

## Sequence Variation and Immunologic Cross-Reactivity among *Babesia bovis* Merozoite Surface Antigen 1 Proteins from Vaccine Strains and Vaccine Breakthrough Isolates

Tanya LeRoith,<sup>1\*</sup> Kelly A. Brayton,<sup>1</sup> John B. Molloy,<sup>2</sup> Russell E. Bock,<sup>3</sup> Stephen A. Hines,<sup>1</sup> Ala E. Lew,<sup>2</sup> and Terry F. McElwain<sup>1</sup>

Program in Vector-Borne Diseases, Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, Washington 99164-7040<sup>1</sup>; Emerging Technologies/Delivery, Queensland Department of Primary Industries and Fisheries, Moorooka 4105, Queensland, Australia<sup>2</sup>; and Tick Fever Centre, Biosecurity, Queensland Department of Primary Industries and Fisheries, Wacol, Queensland, Australia<sup>3</sup>

Received 21 January 2005/Returned for modification 8 March 2005/Accepted 6 May 2005

The *Babesia bovis* merozoite surface antigen 1 (MSA-1) is an immunodominant membrane glycoprotein that is the target of invasion-blocking antibodies. While antigenic variation has been demonstrated in MSA-1 among strains from distinct geographical areas, the extent of sequence variation within a region where it is endemic and the effect of variation on immunologic cross-reactivity have not been assessed. In this study, sequencing of MSA-1 from two Australian *B. bovis* vaccine strains and 14 breakthrough isolates from vaccinated animals demonstrated low sequence identity in the extracellular region of the molecule, ranging from 19.8 to 46.7% between the T vaccine strain and eight T vaccine breakthrough isolates, and from 18.7 to 99% between the K vaccine strain and six K vaccine breakthrough isolates. Although MSA-1 amino acid sequence varied substantially among strains, overall predicted regions of hydrophilicity and hydrophobicity in the extracellular domain were conserved in all strains examined, suggesting a conserved functional role for MSA-1 despite sequence polymorphism. Importantly, the antigenic variation created by sequence differences resulted in a lack of immunologic cross-reactivity among outbreak strains using sera from animals infected with the *B. bovis* vaccine strains. Additionally, sera from cattle hyperinfected with the Mexico strain of *B. bovis* and shown to be clinically immune did not cross-react with MSA-1 from any other isolate tested. The results indicate that isolates of *B. bovis* capable of evading vaccine-induced immunity contain an *msa-1* gene that is significantly different from the *msa-1* of the vaccine strain, and that the difference can result in a complete lack of cross-reactivity between MSA-1 from vaccine and breakthrough strains in immunized animals.

Parasites in the genus *Babesia* are tick-borne, apicomplexan hemoparasites that cause severe hemolytic anemia, abortion, cerebral babesiosis, and death in susceptible animals. Vaccines directing the immune response against proteins involved in erythrocyte invasion, including merozoite surface proteins, provide a potential control point that targets the extracellular merozoite stage (4–7, 22). However, antigenic variation poses a challenge to the use of surface antigens in vaccines (14, 19, 23, 25).

*Babesia bovis* merozoite surface antigen 1 (MSA-1) and MSA-2 are part of the variable merozoite surface antigen family (10, 15). These proteins are exposed to the host immune system and have immunodominant CD4<sup>+</sup> T lymphocyte epitopes (4, 26). Monospecific antiserum directed against MSA-1 and MSA-2 is able to block entry of the merozoite into the erythrocyte in vitro (15, 21). However, antigenic polymorphism of variable merozoite surface antigen proteins among strains is a general feature. It is clear that enough differences exist among geographically diverse strains to translate into a complete lack of immunologic cross-reactivity using monoclonal antibodies and postinfection immune sera (14, 19, 23, 25).

Limited sequence information is available for MSA-1, and all sequences obtained to date have been derived from strains isolated in distinct geographic areas (27). As a family, variable merozoite surface antigen genes (*vmsa*) share sequence identity or significant similarity in the 5' and 3' signal sequence regions, and in short stretches outside of these regions. Probes based on the 5' and 3' signal sequences hybridize in Southern blots to all members of the variable merozoite surface antigen family. However, cross-hybridization between alleles does not occur when sequences outside the 5' and 3' regions are used as a probe (15). The extent of the sequence variation among *B. bovis* isolates within an endemic area has not been examined. Whether MSA-1 sequences vary among strains from region where it is endemic to the same degree as among strains from distinct geographic regions, or whether MSA-1 sequences are more stable among strains from regions where it is endemic is unknown.

This study was designed to examine the extent of MSA-1 sequence diversity in a biologically and immunologically relevant system using live attenuated vaccine strains from Australia, and organisms isolated from vaccinated cattle that subsequently developed clinical babesiosis (termed breakthrough or outbreak isolates). The breakthrough isolates have been characterized using genotypic markers and have been shown to be genetically different from the vaccine strain (18). While genotypic differences have been identified among these strains and

\* Corresponding author. Mailing address: Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA 99164-7040. Phone: (509) 335-6339. Fax: (509) 335-8529. E-mail: tleroith@vetmed.wsu.edu.

isolates, the variation in specific genes that may be targets for protective immunity, such as MSA-1 and MSA-2, has not been determined. The hypotheses that breakthrough isolates reflect the same degree of sequence diversity as geographic strains and that the sequence diversity results in a lack of immunologic cross-reactivity among the vaccine strains and breakthrough isolates were tested. To test these hypotheses, we examined the extent of MSA-1 sequence diversity among a large number of vaccine strains and breakthrough isolates and determined the effect of the MSA-1 sequence variation on immunologic cross-reactivity.

**MATERIALS AND METHODS**

**Origin of *B. bovis* strains and isolates.** The Mo7 biological clone of *B. bovis* was derived from the Mexico strain by limiting dilution as described (13, 24) and maintained as a cryopreserved stabilate in liquid nitrogen. Parasites were grown in long-term microaerophilic stationary-phase culture by previously described techniques (17). Attenuated *B. bovis* K and T strains were attained by serial passage through splenectomized calves as previously described (2, 3). Vaccine breakthrough isolates were obtained from cattle vaccinated with the T or K strains that subsequently were diagnosed with clinical babesiosis. The isolates were then passaged once through splenectomized calves as described previously (2). Outbreak isolates from animals immunized with the T vaccine were designated as G isolates (*n* = 8), and those from animals immunized with the K vaccine as E and F isolates (*n* = 6) (2, 3).

