Molecular Cloning of the Gene for a Conserved Major Immunoreactive 28-Kilodalton Protein of *Ehrlichia canis*: a Potential Serodiagnostic Antigen

JERE W. MCBRIDE, XUE-JIE YU, AND DAVID H. WALKER*

Department of Pathology and WHO Collaborating Center for Tropical Diseases, University of Texas Medical Branch, Galveston, Texas 77555-0609

A gene encoding a 28-kDa protein of *Ehrlichia canis* was cloned, sequenced, and expressed, and a comparative molecular analysis with homologous genes of *E. canis*, *Cowdria ruminantium*, and *Ehrlichia chaffeensis* was performed. The complete gene has an 834-bp open reading frame encoding a protein of 278 amino acids with a predicted molecular mass of 30.5 kDa. An N-terminal signal sequence was identified, suggesting that the protein undergoes posttranslational modification to a mature 27.7-kDa protein (P28). The *E. canis p28* gene has significant nucleic acid and amino acid sequence homologies with the *E. chaffeensis* outer membrane protein-1 (*omp-1*) gene family, with the *Cowdria ruminantium map-1* gene, and with other *E. canis* 28-kDa-protein genes. Southern blotting revealed the presence of at least two additional homologous *p28* gene copies in the *E. canis* genome, confirming that *p28* is a member of a polymorphic multiple-gene family. Amino acid sequence analysis revealed that *E. canis* P28 has four variable regions, and it shares similar surface-exposed regions, antigenicity, and T-cell motifs with *E. chaffeensis* P28. The *p28* genes from seven different *E. canis* isolates were identical, indicating that the gene for this major immunoreactive protein is highly conserved. In addition, reactivity of sera from clinical cases of canine ehrlichiosis with the recombinant P28 demonstrated that the recombinant protein may be a reliable serodiagnostic antigen.

Canine ehrlichiosis, also known as canine tropical pancytopenia, is a tick-borne rickettsial disease of dogs that was first described in Africa in 1935 and in the United States in 1963 (9, 10). The disease received more attention and recognition after an epizootic outbreak occurred in U.S. military dogs during the Vietnam War (29). The etiologic agent of canine ehrlichiosis is *Ehrlichia canis*, a small, gram-negative, obligate intracellular bacterium that exhibits tropism for mononuclear phagocytes (18) and is transmitted by the brown dog tick, *Rhipicephalus sanguineus* (11). The progression of canine ehrlichiosis occurs in three phases, acute, subclinical, and chronic. The acute phase is characterized by fever, anorexia, depression, lymphadenopathy, and mild thrombocytopenia (27). Dogs typically recover from the acute phase but become persistently infected carriers of the organism without clinical signs of disease for months or even years (12). A chronic phase characterized by thrombocytopenia, hyperglobulinemia, anorexia, emaciation, and hemorrhage, particularly epistaxis, followed by death develops in some cases (27).

Molecular taxonomic analysis based on the 16S rRNA gene has determined that *E. canis* and *Ehrlichia chaffeensis*, the etiologic agent of human monocytotropic ehrlichiosis, are closely related (2, 3, 6, 8). Considerable cross-reactivity of the 64-, 47-, 40-, 30-, 29-, and 23-kDa antigens of *E. canis* and *E. chaffeensis* has been reported (5, 6, 22, 23). Analysis of immunoreactive antigens with human and canine convalescent-phase sera by immunoblotting has resulted in the identification of immunodominant proteins of *E. canis*, including a 29-kDa protein (5). In addition, a 30-kDa protein of *E. canis* has been described as a major immunodominant antigen recognized early in the immune response and is antigenically distinct from the 30-kDa protein of *E. chaffeensis* (22, 23). Other immunodominant proteins of *E. canis* with molecular masses ranging from 20 to 30 kDa have also been identified (4–6, 17).

Recently, cloning and sequencing of a multigene family (*omp-1*) encoding proteins of 23 to 28 kDa have been described for *E. chaffeensis* (19). The gene (*p28*) for the 28-kDa immunodominant outer membrane protein of *E. chaffeensis*, which is homologous to the *Cowdria ruminantium map-1* gene, was cloned, and mice immunized with recombinant P28 were protected against challenge infection with the homologous strain based on PCR analysis of peripheral blood 5 days after challenge (19). Molecular cloning of two similar, but nonidentical, tandemly arranged *E. canis* 28-kDa-protein genes homologous to the *E. chaffeensis omp-1* gene family and the *C. ruminantium map-1* gene has also been reported (21).

