amplified fragments of each gene at each peak was plated on nylon filters and screened with digoxigenin (DIG)-labeled oligonucleotide probes for the most prevalent variant as determined by sequencing. Screening oligonucleotides were designed that were specific for the block 1 position of the *msp2* HVR and are shown underlined in the corresponding amino acid sequence in Fig. 3A (*msp2* NAV, 5' DIG-GGG GTA ATG CAG TAG AGA ATG CTA CTA AT 3'; *msp2* TTV, 5' DIG-GGG GTA CTA CTG TAG AAG CTG CTA CTA AT 3'). Based on the sequencing results, *msp2* NAV would be expected to identify 100% of the variants present in bacteremic peak 1, and *msp2* TTV would be expected to identify 86% of the variants present in bacteremic peak 2. Finally, all *msp2*-containing clones were identified using a pan-*msp2* probe corresponding to bp 2 to 335. This hybridization identifies the total number of *msp2*-containing clones in the experiment. The *msp3* HVR is less well characterized than the *msp2* HVR, and as such, specific blocks of variation have not been identified. However, the *msp3* sequence present in bacteremic peak 1 was substantially

different from any sequence found in bacteremic peak 2, and the specific oligonucleotide designed for screening is shown overlined in the corresponding amino acid sequence in Fig. 3B and is designated *msp3* 711 (5' DIG-CTG GGA TTG GAA AGA CTG G 3'). An oligonucleotide was designed, based on the sequence results, to detect 100% of the *msp3* variants found in peak 2, and it hybridizes to at least two regions of the HVR. These regions of the HVR are shown underlined in the corresponding amino acid sequence in Fig. 3B. This oligonucleotide is designated *msp3* 718 (5' DIG-CTA GAG GAA CTA GCT GCA ATA AG 3').

Between 619 and 1,449 CFU from each bacteremic peak for each gene were screened by colony hybridization with DIG-labeled oligonucleotides specific for the dominant HVR sequences found at each bacteremic peak. Filters were hybridized at 48°C in a solution containing 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 1% blocking reagent (Roche), 0.1% *N*-lauroylsarcosine, 0.02% sodium dodecyl sulfate (SDS), 0.1 mg of poly(A) per ml, and 40 pmol of DIG-labeled oligonucleotide per ml. Filters were washed twice in 2× SSC–0.1% SDS and twice in 0.2× SSC–0.1% SDS; each wash was for 5 min and at 48°C. Signal was detected according to the manufacturer's directions (Roche). Colonies hybridizing with each probe were counted and divided by the total number of *msp2*- or *msp3*-containing colonies for each bacteremic peak. In the case of *msp2*, the total number of *msp2*-containing clones was identified using the pan-*msp2* probe. Hybridization results of gene-specific clones from each peak show that the predominant *msp2* or *msp3* variant at peak 1 as determined by