**Generation of postinfection sera.** Calves were hyperinfected with the Mexico strain of *B. bovis* as previously described (13). Briefly, three 4- to 5-month-old Holstein steers were infected by intravenous subinoculation of 5 × 10<sup>7</sup> infected red blood cells from a splenectomized calf. The cattle were then challenged three more times intravenously with 10<sup>8</sup> infected red blood cells once by subinoculation, and twice using culture-derived parasites. Two weeks after the fourth inoculation, the calves were bled and the serum was used in subsequent immunoblots as described previously (9). Sera were also obtained from animals inoculated with stabilate of the T or K vaccine strains or with blood from an animal experimentally infected with G06 stabilate. Postinfection sera from hyper-infected cattle as well as cattle infected once with the T or K vaccine strain, or the G06 outbreak strain were analyzed for reactivity against homologous recombinant protein and cross-reactivity against heterologous recombinant protein by immunoblot essentially as described (9, 20).

**Sequencing of *msa-1* from vaccine strains and breakthrough isolates.** Genomic DNA from *B. bovis* T and K vaccine strains, and the E86, F3, F28, F35, F40, F64 (K vaccine breakthroughs), G06, G19, G36, G42, G45, G52, G56, and G51 (T vaccine breakthroughs) isolates was used, prepared as previously described by Lew (18). Genomic DNA was suspended in 0.01 M Tris, pH 7.5, and approximately 25 ng was used for PCR. The oligonucleotide primers used to clone *msa-1* from the K, T, E86, F28, F35, F40, F3, F64, G06, G19, G36, G42, and G52 isolates were MSA1 pro (5'-AGAGTGCCTGTGTGCATATAATTC-3') and MSA1 rev2 (5'-TAGCCACAGTCAATCCGCC-3'). The primers used to clone *msa-1* from the G45, G51, and G56 isolates were MSA1 5' flank (5'-GGGAA ACACATTCAGTCAAGC-3'), MSA-1 3' flank (5'-GATGACTATGTCA TTGCCGCTG-3'), and MSA1 rev2. The fragments were cloned into the p0Blunt or pCR4 TOPO cloning vector (Invitrogen). Plasmid DNA was prepared using the Roche High Prime Kit and was sequenced using ABI chemistries. The sequences were assembled and analyzed with the Vector Suite NTI (InforMax).

Phylogenetic analysis was performed with PHYLIP (Phylogeny Inference Package, <http://evolution.genetics.washington.edu/phylip.html> [8]). TMpred ([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)) was used to predict secondary structure and hydrophilicity. NNpredict (<http://cmpharm.ucsf.edu/cgi-bin/nnpredict.pl>) was used to predict tertiary structure and the Vector Suite NTI was used to predict antigenicity.

**Expression of recombinant proteins.** The MSA-1 extracytoplasmic regions of strains Mo7, T, K, G45, G06, G36, F35, and G52 were amplified from sequencing vectors by PCR using primers that lie inside the conserved 5' and 3' regions (Table 1). Amplicons were cloned into the pTrcHis expression vector using the pTrcHis TOPO TA cloning kit (Invitrogen) for expression of rMSA-1. The primers were designed to allow in-frame cloning into the vector. For expression of recombinant protein, a single *Escherichia coli* colony was inoculated into 10 ml of Luria broth containing 50 µg of carbenicillin per ml and incubated overnight at 37°C with shaking. Three ml of the culture was subinoculated into 100 ml Luria broth with 50 µg/ml carbenicillin and incubated at 37°C for 3 h with

TABLE 1. Oligonucleotides used for amplification of *msa-1* genes for expression of the extracellular domain

Primer	Sequence
Mo7-5'	5'-GCCGATACTTCAATCGTCCTTCC-3'
Mo7-3'	5'-ATGTACCCTGTTGCTTGGAGG-3'
T-5'	5'-ATTAATCCTCCATCCGATGCTCC-3'
T-3'	5'-ATTGTCCATGAGGTTACTTGTAGATTG-3'
K-5'	5'-GTCCGATTTTACCCCGATGCTCC-3'
K-3'	5'-ACTGACCATTGAGGTTTCTGTAG-3'
G06-5'	5'-CCTGCTTATGTACTTGTGCTTC-3'
G06-3'	5'-AGCCATTGAGGTTTCCCTGCAGG-3'
G36-5'	5'-ATTGATACTTCATCGCCTCTTCC-3'
G36-3'	5'-ATTGTCCATTGAGGTTACCTTCAGG-3'
G45-5'	5'-CCTGCTTATCTACTGTTGCTTC-3'
G45-3'	5'-AAGAACCATCGGGTTTTGGAGG-3'
G52-5'	5'-GTCCGATTTTACCCGATCTTCC
G52-3'	5'-ATGGTCCACTGAGGTTTCTGTG-3'
F35-5'	5'-GTTAGGATTAATCC-3'
F35-3'	5'-GGTTTCTTGTGGGGTTGTC-3'

shaking. Production of recombinant protein was induced with 1 mM isopropylthiogalactopyranoside (IPTG) and cells were incubated an additional 4 h at 37°C. The cells were pelleted and the pellet was frozen at -20°C overnight. The pellet was thawed and resuspended in 4 ml binding buffer (500 mM NaCl, 20 mM K<sub>2</sub>HPO<sub>4</sub>, 20 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.8) and sonicated to lyse the cells. Recombinant protein was purified on a Ni<sup>2+</sup> column using ProBond resin (Invitrogen), eluted in 1.5 ml fractions with elution buffer (500 mM NaCl, 20 mM K<sub>2</sub>HPO<sub>4</sub>, 20 mM KH<sub>2</sub>PO<sub>4</sub>, pH 4) into 75 µl 1 M Tris, pH 8, and quantified using a microbicinchoninic acid protein reagent kit (Pierce). Relative purity of the antigen was confirmed in Coomassie blue-stained sodium dodecyl sulfate-polyacrylamide gels.

**Nucleotide sequence accession numbers.** These sequences were submitted to GenBank under accession numbers DQ028735 to DQ028747. AF275908: *Babesia bovis*, clone of Mexico Mo7 strain, merozoite surface antigen-1 cDNA sequence. AF275909: *Babesia bovis*, Argentina S2P strain, merozoite surface antigen-1 cDNA sequence. AF275910: *Babesia bovis*, Argentina R1A strain, merozoite surface antigen-1 cDNA sequence. DQ028735: *Babesia bovis*, Australia T vaccine strain, merozoite surface antigen-1 genomic sequence. DQ028736: *Babesia bovis*, Australia K vaccine strain, merozoite surface antigen-1 genomic sequence. DQ028737: *Babesia bovis*, Australia G19 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028738: *Babesia bovis*, Australia F64 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028739: *Babesia bovis*, Australia E86 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028740: *Babesia bovis*, Australia F35 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028741: *Babesia bovis*, Australia G06 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028742: *Babesia bovis*, Australia G52 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028743: *Babesia bovis*, Australia F28 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028744: *Babesia bovis*, Australia F40 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028745: *Babesia bovis*, Australia G36 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028746: *Babesia bovis*, Australia G51 outbreak isolate, merozoite surface antigen-1 genomic sequence. DQ028747: *Babesia bovis*, Australia G45 outbreak isolate, merozoite surface antigen-1 genomic sequence.