In this study, we describe the molecular cloning, sequencing, characterization, and expression of the gene (designated *p28*) for a conserved mature 28-kDa immunoreactive protein of *E. canis* and the presence of a *p28* polymorphic multigene family in *E. canis*. Comparison with *E. chaffeensis* and other *E. canis* 28-kDa-protein genes revealed that this gene has the most amino acid homology with the *E. chaffeensis omp-1* multigene family. *E. canis* P28 is a highly conserved major immunodominant protein, and reactivity of sera from clinical canine ehrlichiosis cases with recombinant P28 suggests that the recombinant P28 may be a reliable serodiagnostic antigen.

MATERIALS AND METHODS

Ehrlichiae and purification. The *E. canis* Florida strain and isolates Demon, DJ, Jake, and Fuzzy were kindly provided by Edward Breitschwerdt, (College of Veterinary Medicine, North Carolina State University, Raleigh). The *E. canis* Louisiana strain was kindly provided by Richard E. Corstvet (School of Veterinary Medicine, Louisiana State University, Baton Rouge), and the *E. canis* Oklahoma strain was kindly provided by Jacqueline Dawson (Centers for Disease Control and Prevention, Atlanta, Ga.). Propagation of ehrlichiae was per-

* Corresponding author. Mailing address: Department of Pathology, University of Texas Medical Branch, 301 University Blvd., Galveston, TX 77555-0609. Phone: (409) 772-2856. Fax: (409) 772-2500. E-mail: dwalker@utmb.edu.

formed in DH82 cells with Dulbecco modified Eagle medium supplemented with 10% bovine calf serum and 2 mM L-glutamine at 37°C. The intracellular growth in DH82 cells was monitored by the presence of *E. canis* morulae by using general cytologic staining methods. Cells were harvested when 100% of the cells were infected with ehrlichiae and were then pelleted in a centrifuge at 17,000 × *g* for 20 min. Cell pellets were disrupted with a Braun-Sonic 2000 sonicator twice at 40 W for 30 s on ice. Ehrlichiae were purified as described previously (30). The lysate was loaded onto discontinuous gradients of 42, 36, and 30% Renografin and centrifuged at 80,000 × *g* for 1 h. Heavy and light bands containing ehrlichiae were collected, washed with sucrose-phosphate-glutamate buffer (218 mM sucrose, 3.8 mM KH_2PO_4, 7.2 mM K_2HPO_4, 4.9 mM glutamate, pH 7.0), and pelleted by centrifugation.

Nucleic acid preparation. *E. canis* genomic DNA was prepared by resuspending the Renografin-purified ehrlichiae in 600 µl of 10 mM Tris-HCl buffer (pH 7.5) with 1% (wt/vol) sodium dodecyl sulfate (SDS) and 100 ng of proteinase K per ml as described previously (15). This mixture was incubated for 1 h at 56°C, and the nucleic acids were extracted twice with phenol-chloroform-isoamyl alcohol (24:24:1). DNA was pelleted by absolute ethanol precipitation, washed once with 70% ethanol, dried, and resuspended in 10 mM Tris (pH 7.5). Plasmid DNA was purified by using a High Pure Plasmid Isolation Kit (Boehringer Mannheim, Indianapolis, Ind.), and PCR products were purified by using a QIAquick PCR Purification Kit (Qiagen, Santa Clarita, Calif.).

PCR amplification of the *E. canis* *p28* gene. Regions of the *E. canis* *p28* gene selected for PCR amplification were chosen based on homology (>90%) observed in the consensus sequence generated from Jotun-Hein algorithm alignment of the *E. chaffeensis* *p28* and *C. ruminantium* *map-1* genes. Forward primer 793 (5'-GCAGGAGCTGTTGGTTACTC-3') and reverse primer 1330 (5'-CCT TCCTCCAAGTTCTATGCC-3') corresponded to nucleotides 313 to 332 and 823 to 843, respectively, of *C. ruminantium* *map-1* and to nucleotides 307 to 326 and 814 to 834, respectively, of *E. chaffeensis* *p28*. *E. canis* DNA (from a North Carolina isolate, Jake) was amplified with primers 793 and 1330 with a thermal cycling profile of 95°C for 2 min and 30 cycles of 95°C for 30 s, 62°C for 1 min, and 72°C for 2 min, followed by a 72°C extension for 10 min and a 4°C hold. PCR products were analyzed in 1% agarose gels. This amplified PCR product was sequenced directly with primers 793 and 1330.