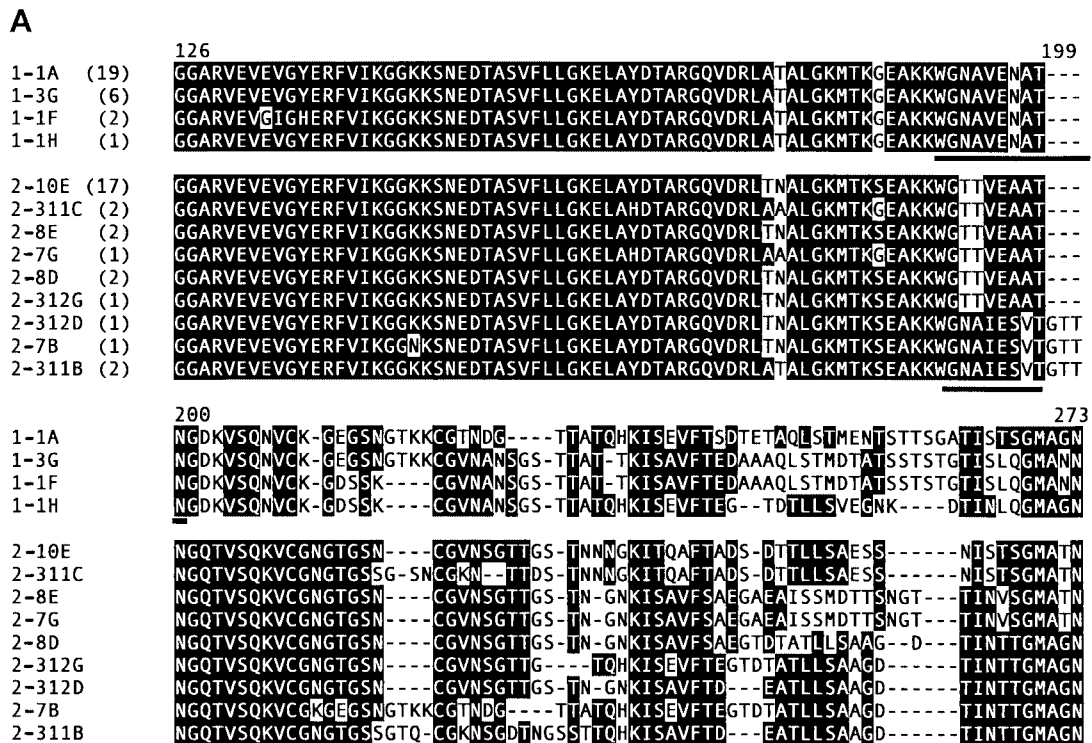


FIG. 3. Expression site sequences for *msp2* and *msp3*. (A) Amino acid sequences deduced from the *msp2* genes found at bacteremic peaks 1 and 2. (B) Amino acid sequences deduced from the *msp3* genes found at bacteremic peaks 1 and 2. All sequences are truncated to show the HVRs only. Regions corresponding to oligonucleotide probes are underlined (*msp2* peak 1 and 2; *msp3* peak 2) or overlined (*msp3* peak 1). Numbers above the sequence correspond to the amino acid positions relative to those of pCKR11.2 (18) for *msp2* and *msp3E-SM16D3* for *msp3* (16). Clone designations begin with 1 or 2 to indicate that the clone was obtained from bacteremic peak 1 or 2, respectively. The numbers in parentheses indicate the number of sequences obtained for each variant. Sequence conservation in the different clones is indicated by white type on black background. Gaps introduced to maximize alignment are indicated by dashes.

sequence analysis was found at levels >90% when a larger population survey is done. Importantly, these variants were found in peak 2 at levels of less than 5% (Table 1), indicating rapid clearance of these variants. Conversely, the predominant variant for each gene found at peak 2 was present at very low levels or absent at peak 1 (Table 1), indicating their rapid emergence under immune selection. As predicted from the sequence analysis, *msp2* TTV was present in only ~84% of the clones found at peak 2. This smaller proportion of clones bearing the TTV marker was due to the presence of two block 1 marker sequences (TTV and NAI) at peak 2. Although we measured only the more prevalent TTV marker sequence, the second NAI marker sequence found in this peak was also distinct from the NAV marker sequence found in clones at peak 1.

The results demonstrate a dramatic simultaneous shift in the expressed variants for both *msp2* and *msp3* within 7 days, with the dominant variant essentially disappearing to be replaced by new variants. As seen in Fig. 1, the population of *A. marginale* at bacteremic peak 1 was controlled, as the organisms expressing distinct *msp2* and *msp3* variants were cleared. The organisms that evaded the immune response survived and replicated within a few days to create the next bacteremic peak and were expressing new variants for both *msp2* and *msp3*. The simultaneous clearance of variants for the two most immunodominant surface proteins of *A. marginale*, followed by the emer-

gence of unique variants, indicates that the switch rates and immune selection for MSP2 and MSP3 are sufficiently similar to explain the waves of bacteremia observed during infection in the immunocompetent host. The absence of conservation between the HVRs of MSP2 and MSP3 means that immune selection acts independently on the variants of each surface protein and is consistent with the ability of the immune system to generate neutralizing antibody against two separate antigens simultaneously. The switching itself may result from independent recombination events for each gene, or there may be a specific mechanism to coordinate recombination for both genes. The former explanation appears most likely, as four *msp2* variants were obtained in peak 1 when only a single variant of

TABLE 1. Quantification of *msp2* and *msp3* variants in sequential bacteremic peaks