**RESULTS**

**Comparative analysis of MSA-1 sequence and predicted structure.** Limited sequence information is available for MSA-1 and has been obtained only among strains from distinct geographic areas (27). The extent of sequence diversity within an region where it is endemic has not been examined. In addition, while genotypic differences have been identified between vaccine strains and breakthrough isolates, it is unknown whether MSA-1 from the breakthrough isolates varies from vaccine strains. To address these questions, the *msa-1* gene

```

(1) 1 10 20 30 40 50 60 70 80
MSA1 G06 (1) AVLTFLIT V CAVSFASCA-----QENPQSIETPPAYVPVASSLLYD KRTFFYNLMGKLYIGVDEDIKKI LAGE
MSA1 Mo7 (1) ATEFALPIS L CVLAIISAG-----EELTQSDVRNADTIVLPEGSFYD MSKFYGAVGSFDC--TKLYSVLSANF
MSA1 G52 (1) ATEFVLPIS L CVSAVVSSC-----PRSCENVIRVVDSSSDLPEGYLYD MAKFYGAVESFDR--TKMYGVIAANF
MSA1 G19 (1) ATEFVLPIS L CVSAVVSSC-----PRSCENVIRVVDSSSDLPEGYLYD MAKFYGAVESFDR--TKMYGVIAANF
MSA1 F64 (1) ATEFVLPIS L CVSAVVSSC-----PRSCENVIRVVDSSSDLPEGYLYD MAKFYGAVESFDR--TKMYGVIAANF
MSA1 K (1) VTFSLPIT L CVSAVTSTE-----PESTQSFRLVDSSPDAPFEGFLYD MINFYGAVESFGR--NRLYGLISANF
MSA1 F40 (1) VTFSLPIT L CVSAVTSTE-----PESTQSFRLVDSSPDAPFEGFLYD MINFYGAVESFGR--NRLYGLISANF
MSA1 G36 (1) ANVFVFPIS L CFSAVTSSAEVFP---EQPAPSFVRVDTSSPLPEGSLYD MNKFYGAVESFDR--GRLYGVIAANF
MSA1 F35 (1) VTFALPIS L FVSAVTSAG-----EVPAPAPINVRINPPSDAPKGLIVD MTKFYGAVESFDR--NRLYGLISANF
MSA1 T (1) VTFALPIS L FVSAVTSAG-----EVPAPAPINVRINPPSDAPKGLIVD MTKFYGAVESFDR--NRLYGLISANF
MSA1 E86 (1) VTFALPIS L FVSAVTSAG-----EVPAPAPINVRINPPSDAPKGLIVD MTKFYGAVESFDR--NRLYGLISANF
MSA1 S2P (1) ATEFALPIS L CVSATASSSEDLAQESAQISTPINVRIINPPYDAPKGLIED MTKFYGAVESFDR--NRLYGLISANF
MSA1 R1A (1) ATEFALPIS L CVSATASSSEDLAQESAQISTPINVRIINPPYDAPKGLIED MTKFYGAVESFDR--NRLYGLISANF
MSA1 G45 (1) AVLTFLMT A CAVSLATST-----LQNSNDILPPAYVPVAFSGILYE KRTFFYNLMIKLESSLNELEKVVHGS
MSA1 F28 (1) AVLTFLMT A CAVSLATST-----LQNSNDILPPAYVPVAFSGILYE KRTFFYNLMIKLESSLNELEKVVHGS
MSA1 G51 (1) AVLTFLMT A CAVSLATST-----LQNSNDILPPAYVPVAFSGILYE KRTFFYNLMIKLESSLNELEKVVHGS

(81) 81 90 100 110 120 130 140 150 160
MSA1 G06 (73) DNVDAIVRDAKNGIVRNRKVAELL EDKCFSSS---GFTITSTEARAKSVEGD FESIMY VTCIYHRCVALNKVNR FVNA
MSA1 Mo7 (71) KAARKMLDQKVKDTFKNLYKVNALI NNPMIRPLENATIVSGFSTKNEEK NAIFD IKGMYYRACHMDKYLKS RWNT
MSA1 G52 (71) RNGGLADEKVKKEAFISLYKVSASI VNPMSINSLKFAAVSFSGKTEKDK NAIFE IKTMVRRICHMOTFLKGNWSS
MSA1 G19 (71) RNGGLADEKVKKEAFISLYKVSASI VNPMSINSLKFAAVSFSGKTEKDK NAIFE IKTMVRRICHMOTFLKGNWSS
MSA1 F64 (71) RNGGLADEKVKKEAFISLYKVSASI VNPMSINSLKFAAVSFSGKTEKDK NAIFE IKTMVRRICHMOTFLKGNWSS
MSA1 K (71) TIAGMKACQVENTFTYLYKVDMI KNPITISSEKFNETAASKFSDKTEKEK DAIFD VKSMIKRITHMOTFLKT DWSS
MSA1 F40 (71) TIAGMKACQVENTFTYLYKVDMI KNPITISSEKFNETAASKFSDKTEKEK DAIFD VKSMIKRITHMOTFLKT DWSS
MSA1 G36 (73) KAVRMDQKVENAFTYLYKVDMI KNPITISSEKFNETAASKFSDKTEKEK DAIFD VKSMIKRITHMOTFLKT DWSS
MSA1 F35 (73) HVADMSAQCQVENTFTYLYKVKEMI NNPMISSEKFNETAASKFSDKTEKEK NAIFD LMSMNNIHHMKFLKS KWEF
MSA1 T (73) HVADMSAQCQVENTFTYLYKVKEMI NNPMISSEKFNETAASKFSDKTEKEK NAIFD LMSMNNIHHMKFLKS KWEF
MSA1 E86 (73) HVADMSAQCQVENTFTYLYKVKEMI NNPMISSEKFNETAASKFSDKTEKEK NAIFD LMSMNNIHHMKFLKS KWEF
MSA1 S2P (79) HVADMSTCQVDNTFTYLYKVKEMI NNPMISSEKFNETAASKFSDKTEKEK NAIFD LMSMNNIHHMKFLKS KWEF
MSA1 R1A (79) HVADMSTCQVDNTFTYLYKVKEMI NNPMISSEKFNETAASKFSDKTEKEK NAIFD LMSMNNIHHMKFLKS KWEF
MSA1 G45 (73) DLMYLEIDKAKKGLKMKETAVSL DNKFFTAL---GFVTSAGLVASVNGN ETEMD VTCIYHRCVALNKVNR FVNG
MSA1 F28 (73) EQMAYCVANIKDGLIKMKETAKSL DDNCFSE---GFVTSAEFAKRVVGD DTEMG VTCIYHRCVALNKVNR FVNG
MSA1 G51 (73) EQMAYCVANIKDGLIKMKETAKSL DDNCFSE---GFVTSAEFAKRVVGD DTEMG VTCIYHRCVALNKVNR FVNG