Sequencing of unknown 5′ and 3′ regions of the *p28* gene. The full-length sequence of *E. canis* *p28* was determined by using a Universal GenomeWalker Kit (Clontech, Palo Alto, Calif.) according to the protocol supplied by the manufacturer. Genomic *E. canis* DNA (Jake isolate) was digested completely with five restriction enzymes (*Dra*I, *Eco*RV, *Pvu*II, *Sca*I, and *Stu*I) which produce blunt-ended DNA. An adapter (AP1) supplied in the kit was ligated to each end of the *E. canis* DNA. The genomic libraries were used as templates to find the unknown DNA sequence of the *p28* gene by PCR with a primer complementary to a known portion of the *p28* sequence and a primer specific for the adapter AP1. Primers specific for *p28*, used for genome walking, were designed from the known DNA sequence derived from PCR amplification of *E. canis* *p28* with primers 793 and 1330. Primers 394 (5'-GCATTTCCACAGGATCATAGGTA A-3'; nucleotides 687 to 710) and 394C (5'-TTACCTATGATCCTGTGGAAA TGC-3'; nucleotides 710 to 687) were used in conjunction with supplied primer AP1 to amplify the unknown 5′ and 3′ regions of the *p28* gene by PCR. A PCR product corresponding to the 5′ region of the *p28* gene amplified with primers 394C and AP1 (2,000 bp) was unidirectionally sequenced with primer 793C (5'-GAGTAACCAACAGCTCCTGC-3'). A PCR product corresponding to the 3′ region of the *p28* gene amplified with primers 394 and AP1 (580 bp) was bidirectionally sequenced with the same primers. Noncoding regions on the 5′ and 3′ regions adjacent to the open reading frame were sequenced, and primers EC28OM-F (5'-TCTACTTTGCACTTCCACTATTGT-3') and EC28OM-R (5'-ATTCTTTTGCCACTATTTTTCTTT-3') complementary to these regions were designed in order to amplify the entire *p28* gene.

DNA sequencing. DNA was sequenced with an ABI Prism 377 DNA sequencer (Perkin-Elmer Applied Biosystems, Foster City, Calif.).

Sequencing of *p28* genes from *E. canis* isolates. The entire *p28* gene from seven *E. canis* isolates (four from North Carolina and one each from Oklahoma, Florida, and Louisiana) were amplified by PCR with primers EC28OM-F and EC28OM-R with a thermal cycling profile of 95°C for 5 min; 30 cycles of 95°C for 30 s, 62°C for 1 min, and 72°C for 2 min; and a 72°C extension for 10 min. The resulting PCR products were bidirectionally sequenced with the same primers.

Cloning and expression of *E. canis* *p28*. The entire *E. canis* *p28* gene was PCR amplified with primers EC28OM-F and EC28OM-R and cloned into the pCR2.1-TOPO TA cloning vector to obtain the desired set of restriction enzyme cleavage sites (Invitrogen, Carlsbad, Calif.). The insert was excised from pCR2.1-TOPO with *Bst*XI and ligated into the pcDNA 3.1 eukaryotic expression vector (Invitrogen) (designated pcDNA3.1/EC28 for subsequent studies). The pcDNA3.1/EC28 plasmid was amplified, and the gene was excised with a *Kpn*I-*Xba*I double digestion and directionally ligated into the pThioHis prokaryotic expression vector (Invitrogen). The clone (designated pThioHis/EC28) produced a recombinant thioredoxin fusion protein in *Escherichia coli* BL21. The recombinant fusion protein was crudely purified in the insoluble phase by centrifugation. The control thioredoxin fusion protein was purified from soluble cell lysates under native conditions by using nickel-nitrilotriacetic acid spin columns (Qiagen).

Serodiagnosis. The recombinant *E. canis* P28 fusion protein was subjected to SDS-polyacrylamide gel electrophoresis on 4 to 15% Tris-HCl gradient gels (Bio-Rad, Hercules, Calif.) and transferred to pure nitrocellulose (Schleicher & Schuell, Keene, N.H.) by using a semidry transfer cell (Bio-Rad). The membrane was incubated with convalescent-phase antisera (1:50) from six *E. canis*-infected dogs for 1 h, washed, and then incubated with an anti-canine immunoglobulin G (heavy plus light chains) alkaline phosphatase-conjugated affinity-purified secondary antibody (Kirkegaard & Perry Laboratories, Gaithersburg, Md.) at 1:1,000 for 1 h. Bound antibody was visualized with 5-bromo-4-chloro-3-indolyl phosphate–nitroblue tetrazolium substrate (Kirkegaard & Perry Laboratories).