Probe	No. ^a of <i>msp2</i> or <i>msp3</i> variants found at:	
	Peak 1	Peak 2
<i>msp2</i> NAV	574/619 (92.7)	26/621 (4.2)
<i>msp2</i> TTV	22/619 (3.6)	519/621 (83.6)
<i>msp3</i> 711	1,208/1,273 (95.9)	7/1,459 (0.5)
<i>msp3</i> 718	0/1,273 (0)	1,404/1,459 (96.2)

^a Number of colonies hybridizing with each probe/total number of colonies containing *msp2* or *msp3* and the percentage (shown in parentheses).

B

262 386
 1-23 (23) LGELKWL EQL ETL ETEELR GELL-----R-----L KARKASOELRTI LAF S KQH LNADE I EGELKSAI GLERL E K LAELE IV-----KORLES KLE K E K K I BEAQ
 2-23 (1) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEVE I FLGL LHNDTETTKI K E K R E L G A G K L K E L V K K L E E L A K K L B P V
 2-6 (2) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEVTVKVG F N E L G L E A K W E K L G L T E L G-----E K L M K T L E K L D S
 2-8 (1) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEVTVKVG F N E L G L E A K W E K L G L T E L G-----E K L M K T L E K L D S
 2-17 (1) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEVTVKVG F N E L G L E A K W E K L G L T E L G-----E K L M K T L E K L D S
 2-10 (2) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEV A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-14 (1) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEV A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-16 (1) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEV A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-21 (1) LGELKWL EQL ETL ETEELR-----VKDRIGE I EAKKI TEV KAAKGEV A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-9 (1) VGELKWL EQL ETL ETEELR-----VKDRIGE I EANK I TEV KAAKGEV A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-15 (3) LGELKWL EQL ETL ETEELR E Q G L E G A L K A L G V E K D K A A L V K F E K E I D N G K T P E E I K L S M I K E I E K K K I E V K A A K G E V A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-20 (1) LGELKWL EQL ETL ETEELR E Q G L E G A L K A L G V E K D K A A L V K F E K E I D N G K T P E E I K L S M I K E I E K K K I E V K A A K G E V A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-13 (1) LGELKWL EQL ETL ETEELR E Q G L E G A L K A L G V E K D K A A L V K F E K E I D N G K T P E E I K L S M I K E I E K K K I E V K A A K G E V A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-18 (1) LGELKWL EQL ETL ETEELR E Q G L E G A L K A L G V E K D K A A L V K F E K E I D N G K T P E E I K L S M I K E I E K K K I E V K A A K G E V A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-2 (1) LGELKWL EQL ETL ETEELR E P G L E G A L K A L G V E K D K A A L V K F E K E I D N G K T P E E I K L S M I K E I E K K K I E V K A A K G E V A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-1 (1) LGELKWL EQL ETL ETEELR E Q G L E G A L K A L G V E K D K A A L V K F E K E I D N G K T P E E I K L S M I K E I E K K K I E V K A A K G E V A L K T W E Q V K L K G G-----E K W K L E E L R G D K L K E L E K K L G T W I V R B K