(161) 161 170 180 190 200 210 220 230 240
MSA1 G06(150) KESEESIAEMKKYFKKNIYDKD-TVSNQGLIAGDA VK-RED DSIVRTAAEYFKKMGAPGGGHSFEARESEGPAGFSA
MSA1 Mo7(151) DIVEEDREKAVEYFKKHVYTCHEV DVNGMAGVCKE LSPASD YKLVESEFAFAKAVHAQVGNFVKPCDITAPPEVIT
MSA1 G52(151) GLTIDGRKRVKVEDYFKKHVYTCHEV DVNGMAGVCKE LSEGSN NKLWVSEDFEARGRHYGHAEDYVEPNFDAITLPEGLL
MSA1 G19(151) GLTIDGRKRVKVEDYFKKHVYTCHEV DVNGMAGVCKE LSEGSN NKLWVSEDFEARGRHYGHAEDYVEPNFDAITLPEGLL
MSA1 F64(151) GLTIDGRKRVKVEDYFKKHVYTCHEV DVNGMAGVCKE LSEGSN NKLWVSEDFEARGRHYGHAEDYVEPNFDAITLPEGLL
MSA1 K(151) DIKEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LGEESY YKLWVNVDDFETGKRLARVDHFVTPESAIMPPEELI
MSA1 F40(151) DIKEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LGEESY YKLWVNVDDFETGKRLARVDHFVTPESAIMPPEELI
MSA1 G36(155) QMTORNRKAEYFKKHVYTCHEV DVNGMAGVCKE LSVVSY YKLAVYEDDFSKAKHEAGNGNFTPTPETVPPPEELI
MSA1 F35(153) QMTEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LSVVSY YKLAVYEDDFVIAKHLARVNTFVTPESLIVPPEELI
MSA1 T(153) QMTEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LSVVSY YKLAVYEDDFVIAKHLARVNTFVTPESLIVPPEELI
MSA1 E86(153) QMTEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LSVVSY YKLAVYEDDFVIAKHLARVNTFVTPESLIVPPEELI
MSA1 S2P(159) QMTEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LSVVSY YKLAVYEDDFVIAKHLARVNTFVTPESLIVPPEELI
MSA1 R1A(159) QMTEEDRRKAEYFKKHVYTCHEV DVNGMAGVCKE LSVVSY YKLAVYEDDFVIAKHLARVNTFVTPESLIVPPEELI
MSA1 G45(150) QETENIKEMNKYFNKNVYNSD-N DNKALVITGNM VK-AEN DSIVRTAAEYFKKMGAGGSGNSMAPSDNSSTGGSG
MSA1 F28(150) EERKENIEMKTYFNKNVYNSD-D EGKALVNVGSM LE-GEN DSVVGTAAEYFKKMGAGGSGNSMAPSDNSSTGGSG
MSA1 G51(150) EERKENIEMKTYFNKNVYNSD-D EGKALVNVGSM LE-GEN DSVVGTAAEYFKKMGAGGSGNSMAPSDNSSTGGSG

(241) 241 250 260 270 280 290 300 310 320
MSA1 G06(228) TQERTPEFPS-----PQG-----PPAESSQAECFAGNLNCPG
MSA1 Mo7(231) DALFKELQECKP-ARSPSTEVPAFC-----DASGVQPPASGTSQCF-APTTPSPSPSSS---GNLQGGCQ
MSA1 G52(231) GELERQLATIKP--AEEINRSPGTGSDPHPPADGCAQOETPAPSSQPCPTGAQSQTQCAENSSPESTPGNLSGPKG
MSA1 G19(231) GELERQLATIKP--AEEINRSPGTGSDPHPPADGCAQOETPAPSSQPCPTGAQSQTQCAENSSPESTPGNLSGPKG
MSA1 F64(231) GELERQLATIKP--AEEINRSPGTGSDPHPPADGCAQOETPAPSSQPCPTGAQSQTQCAENSSPESTPGNLSGPKG
MSA1 K(231) AETLEREQAPSPSGCTENGDLAAPS-----E---PGAGAEPSQGTPTGTOQFCAPAPQSPPTQES--TGNLNGGQ
MSA1 F40(231) AETLEREQAPSPSGCTENGDLAAPS-----E---PGAGAEPSQGTPTGTOQFCAPAPQSPPTQES--TGNLNGGQ
MSA1 G36(235) IAVESEQEEPC-----RTENGDLRQCCAGSGSGE-GVSVVQPTGKQDQSTVQFAAPSQCADQIKP--EGNLNGQ
MSA1 F35(233) ASITEREAPPQ-----GSTVLE-----TGC-QPQOETFEARFGAPSLPSADQAPQC--EGNLGQPSKPA
MSA1 T(233) ASITEREAPPQ-----GSTVLE-----TGC-QPQOETFEARFGAPSLPSADQAPQC--EGNLGQPSKPA
MSA1 E86(233) ASITEREAPPQ-----GSTVLE-----TGC-QPQOETFEARFGAPSLPSADQAPQC--EGNLGQPSKPA
MSA1 S2P(239) ASITEREQPQETSACQETTGCPUTAGQPSPPGS-PKEPAGEQSQENSSGLFAAHNTPSADQESKP--EGNLGQ
MSA1 R1A(239) ASITEREQPQETSACQETTGCPUTAGQPSPPGS-PKEPAGEQSQENSSGLFAAHNTPSADQESKP--EGNLGQ
MSA1 G45(228) TGAFGSG-----PAGQOEFEARFSGNLOKFDGNSPS
MSA1 F28(228) PTFSGSN-----PTER-ESEARSFSGNLOKPE---TP
MSA1 G51(228) PTFSGSN-----PTER-ESEARSFSGNLOKPE---TP

(321) 321 338
MSA1 G06(261) ---K--P--FTFGGLTVA
MSA1 Mo7(293) ---TKPAGSFTFGGLTVA
MSA1 G52(306) ---PSKTSFTFGGLTVA
MSA1 G19(306) ---PSKTSFTFGGLTVA
MSA1 F64(306) ---PSKTSFTFGGLTVA
MSA1 K(297) ---KPAGSFTFGGLTVA
MSA1 F40(297) ---KPAGSFTFGGLTVA
MSA1 G36(304) ---EPAGSFTFGGLTVA
MSA1 F35(291) ETPKPTGFTFGGLTVA
MSA1 T(296) ---NPKPFTFGGLTVA
MSA1 E86(296) ---NPKPFTFGGLTVA
MSA1 S2P(312) ---AGSFTFGGLTVA
MSA1 R1A(313) ---SPKPTGFTFGGLTVA
MSA1 G45(280) ---KPAGSFTFGGLTVA
MSA1 F28(256) ---KPTGFTFGGLTVA
MSA1 G51(256) ---KPTGFTFGGLTVA