Southern blot analysis. To determine if multiple genes homologous to the *p28* gene were present in the *E. canis* genome, a genomic Southern blot analysis was performed by a standard procedure (25). *E. canis* genomic DNA was digested completely with each of the restriction enzymes *Ban*II, *Eco*RV, *Hae*II, *Kpn*I, and *Spe*I, which do not cut within the *p28* gene, and with *Ase*I, which digests *p28* at nucleotides 34, 43, and 656. The probe was produced by PCR amplification with primers EC28OM-F and EC28OM-R and digoxigenin (DIG)-labeled deoxynucleotide triphosphates (Boehringer Mannheim) and digested with *Ase*I. The digested probe (566 bp) was separated by agarose gel electrophoresis, gel purified, and then used for hybridization. The completely digested genomic *E. canis* DNA was electrophoresed, transferred to a nylon membrane (Boehringer Mannheim), and hybridized at 40°C for 16 h with the *p28* gene DIG-labeled probe in DIG Easy Hyb buffer according to the protocol of the manufacturer (Boehringer Mannheim). Bound probe was detected with an anti-DIG alkaline phosphatase-conjugated antibody and a luminescent substrate (Boehringer Mannheim) and exposed to BioMax scientific imaging film (Eastman Kodak, Rochester, N.Y.).

Sequence analysis. *E. chaffeensis* *p28* and *C. ruminantium* *map-1* DNA sequences were obtained from the National Center for Biotechnology Information (16a). Nucleotide and deduced amino acid sequence analyses and protein and phylogenetic analyses were performed with LASERGENE software (DNASTAR, Inc., Madison, Wis.). Analysis of post-translational processing was performed by the method of McGeoch (16) and von Heijne (28) for signal sequence recognition with the PSORT program (20a).

Nucleotide sequence accession numbers. The GenBank accession numbers for the nucleic acid and amino acid sequences of the *E. canis* *p28* genes described in this report are as follows: Jake, AF082744; Louisiana, AF082745; Oklahoma, AF082746; Demon, AF082747; DJ, AF082748; Fuzzy, AF082749; and Florida, AF082750.