387 511
 1-23 LTAEEL R E N R E K K L E L A G R G E D P A G L K K A I T E A N A D E L K K G L K E V I K T L E G K K S E L V K G L E G H-----L E K R E K K A L A D K A L E L L K-----K E L K K V L E A E F G E P E E K
 2-23 GGNTG Q G Q V G L V K D K K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-6 --SPGL K E K Q K M K N K K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-8 --SPGL K E K Q K M K N K K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-17 --SPGL K E K Q K M K N K K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V A E L K T W E Q V K L K G G E K W K L E E L R G D K L K E L E K K L G T W I V R B K
 2-10 -LGGKWK R O G E M L K N L K S K L E E L A A I R G L K E E I K I N E P-----G E L K A L E A K K L E E V E A A R G-----V V E K I E K V K E K V G D T G L K K V E L G-----E K F K R I L-----E O D S A
 2-14 -LGGKWK R O G E M L K N L K S K L E E L A A I R G L K E E I K I N E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-16 -LGGKWK R O G E M L K N L K S K L E E L A A I R G L K E E I K I N E P-----G E L K A L E A K K L E E V E A A R G-----V V E K I E K V K E K V G D T G L K K V E L G-----E K F K R I L-----E O D S A
 2-21 -LGGKWK R O G E M L K N L K S K L E E L A A I R G L K E E I K I N E P-----G E L K A L E A K K L E E V E A A R G-----V V E K I E K V K E K V G D T G L K K V E L G-----E K F K R I L-----E O D A A
 2-9 -LGGKWK R O G E M L K N L K S K L E E L A A I R G L K E E I K I N E P-----G E L K A L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-15 -LGGKWK R O G E M L K N L K S K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-20 -LGGK V E A P P R G D V E P E O P O T R G T S C N K G I E G G I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-13 E A R G V E P P R G D V E P E O P O T R G T S C N K G I E G G I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-18 GGNTG Q G Q V G L V K D K K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L
 2-2 --SPGL K E K Q K M K N K K L E E L A A I R G L K E E I K I T E P-----G E L G A K A K K L E E V E A A K G-----E V V P R R T H K S D G S A E K S W A K V K E L G S Q P E T K D L G G L V Q E V K E L E E K V-----K D L Q A
 2-1 GGNTG Q G Q V G L V K D K K L E E L A A I R G L K E E I D K L S E P-----G D L E Q L E A K K L E E V E A A K G-----E V V T K V G F N E L G L K E A E-----W K K L G-----S T E K L

512 636
 1-23 K M E K L E G K I D E L R E K E L I A V I R-----T L K E G P E L V O M L K E I R L E F F O L G L A E T K E V K A L A E K R Q A G G L E L Q E V L Q L T I R T V A L E-----R-----
 2-23 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-6 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-8 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V K G L A E K Q V G G L E I O E I L Q L T I R T V A L E R-----
 2-17 K E K L G K M-----R G E M L K N-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-10 E L E K F A-----K L L G Q E R K N-----L K K L E E L A A I R E L R A L G L E E R L R E L A E T K E V K R L A E R Q N A G E F G I O E G L Q L T E K I T A L S V G I V G G K R L E A K K L G E A D G R L V E K-----L E R L
 2-14 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-16 E L E E K V V K E A K L L G E Q E R K N-----L K K L E E L A A I R E L R A L G L E E R L R E L A E T K E V K R L A E R Q N A G E F G I O E G L Q L T E K I T A L S V G I V G G K R L E A K K L G E A D G R L V E K-----L E R L
 2-21 E L E K F A-----K L L G Q E R K N-----L K K L E E L A A I R E L R A L G L E E R L R E L A E T K E V K R L A E R Q N A G E F G I O E G L Q L T E K I T A L S V G I V G G K R L E A K K L G E A D G R L V E K-----L E R L
 2-9 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-15 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-20 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-13 K R V Q E V A E K-----R K L G E L K D K-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E V K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-18 K M K E L A-----K K L G E Q W K K G F K Q V D M L K N-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E T K E V K L A E K Q V G G L E I O E I L Q L T I R T V A L E R-----
 2-2 K K L G E Q W K G-----F K G Q V D M L K N-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E T K E V K L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L
 2-1 K K L G E Q W K G-----F K G Q V D M L K N-----L K S K L E E L A A I R E L R A L G L E E R L R E L A E T K E V R A L A E K R-----E L D V O G G L Q L T E R I A L G-----G O L D R L E A R R L A E S T E G N L E E V W I L G K L