```

from two Australian vaccine strains, designated K and T (2), and multiple breakthrough isolates were cloned and sequenced. E and F outbreak isolates (E86, F28, F35, F40, F3, and F64) were obtained from cattle immunized with the K strain (2), and G outbreak isolates (G06, G19, G36, G42, G45, G51, G56, and G52) were obtained from cattle immunized with the T strain (3).

The deduced amino acid sequences of the MSA-1 extracellular domain among all isolates in this region of endemicity in Australia vary extensively, ranging from only 17% sequence identity between the G51 and G64 isolates to 99% sequence identity between the G19 and F64 isolates. When vaccine strains are compared directly to the associated outbreak isolates, sequence identity ranged from 19.8 to 46.7% between the T vaccine strain and the eight T breakthrough isolates, and from 18.7 to 99% between the K vaccine strain and the six K breakthrough isolates. Additionally, a phylogenetic analysis of all MSA-1 sequences (PHYLIP, <http://evolution.genetics.washington.edu/phylip.html> [8]) indicates that all sequences cluster into three to four major groups and MSA-1 from the vaccine breakthrough isolates, with one exception (F40), are in distinct branches from their respective vaccine strain (Fig. 2). The F40 MSA-1 sequence shares 99% identity with the K vaccine strain. Interestingly, some of the Australia MSA-1 sequences are more closely related to the Mo7 cloned laboratory strain or Argentina R1A and S2P strains than to other Australia isolates (Fig. 1 and 2).

The majority of strictly conserved amino acids are in the amino terminal and carboxy terminal signal sequence regions, as previously demonstrated (15). In the amino terminal leader sequence, 3 of 21 amino acids are strictly conserved, while in all strains and isolates except G51, six of nine amino acids in the carboxy-terminal glycerylphosphatidylinositol signal sequence are strictly conserved (Fig. 1). Throughout the rest of the molecule, only single amino acids or small clusters of amino acids are conserved in all isolates, including a strictly conserved YFK motif at amino acids 174 to 176 (Fig. 1). The last strictly conserved amino acid in the extracellular domain is at position 205, and there is significantly more variation in the carboxy terminus (approximately 100 amino acids) of the molecule.

Even though overall sequence conservation is relatively low among different strains and isolates, predictive algorithms (TMpred and NNpredict) suggest that the molecules have a similar membrane structure. Six representative analyses are shown in Fig. 3. Each MSA-1 has a predicted amino and carboxy terminal hydrophobic domain, and similarly located  $\alpha$ -helices,  $\beta$ -pleated sheets (data not shown), and hydrophilic peaks within the extracytoplasmic region. All strains contain a major hydrophilic peak that corresponds approximately to amino acids 130 to 180, and contains the strictly conserved YFK motif. Additionally, most strains and isolates contain two to three additional, often similarly located, hydrophilic peaks between amino acids 30 to 130 and 180 to 280 (Fig. 3).

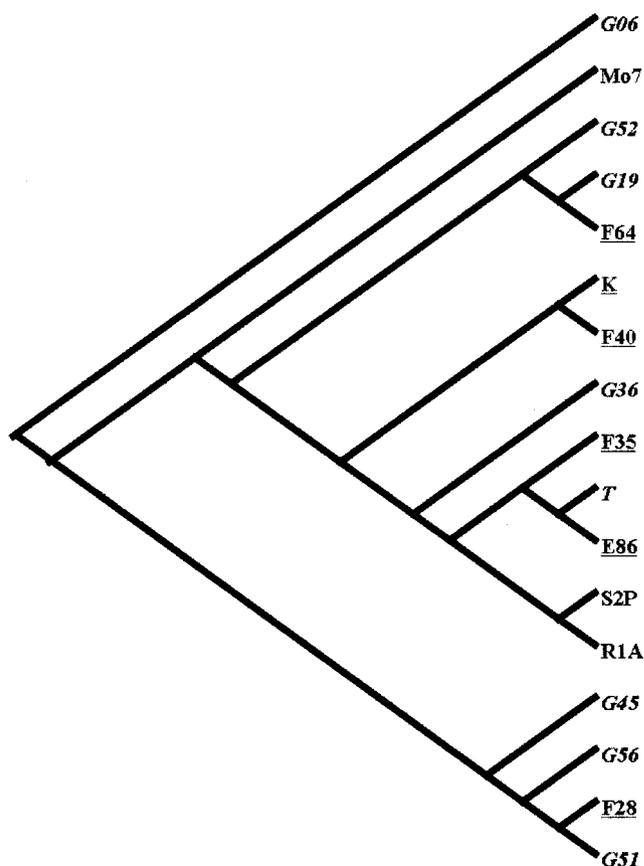


FIG. 2. Dendrogram of MSA1 sequences from various *B. bovis* strains, including Australian vaccine strains and outbreak isolates. Mo7 is the biological clone of a Mexico isolate; R1A and S2P are Argentina strains; T and K strains are Australian vaccine strains; G isolates are T vaccine breakthrough isolates; and E and F are K vaccine breakthrough isolates.

**Antigenic cross-reactivity among *B. bovis* strains – postinfection sera.** Previous studies suggest that there is a complete lack of MSA-1 cross-reactivity among strains from distinct geographic areas (14, 23, 25). The availability of postinfection sera against the Australia vaccine strains allowed us to determine whether the sequence diversity in breakthrough isolates also resulted in a lack of cross-reactivity. Recombinant his-tagged MSA-1 from the Mo7 laboratory strain, the T and K vaccine strains, and the G06, G45, G36, G52, and F35 breakthrough isolates were used for these experiments. Postinfection serum against the T vaccine strain and the T vaccine outbreak isolate G06 only reacted with homologous MSA-1 antigen (Fig. 4A and B). Antiserum against the K vaccine strain reacted with rMSA-1 from both the K strain and G45 isolate, a T vaccine outbreak isolate (Fig. 4C). Interestingly, the G45 isolate

FIG. 1. Sequence alignments of MSA1 from selected *B. bovis* strains and isolates. Strictly conserved amino acids are in yellow boxes, partially conserved amino acids are in blue and green. Mo7 is the biological clone of the Mexico strain; the T and K strains are Australia vaccine strains; G isolates are T vaccine breakthroughs; F isolates are K vaccine breakthroughs. R1A and S2P are Argentina strains. Mo7, R1A, and S2P have been previously published (27).