RESULTS

PCR amplification, cloning, sequencing, and expression of the *E. canis* *p28* gene. Alignment of nucleic acid sequences from *E. chaffeensis* *p28* and *C. ruminantium* *map-1* by using the Jotun-Hein algorithm produced a consensus sequence with regions of high homology (>90%). These homologous regions (nucleotides 313 to 332 and 823 to 843 of *C. ruminantium* *map-1* and 307 to 326 and 814 to 834 of *E. chaffeensis* *p28*) were targeted as primer annealing sites for PCR amplification. PCR amplification of the *E. canis* and *E. chaffeensis* *p28* genes was accomplished with primers 793 and 1330, resulting in a 518-bp PCR product. The nucleic acid sequence of the *E. canis* PCR product was obtained by sequencing the product directly with primers 793 and 1330. Analysis of the sequence revealed an open reading frame encoding a protein of 170 amino acids, and comparison of the 518-bp sequence obtained from PCR amplification of *E. canis* with the DNA sequence of the *E. chaffeensis* *p28* gene revealed homology greater than 70%. Adapter PCR with primers 394 and 793C was performed to obtain the 5′ and 3′ segments of the sequence of the entire gene. Primer 394 produced four PCR products (3, 2, 1, and 0.8 kb), and the 0.8-bp product was sequenced bidirectionally with primers 394 and AP1. The deduced sequence overlapped with the 3′ end of the 518-bp product, extending the open reading frame 12 bp to a termination codon. An additional 625 bp of noncoding sequence at the 3′ end of the *p28* gene was also sequenced. Primer 394C was used to amplify the 5′ end of the *p28* gene with supplied primer AP1. Amplification with these primers resulted in three PCR products (3.3, 3, and 2 kb). The 2-kb fragment was sequenced unidirectionally with primer 793C. The sequence provided the putative start codon of the *p28* gene and completed the 834-bp open reading frame encoding a protein of 278 amino acids. An additional 144 bp of

```
   1 ATTTTATTTATTACCAATCTTATATAATATATTAAATTTCTCTTACAAAAATCTCTAATG  60
  61 TTTTATACCTAATATATATATTCTGGCTTGTATCTACTTTGCACTTCCACTATTGTTAAT 120
 121 TTATTTTCACTATTTTAGGTGTAATATGAATTGCAAAAAAATTCTTATAACAACTGCATT 180
                                 M   N   C   K   K   I   L   I   T   T   A   L

 181 AATATCATTAATGTACTCTATTCCAAGCATATCTTTTTCTGATACTATACAAGATGGTAA 240
      I   S   L   M   Y   S   I   P   S   I   S   F   S   D   T   I   Q   D   G   N

 241 CATGGGTGGTAACTTCTATATTAGTGGAAAGTATGTACCAAGTGTCTCACATTTTGGTAG 300
      M   G   G   N   F   Y   I   S   G   K   Y   V   P   S   V   S   H   F   G   S

 301 CTTCTCAGCTAAAGAAGAAAGCAAATCAACTGTTGGAGTTTTTGGATTAAAACATGATTG 360
      F   S   A   K   E   E   S   K   S   T   V   G   V   F   G   L   K   H   D   W

 361 GGATGGAAGTCCAATACTTAAGAATAAACACGCTGACTTTACTGTTCCAAACTATTCGTT 420
      D   G   S   P   I   L   K   N   K   H   A   D   F   T   V   P   N   Y   S   F

 421 CAGATACGAGAACAATCCATTTCTAGGGTTTGCAGGAGCTATCGGTTACTCAATGGGTGG 480
      R   Y   E   N   N   P   F   L   G   F   A   G   A   I   G   Y   S   M   G   G

 481 CCCAAGAATAGAATTCGAAATATCTTATGAAGCATTCGACGTAAAAAGTCCTAATATCAA 540
      P   R   I   E   F   E   I   S   Y   E   A   F   D   V   K   S   P   N   I   N

 541 TTATCAAAATGACGCGCACAGGTACTGCGCTCTATCTCATCACACATCGGCAGCCATGGA 600
      Y   Q   N   D   A   H   R   Y   C   A   L   S   H   H   T   S   A   A   M   E

 601 AGCTGATAAATTTGTCTTCTTAAAAAACGAAGGGTTAATTGACATATCACTTGCAATAAA 660
      A   D   K   F   V   F   L   K   N   E   G   L   I   D   I   S   L   A   I   N

 661 TGCATGTTATGATATAATAAATGACAAAGTACCTGTTTCTCCTTATATATGCGCAGGTAT 720
      A   C   Y   D   I   I   N   D   K   V   P   V   S   P   Y   I   C   A   G   I

 721 TGGTACTGATTTGATTTCTATGTTTGAAGCTACAAGTCCTAAAATTTCCTACCAAGGAAA 780
      G   T   D   L   I   S   M   F   E   A   T   S   P   K   I   S   Y   Q   G   K

 781 ACTGGGCATTAGTTACTCTATTAATCCGGAAACCTCTGTTTTCATCGGTGGGCATTTCCA 840
      L   G   I   S   Y   S   I   N   P   E   T   S   V   F   I   G   G   H   F   H

 841 CAGGATCATAGGTAATGAGTTTAGAGATATTCCTGCAATAGTACCTAGTAACTCAACTAC 900
      R   I   I   G   N   E   F   R   D   I   P   A   I   V   P   S   N   S   T   T

 901 AATAAGTGGACCACAATTTGCAACAGTAACACTAAATGTGTGTCACTTTGGTTTAGAACT 960
      I   S   G   P   Q   F   A   T   V   T   L   N   V   C   H   F   G   L   E   L

 961 TGGAGGAAGATTTAACTTCTAATTTTATTGTTGCCACATATTAAAAATGATCTAAACTTG 1020
      G   G   R   F   N   F
1021 TTTTTAWTATTGCTACATACAAAAAAAGAAAAATAGTGGCAAAAGAATGTAGCAATAAGA 1080
1081 GGGGGGGGGGGGGACCAAATTTATCTTCTATGCTTCCCAAGTTTTTTTCYCGCTATTTATGA 1140
1141 CTTAAACAACAGAAGGTAATATCCTCACGGAAAACTTATCTTCAAATATTTTATTTATTA 1200
1201 CCAATCTTATATAATATATTAAATTTCTCTTACAAAAATCACTAGTATTTTATAGCAAAA 1260
1261 TATATATTCTGACTTGCTTTTCTTCTGCACTTCTACTATTTTTAATTTATTTGTCACTAT 1320
1321 TAGGTTATAATAAWATGAATTGCMAAAGATTTTTCATAGCAAGTGCATTGATATCACTAA 1380
1381 TGTCTTTCTTACCTAGCGTATCTTTTTCTGAATCAATACATGAAGATAATATAAATGGTA 1440
1441 ACTTTTACATTAGTGCAAAGTATATGCCAAGTGCCTCACACTTTGGCGTATTTTCAGTTA 1500
1501 AAGAAGAGAAAAACACAACAACTGGAGTTTTCGGATTAAAACAAGATTGGGACGGAGCAA 1560
1561 CACTAAAGGATGCAAGCWGCAGCCACACAWTAGACCCAAGTACAATG              1607
```