637 761
 1-23 -----K L V E K V S K L D R K-----S L I S Q K V L E E R K K E L E Q E K L-----L T K E L A R K S D K G F G A L L D L G K A L V L G A G E S K E L-----K-----
 2-23 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-6 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-8 -----N-----L A E K V S K L D R-----E K E Q E M L R L E-----K-----E K L K-----E L E K V Q W H E N R L A G A K O G M I L G E V L S S L E A M H I
 2-17 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-10 G R L D E L V P G K R A Q A L E M K A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-14 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-16 G R L D E L V P G K R A Q A L E M K A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-21 G R L D E L V P G K R A Q A L E M K A T L K Q L-----S V G I F G T I F G L V E R T Y G R E T L G N V L K I Q V A K L E N L O A K L K B A T E K K E I K L R E V K A E I A E Q A S A K F K E K N L V A Q
 2-9 G O L D E L A T G K K L R E S E A T R A T L K Q L-----E T R Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-15 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-20 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-13 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-18 -----N-----L A E K V S K L D R-----E K E Q E M L R L E-----K-----E K L K-----E L E K V Q W H E N R L A G A K O G M I L G E V L S S L E A M H I
 2-2 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E
 2-1 G O L D E L A T G K K L R E S E A T R A T L K Q L K E I K L V D L E R K L E T M G E L K K E I E K L K Q G D L E Q L E A K K I K E V E A A L E K V K S R D R G V S N L Q E K I K E V K E A G K L K E A K I T K L K E L V K E L E K P G T V S O G E L K E

762 825
 1-23 -E L I L K G O T A R I K A L E-----G-----G D I A K L G R I N R I K W L D L A V S K K L K A A L A S A M E I G K N R G
 2-23 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-6 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-8 G T E I V G R A L K D K E V V K L E R-----Q I A K L E R K T I E R I R G L D L A V S K K L K A A L A S A M E I A K N R G
 2-17 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-10 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-14 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-16 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-21 G L N A L E E L E K I K L G L K V E K L A E G L K A L A K I K A L E E L T E I A E K R G V A T V M K A A L T N A M E I T K N R G
 2-9 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-15 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-20 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-13 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-18 G T E I V G R A L K D K E V V K L E R-----Q I A K L E R K T I E R I R G L D L A V S K K L K A A L A S A M E I A K N R G
 2-2 G K E N L K K R L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G
 2-1 G K E N L K K K L E E L A A I R-----E L K E L G L K D W L R L R T L E E L T E I A E K R G V A T V M K A A L T N A M E M A K N R G

FIG. 3—Continued.

Downloaded from http://iai.asm.org/ on October 20, 2020 by guest

msp3 was obtained, indicating that variation is not linked such that each recombination event in *msp2* is coupled to a recombination event in *msp3*. In addition, the physical separation of the two expression site loci within the genome excludes coordinated *cis* regulation. However, the possibility of coordinated *trans* regulation cannot be excluded and may be controlled through the pseudogene complex, which contains a *msp2* and *msp3* pseudogene in a paired tail-to-tail arrangement flanked by a 600-bp repeat (6). If this mechanism were operative, recombination of a *msp2* pseudogene from a particular pseudogene complex into the *msp2* expression site would be accompanied by recombination of the *msp3* pseudogene from the same complex into the *msp3* expression site.

In summary, these findings support the applicability of the single immunodominant surface protein model of antigenic variation to bacteria with a more complex surface coat composed of multiple highly immunogenic outer membrane proteins. The conserved features are a constant rate of switching of the surface antigens combined with strong immune selective pressure (4, 5, 9, 24). Organisms expressing variant surface antigens, the single VSG for trypanosomes or both MSP2 and MSP3 for *Anaplasma*, survive to create the next bloom or peak of infection until that variant is recognized and controlled by the immune response—a pattern and mechanism remarkably similar among vector-borne eukaryotic and prokaryotic pathogens.

We thank Carla Robertson, Pete Hetrick, and Ralph Horn for excellent technical assistance.

This work was supported in part by NIH grants RO1 AI45580 and RO1 AI44005 and USDA-ARS-CRIS 5348-32000-012-00D and USDA-SCA 58-5348-8-044.