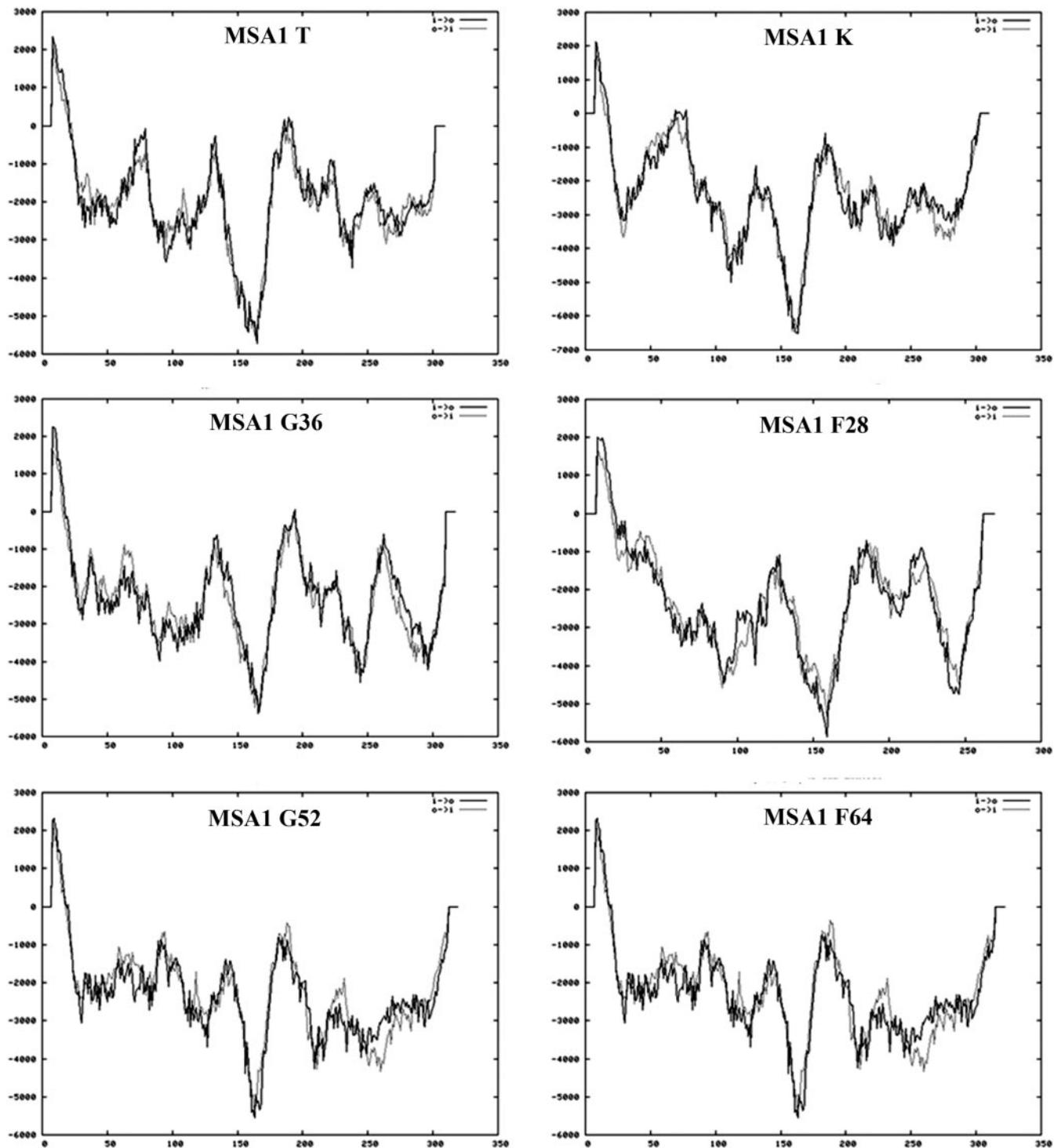


FIG. 3. Predicted membrane structure of MSA1 from T and K vaccine strains, and G36, F28, G52, and F64 breakthrough isolates. The predicted membrane orientation was calculated with TMpred and is plotted on the y axis against the amino acid position on the x axis. The solid line,  $i \rightarrow o$ , is the calculation performed from N terminus to C terminus. The dashed line,  $o \rightarrow i$ , is calculated from C terminus to N terminus.

MSA-1 is more distantly related to the K vaccine strain than G52 (Fig. 2) to which there was no cross-reactivity with K vaccine specific antiserum.

The ability to induce more extensive cross-reactivity after multiple infections with the same strain was examined. Immu-

noblot analysis was performed using serum from three different cattle (only data from one animal shown), each infected four times with the *B. bovis* Mexico strain (13) and demonstrated to be immune to challenge. The sera were used as a probe against recombinant protein from the homologous Mo7

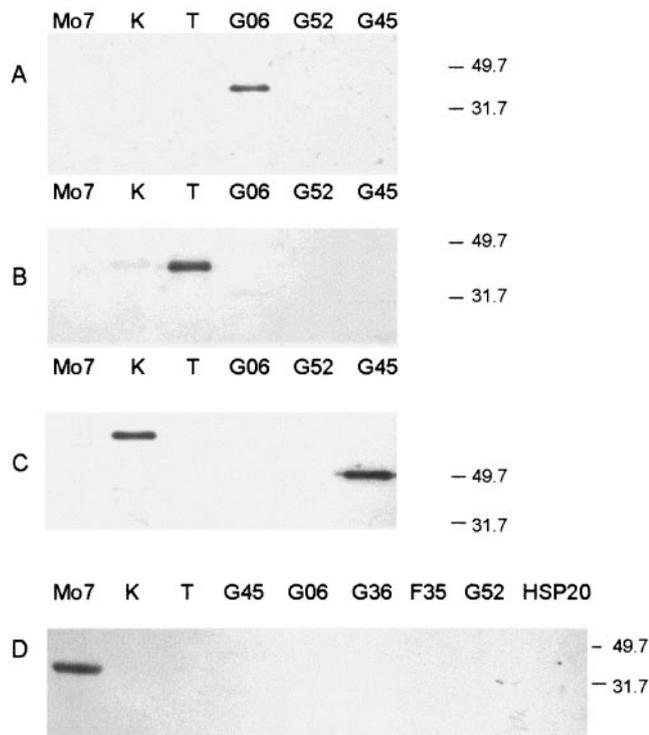


FIG. 4. MSA-1 cross-reactivity of postinfection sera. Recombinant MSA-1 from multiple strains and isolates (as noted) and from recombinant *B. bovis* heat shock protein 20 (HSP20) (as a negative control) were probed with postinfection bovine serum from a calf infected once with the G06 isolate (panel A), the T vaccine strain (panel B), or the K vaccine strain (panel C). Panel D is an immunoblot of the same antigens infected with the Mexico strain and challenged three times with the homologous strain.