FIG. 1. Nucleic acid sequence of the *E. canis p28* gene, including adjacent 5′ and 3′ noncoding sequences. The ATG start codon and TAA termination codon are shown in boldface, and the 23-amino-acid leader signal sequence is underlined.
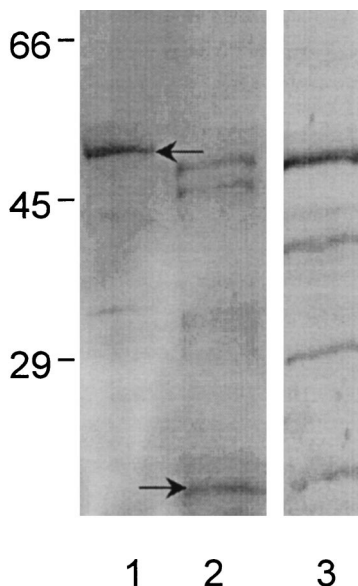
FIG. 2. SDS-polyacrylamide gel electrophoresis of expressed 50-kDa recombinant *E. canis* P28-thioredoxin fusion protein (lane 1, arrow) and 16-kDa thioredoxin control (lane 2, arrow) and corresponding immunoblot of recombinant *E. canis* P28-thioredoxin fusion protein recognized by convalescent-phase *E. canis* canine antiserum (lane 3). The thioredoxin control antigen did not react with the *E. canis* antiserum (not shown). Numbers on the left are molecular masses in kilodaltons.

readable sequence in the 5′ noncoding region of the *p28* gene was generated. Primers EC28OM-F and EC28OM-R were designed from complementary noncoding regions adjacent to the *p28* gene. The PCR product amplified with these primers was sequenced directly with the same primers. The complete DNA sequence for the *E. canis p28* gene is shown in Fig. 1. The *p28* PCR fragment amplified with these primers contained the entire open reading frame and sequence encoding 17 additional amino acids from the 5′ noncoding primer region. The gene was directionally subcloned into the pThioHis expression vector, and *E. coli* (BL21) was transformed with this construct. The expressed P28-thioredoxin fusion protein was insoluble. The expressed protein had an additional 114 amino acids associated with the thioredoxin, 5 amino acids for the enterokinase recognition site, and 32 amino acids from the multiple cloning site and 5′ noncoding primer region at the N terminus. Convalescent-phase antiserum from an *E. canis*-infected dog recognized the expressed recombinant fusion protein but did not react with the thioredoxin control (Fig. 2).