REFERENCES

- Alleman, A. R., and A. F. Barbet. 1996. Evaluation of *Anaplasma marginale* major surface protein 3 (MSP3) as a diagnostic test antigen. *J. Clin. Microbiol.* **34**:270–276.
- 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.
- Barbet, A. F., A. Lundgren, J. Yi, F. R. Rurangirwa, and G. H. Palmer. 2000. Antigenic variation of *Anaplasma marginale* by expression of MSP2 mosaics. *Infect. Immun.* **68**:6133–6138.
- Barry, J. D., and R. McCulloch. 2001. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Adv. Parasitol.* **49**:2–70.
- Borst, P. 1991. Molecular genetics of antigenic variation. *Immunol. Today* **12**:A29–A33.
- Brayton, K. A., D. P. Knowles, T. C. McGuire, and G. H. Palmer. 2001. Efficient use of a small genome to generate antigenic diversity in tick-borne ehrlichial pathogens. *Proc. Natl. Acad. Sci. USA* **98**:4130–4135.
- Brayton, K. A., G. H. Palmer, A. Lundgren, J. Yi, and A. F. Barbet. 2002. Antigenic variation of *Anaplasma marginale msp2* occurs by combinatorial gene conversion. *Mol. Microbiol.* **43**:1151–1159.
- Brown, W. C., K. A. Brayton, C. M. Styer, and G. H. Palmer. 2003. The hypervariable region of *Anaplasma marginale* major surface protein 2 (MSP2) contains multiple immunodominant CD4+ T lymphocyte epitopes that elicit variant-specific proliferative and IFN- γ responses in MSP2 vaccinees. *J. Immunol.* **170**:3790–3798.
- Deitsch, K. W., E. R. Moxon, and T. E. Wellems. 1997. Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol. Mol. Biol. Rev.* **61**:281–293.
- 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 rickettsia. *Infect. Immun.* **66**:1200–1207. (Erratum, **66**:2400.)
- French, D. M., W. C. Brown, and G. H. Palmer. 1999. Emergence of *Anaplasma marginale* antigenic variants during persistent rickettsia. *Infect. Immun.* **67**:5834–5840.
- Kieser, S. T., I. S. Eriks, and G. H. Palmer. 1990. Cyclic rickettsia during persistent *Anaplasma marginale* infection of cattle. *Infect. Immun.* **58**:1117–1119.
- Kyes, S. A., J. A. Rowe, N. Kriek, and C. I. Newbold. 1999. Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **96**:9333–9338.
- Losos, G. J. 1986. Anaplasmosis, p. 743–795. In G. J. Losos (ed.), *Infectious tropical diseases of domestic animals*. Longman House, Essex, United Kingdom.
- Luft, B. J., W. Jiang, P. Munoz, R. J. Dattwyler, and P. D. Gorevic. 1989. Biochemical and immunological characterization of the surface proteins of *Borrelia burgdorferi*. *Infect. Immun.* **57**:3637–3645.
- Meeus, P. F. M., K. A. Brayton, G. H. Palmer, and A. F. Barbet. 2003. Conservation of a gene conversion mechanism in two distantly related paralogues of *Anaplasma marginale*. *Mol. Microbiol.* **47**:633–643.
- Nyindo, M., I. Kakoma, and R. Hansen. 1991. Antigenic analysis of four species of the genus *Ehrlichia* by use of protein immunoblot. *Am. J. Vet. Res.* **52**:1225–1230.
- 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., W. C. Brown, and F. R. Rurangirwa. 2000. Antigenic variation in the persistence and transmission of the ehrlichia *Anaplasma marginale*. *Microbes Infect.* **2**:167–176.
- Richards, F. F. 1984. The surface of the African trypanosomes. *J. Protozool.* **31**:60–64.
- Rossouw, M., A. W. Neitz, D. T. de Waal, J. L. du Plessis, L. van Gas, and S. Brett. 1990. Identification of the antigenic proteins of *Cowdria ruminantium*. *Onderstepoort J. Vet. Res.* **57**:215–221.
- 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.
- Smith, J. D., C. E. Chitnis, A. G. Craig, D. J. Roberts, D. E. Hudson-Taylor, D. S. Peterson, R. Pinches, C. I. Newbold, and L. H. Miller. 1995. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**:101–110.
- Turner, C. M. R. 2002. A perspective on clonal phenotypic (antigenic) variation in protozoan parasites. *Parasitology* **125**:S17–S23.
- Wieler, L. H., C. Szattelberger, R. Weiss, R. Bauerfeind, P. Kutzer, K. Failing, and G. Baljer. 1999. Serum antibodies against particular antigens of canine *Borrelia burgdorferi sensu stricto* and their potential in the diagnosis of canine Lyme borreliosis. *Berl. Munch. Tierarztl. Wochenschr.* **112**:465–471.

Editor: J. M. Mansfield