strain and heterologous Australia strains and isolates. The results demonstrated strong reactivity of the hyper-infection sera with homologous rMSA-1 from the Mo7 strain (Fig. 4D). However, no reactivity with heterologous protein could be detected, suggesting that multiple exposures to homologous MSA-1 during vaccination or natural infection do not result in development of antibody against B-lymphocyte epitopes shared with heterologous MSA-1.

## DISCUSSION

In this study, the sequence variation of *B. bovis* MSA-1 within an endemic area in a biologically relevant system of vaccine strains and vaccine breakthrough isolates were examined. Since escape from clinical immunity almost certainly involves variation in multiple genes and gene products, some of which vary quite rapidly (1), we postulated that MSA-1 would vary in at least some, but probably not all, of the breakthrough isolates. However, extensive MSA-1 sequence variation was detected between the attenuated strains used as modified live vaccines in Australia, and, with one exception, every isolate from vaccinated animals that were subsequently diagnosed with clinical babesiosis. Since MSA-1 is encoded by a single copy gene, it was somewhat surprising that the MSA-1 sequences varied so extensively, and that they varied in all but one breakthrough isolate. Similar, but less extensive variation

has been seen in the single copy surface protein genes of other apicomplexan hemoparasites, including the merozoite surface proteins of *Plasmodium falciparum* (28), the major merozoite-piroplasm surface antigen (Tams1) of *Theileria annulata*, and the polymorphic immunodominant molecule (PIM) of *Theileria parva* (11, 12). Genetic exchange is postulated to be the major means of generating polymorphism in these single copy genes, and in MSA-1, and most likely arises during diploid or polyploid stages in the respective vectors (11, 12). No obvious sites of recombination could be identified among the MSA-1 sequences derived in this study, unlike the related MSA-2 genes in which a conserved site of recombination has been found (10).

Alignment of amino acid sequences of MSA-1 from these strains as well as from laboratory and previously published strains demonstrated that the majority of strictly conserved amino acids were located in the N terminal and C terminal signal sequences, as previously found (15), with only short stretches of strictly or partially conserved amino acids in the extracellular, hydrophilic domains. Despite limited sequence conservation, however, analysis of the proteins from all strains demonstrated a predicted conserved membrane orientation and hydrophilic peak location, suggesting a conserved function for the protein. The region that corresponds to the most hydrophilic domain contains a YFK motif that is strictly conserved in all MSA-1 sequences from all strains and isolates evaluated. The strict conservation of this tripeptide, as well as the nine additional individual amino acids in the extracytoplasmic domain that are strictly conserved, suggests that they may be essential for correct structure and function of the protein. The Tams1 polypeptide of *Theileria annulata* contains a strictly conserved KEK motif, as well as other conserved K residues (16). Lysine is thought to be involved in intra- and intermolecular cross-linking, and may be involved in erythrocyte binding. Of the other strictly conserved amino acids, 9/12 are in predicted antigenic regions (data not shown), suggesting they may contribute to epitopes that can induce cross-reactive antibody.

A dendrogram generated using all known MSA-1 sequences demonstrated that, with one exception (F40 and K vaccine), each outbreak isolate was in a separate branch from its associated vaccine strain. Interestingly, previous studies have demonstrated similarities between F40 and K, including the ability of the K vaccine to protect against F40 challenge (2, 18), and suggested that F40 may not truly represent a breakthrough isolate. All other isolates in this and previous studies have been capable of overcoming vaccine-induced immunity in challenge trials (2, 18).

Selection for MSA-1 diversity could be driven by immune pressure, by a separate biological process, or both. The overall population structure of MSA-1 is unknown in this endemic area. It is possible that the sequence differences are simply a result of random sampling in a markedly divergent population. However, since all but one of the MSA-1 sequences from breakthrough isolates were different from the vaccine strain, and immune sera from cattle infected with a vaccine strain did not cross-react with recombinant MSA-1 protein from any of the associated outbreak isolates evaluated, immune selection of MSA-1 variants may be a significant factor in development of vaccine breakthroughs. Interestingly, cross-reactivity could

not be generated in postinfection serum even after infection and three challenges (13). These data suggest that once a response to particular MSA-1 B-lymphocyte epitopes is established in an infected or vaccinated animal, hyperinfection results in an anamnestic response to the same epitopes but does not induce a response to additional epitopes that are shared with heterologous MSA-1.

The role of MSA-1 in the merozoite membrane, either as a structural protein or as a critical part of the erythrocyte attachment and invasion process is unknown. The ability of antibody against MSA-1 to block erythrocyte invasion of both the merozoite and the sporozoite in vitro suggests a critical function in host cell entry. In this region where it is endemic, there is clearly selection for parasites that express allelic variants of MSA-1. The lack of cross-reactivity observed using postinfection serum suggests that immune pressure may play a role in selection of these parasites.

#### ACKNOWLEDGMENTS

The skillful technical assistance of Bev Hunter, Carla Robertson, and Deb Alperin is acknowledged. Additionally, we thank Guy Palmer for helpful discussions and critical review of the manuscript, and the anonymous reviewer for valuable comments.

This work was supported by NIH K08 Award 1 K08 A151391, and USDA SCA58-5348-2-683.