Nucleic acid sequence homology. The nucleic acid sequences of *E. canis p28* (834 bp) and the *E. chaffeensis omp-1* family of genes (*p28* and *omp-1A, -1B, -1C, -1D, -1E,* and *-1F*), including signal sequences, were aligned by using the Clustal method to examine homology between these genes (alignment not shown). Nucleic acid homology was equally conserved (68.9%) between *E. canis p28* and *E. chaffeensis p28* and *omp-1F*. Other putative outer membrane protein genes in the *E. chaffeensis omp-1* family, i.e., *omp-1D* (68.2%), *omp-1E* (66.7%), and
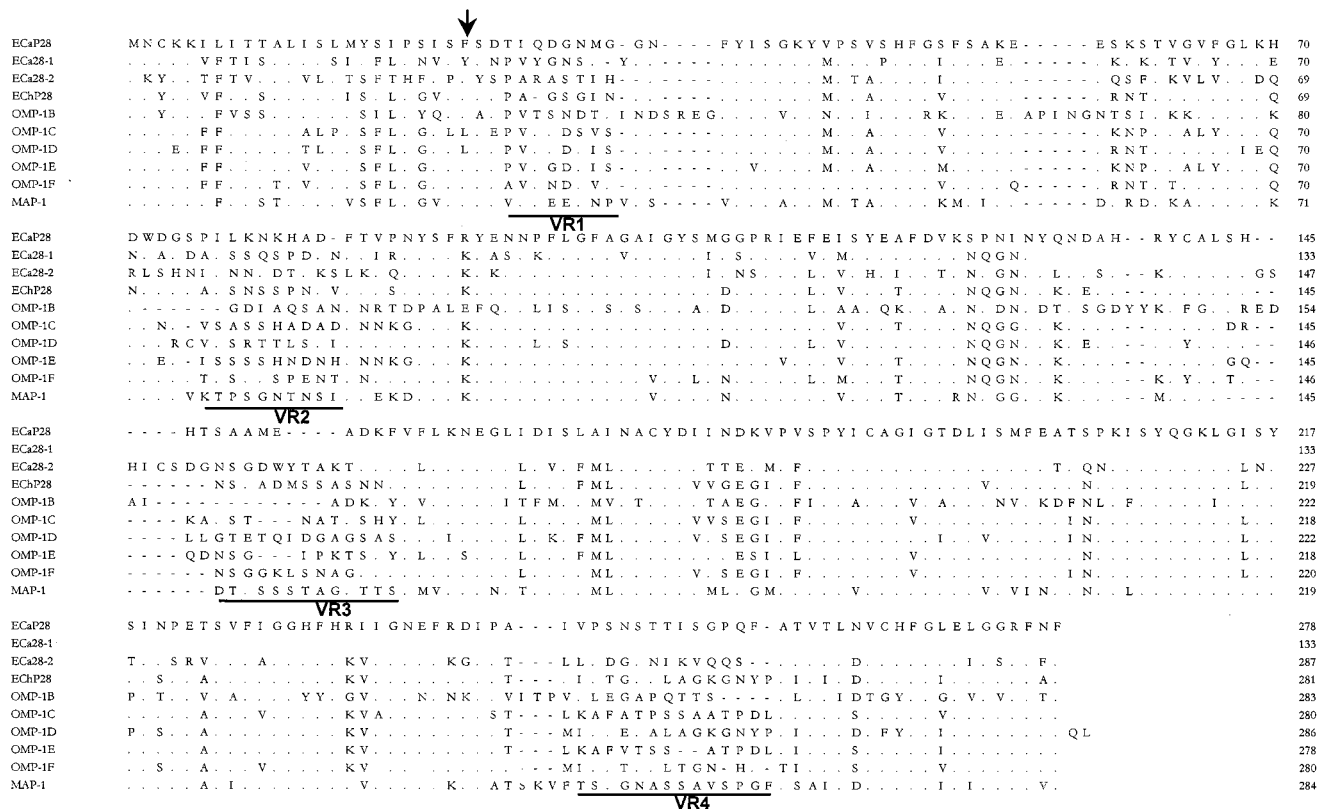

FIG. 3. Alignment of *E. canis* P28 (ECaP28), *E. canis* 28-kDa protein 1 (ECa28-1) (complete) and 28-kDa protein 2 (ECa28-2) (partial), *E. chaffeensis* (ECh) OMP-1 family, and *C. ruminantium* MAP-1 amino acid sequences. The *E. canis* P28 amino acid sequence is presented as the consensus sequence. Amino acids not shown are identical to those of *E. canis* P28 and are represented by dots. Divergent amino acids are shown with the corresponding one-letter abbreviations. Gaps introduced for maximal alignment of the amino acid sequences are denoted with dashes. Variable regions are underlined and denoted VR1, VR2, VR3, and VR4. The arrow indicates the predicted signal peptidase cleavage site for the signal peptide.
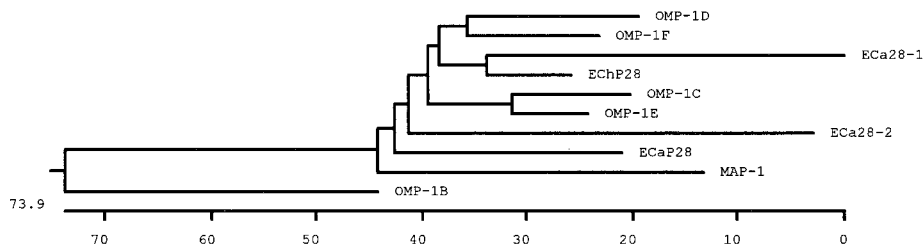
FIG. 4. Phylogenetic relatedness of *E. canis* P28 (ECaP28) with the *E. canis* 28-kDa protein 1 (ECa28-1) (complete) and 28-kDa protein 2 (ECa28-2) (partial), six members of the *E. chaffeensis* (ECh) OMP-1 family, and *C. rumanintium* MAP-1 from deduced amino acid sequences by utilizing unbalanced tree construction. The length of each pair of branches represents the distance between the amino acid sequences of the pairs. The scale measures the distance between sequences.