#### REFERENCES

- Allred, D. R., and B. Al-Khedery. 2004. Antigenic variation and cytoadhesion in *Babesia bovis* and *Plasmodium falciparum*: different logics achieve the same goal. *Mol. Biochem. Parasitol.* **134**:27–35.
- Bock, R. E., de Vos, A. J., T. G. Kingston, I. A. Shiels, and R. J. Dalgliesh. 1992. Investigations of breakdowns in protection provided by living *Babesia bovis* vaccine. *Vet. Parasitol.* **43**:45–56.
- Bock, R. E., de A. J. Vos, A. Lew, T. G. Kingston, and I. R. Fraser. 1995. Studies on failure of T strain live *Babesia bovis* vaccine. *Australian Vet. J.* **72**:296–300.
- Brown, W. C., and G. H. Palmer. 1999. Designing blood-stage vaccines against *Babesia bovis* and *Babesia bigemina*. *Parasitol. Today* **15**:275–281.
- Brown, W. C., and Rice-A. C. Ficht. 1994. Use of helper T cells to identify potential vaccine antigens of *Babesia bovis*. *Parasitol. Today* **10**:145–149.
- Chitnis, C. E., and M. J. Blackman. 2000. Host cell invasion by malaria parasites. *Parasitol. Today* **16**:407–408.
- Dubremetz, J. F., Garcia-Reguet, N., V. Conseil, and M. N. Fourmaux. 1998. Apical organelles and host-cell invasion by Apicomplexa. *Int. J. Parasitol.* **28**:1007–1013.
- Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, Wash.
- Fisher, T. G., T. F. McElwain, and G. H. Palmer. 2001. Molecular basis for variable expression of merozoite surface antigen gp45 among American isolates of *Babesia bigemina*. *Infect. Immun.* **69**:3782–3790.
- Florin-Christensen, M., C. E. Suarez, S. A. Hines, G. H. Palmer, W. C. Brown, and T. F. McElwain. 2002. The *Babesia bovis* Merozoite Surface Antigen 2 locus contains four tandemly arranged and expressed genes encoding immunologically distinct proteins. *Infect. Immun.* **70**:3566–3575.
- Geysen, D., T., B. Bazarusanga, B. Brandt, and T. T. Dolan. 2004. An unusual mosaic structure of the PIM gene of *Theileria parva* and its relationship to allelic diversity. *Mol. Biochem. Parasitol.* **133**:163–174.
- Gubbels, M. J., F. Katzer, G. Hide, F. Jongejan, and B. R. Shiels. 2000. Generation of a mosaic pattern of diversity in the major merozoite-piroplasm surface antigen of *Theileria annulata*. *Mol. Biochem. Parasitol.* **110**:23–32.
- Hines, S. A., T. F. McElwain, G. M. Buening, and G. H. Palmer. 1989. Molecular characterization of *Babesia bovis* merozoite surface proteins bearing epitopes immunodominant in protected cattle. *Mol. Biochem. Parasitol.* **37**:1–10.
- Hines, S. A., G. H. Palmer, D. P. Jasmer, W. L. Goff, and T. F. McElwain. 1995. Immunization of cattle with recombinant *Babesia bovis* merozoite surface antigen-1. *Infect. Immun.* **63**:349–352.
- Hines, S. A., G. H. Palmer, D. P. Jasmer, T. C. McGuire, and T. F. McElwain. 1992. Neutralization-sensitive merozoite surface antigens of *Babesia bovis* encoded by members of a polymorphic gene family. *Mol. Biochem. Parasitol.* **55**:85–94.
- Katzer, F., S. McKellar, M. A. Ferguson, d'Oliveira, C., B. R. and Shiel. 2002. A role for tertiary structure in the generation of antigenic diversity and molecular association of the Tams1 polypeptide in *Theileria annulata*. *Mol. Biochem. Parasitol.* **122**:55–67.
- Levy, M. G., and M. Risti. 1980. *Babesia bovis*: continuous cultivation in a microaerophilous stationary phase culture. *Science* **207**:1218–1220.
- Lew, A. E., R. E. Bock, J. M. Croft, C. M. Minchin, T. G. Kingston, and R. J. Dalgliesh. 1997. Genotypic diversity in field isolates of *Babesia bovis* from cattle with babesiosis after vaccination. *Australian Vet. J.* **75**:575–578.
- Madrugá, C. R., C. E. Suarez, T. F. McElwain, and G. H. Palmer. 1996. Conservation of merozoite membrane and apical complex B cell epitopes among *Babesia bigemina* and *Babesia bovis* strains isolated in Brazil. *Vet. Parasitol.* **61**:21–30.
- Molloy, J. B., S. J. Waldron, and W. K. Jorgensen. 1995. Identification of an immunodominant 40 kDa merozoite antigen common to the Australian T and Dixie vaccine strains of *Babesia bovis* and the development of diagnostic tests specific for these strains. *Vet. Parasitol.* **60**:229–240.
- Mosqueda, J., T. F. McElwain, and G. H. Palmer. 2002. *Babesia bovis* merozoite surface antigen 2 proteins are expressed on the merozoite and sporozoite surface, and specific antibodies inhibit attachment and invasion of erythrocytes. *Infect. Immun.* **70**:6448–6455.
- Palmer, G. H., and T. F. McElwain. 1995. Molecular basis for vaccine development against anaplasmosis and babesiosis. *Vet. Parasitol.* **57**:233–253.
- Palmer, G. H., T. F. McElwain, L. E. Perryman, W. C. Davis, D. R. Reduker, D. P. Jasmer, V. Shkap, E. Pipano, W. L. Goff, and T. C. McGuire. 1991. Strain variation of *Babesia bovis* merozoite surface-exposed epitopes. *Infect. Immun.* **59**:3340–3342.
- Rodriguez, S. D., G. M. Buening, T. J. Green, and C. A. Carson. 1983. Cloning of *Babesia bovis* by in vitro cultivation. *Infect. Immun.* **42**:15–18.
- Shkap, V., E. Pipano, T. F. McElwain, U. Herzberg, Y. Krigel, L. Fish, and Palmer G. H. 1994. Cross-protective immunity induced by *Babesia bovis* clones with antigenically unrelated variable merozoite surface antigens. *Vet. Immunol. Immunopathology* **41**:367–374.
- Stich, R. W., Rice-A. C. Ficht, W. Tuo, and W. C. Brown. 1999. *Babesia bovis*: common protein fractions recognized by oligoclonal B. bovis-specific CD4+ T cell lines from genetically diverse cattle. *Exp. Parasitol.* **91**:40–51.
- Suarez, C. E., Florin-Christensen, M., S. A. Hines, G. H. Palmer, W. C. Brown, and T. F. McElwain. 2000. Characterization of allelic variation in the *Babesia bovis* merozoite surface antigen 1 (MSA-1) locus and identification of a cross-reactive inhibition-sensitive MSA-1 epitope. *Infect. Immun.* **68**:6865–6870.
- Volkman, S. K., D. L. Hartl, D. F. Wirth, K. M. Nielsen, M. Choi, S. Batalov, Y. Zhou, D. Plouffe, Le K. G. Roch, R. Abagyan, and E. A. Winzeler. 2002. Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*. *Science* **298**:216–218.

Editor: W. A. Petri, Jr.