omp-1C (64.1%), *C. ruminantium map-1* (61.8%), and the *E. canis* 28-kDa-protein 1 gene (60%) and 28-kDa-protein 2 gene (partial) (59.5%), were also homologous to *E. canis p28*. *E. chaffeensis omp-1B* had the least nucleic acid homology (45.1%) with *E. canis p28*.

Amino acid sequence homology. Alignment of the predicted amino acid sequences of *E. canis* P28 and *E. chaffeensis* P28 revealed amino acid substitutions resulting in four variable regions. Substitutions or deletions in the amino acid sequence and the locations of variable regions of *E. canis* P28 and the *E. chaffeensis* OMP-1 family were identified (Fig. 3). Amino acid comparison demonstrated that the *E. canis* P28 protein had the most homology with OMP-1F (68%) of the *E. chaffeensis* OMP-1 family, followed by *E. chaffeensis* P28 (65.5%), OMP-1E (65.1%), OMP-1D (62.9%), and OMP-1C (62.9%), *C. ruminantium* MAP-1 (59.4%), *E. canis* 28-kDa protein 1 (55.6%) and 28-kDa protein 2 (partial) (53.6%), and *E. chaffeensis* OMP-1B (43.2%). The phylogenetic relationships based on amino acid sequences show that *E. canis* P28 and *C.*

ruminantium MAP-1, *E. chaffeensis* OMP-1 proteins, and *E. canis* 28-kDa proteins 1 and 2 (partial) are related (Fig. 4).

N-terminal signal sequence. The amino acid sequence analysis revealed that the entire *E. canis* P28 has a deduced molecular mass of 30.5 kDa. The protein has a predicted N-terminal signal peptide of 23 amino acids (MNCKKILITTAL ISLMYSIPSIS) (Fig. 3), which is similar to that predicted for *E. chaffeensis* P28 (MNYKKILITSALISLISSLPGVSFS) and the OMP-1 protein family (19, 31). A preferred cleavage site for signal peptidases (SIS; Ser-X-Ser) (20) is found at amino acids 21, 22, and 23. An additional putative cleavage site at amino acid position 25 (MNCKKILITTALISLMYSIPSISSFS) identical to the predicted cleavage site of *E. chaffeensis* P28 (SFS) was also present and would result in a mature *E. canis* P28 with a predicted molecular mass of 27.7 kDa. The signal cleavage site of the previously reported *E. canis* 28-kDa protein 1 is predicted to be at amino acid 30. However, signal sequence analysis predicted that *E. canis* 28-kDa protein 2 had an uncleavable signal sequence.

Detection of homologous genomic copies of the *E. canis p28* gene. Genomic Southern blot analysis of *E. canis* DNA was performed following complete independent digestions with restriction enzymes *Ban*II, *Eco*RV, *Hae*II, *Kpn*I, and *Spe*I, which do not have restriction endonuclease sites in the *p28* gene. In addition, digestion with *Ase*I, which has internal restriction endonuclease sites at nucleotides 34, 43, and 656, revealed the presence of at least three homologous *p28* gene copies (Fig. 5). Although *E. canis p28* has internal *Ase*I restriction sites, the DIG-labeled probe used in the hybridization experiment targeted a region of the gene within a single DNA fragment generated by the *Ase*I digestion of the gene. Digestion of genomic DNA of *E. canis* with *Ase*I produced three bands (approximately 566 bp, 850 bp, and 3 kb) that hybridized with the *p28* DNA probe, indicating the presence of multiple genes homologous to *p28* genes in the genome. Digestion with *Eco*RV and *Spe*I produced two bands that hybridized with the *p28* gene probe.

Predicted surface probability and immunoreactivity. Analysis of *E. canis* P28 by hydropathy and hydrophilicity profiles predicted surface-exposed regions on P28 (Fig. 6). Eight major surface-exposed regions consisting of three to nine amino acids were identified on *E. canis* P28 and were similar to the profile of surface-exposed regions on *E. chaffeensis* P28 (Fig. 6). Five of the larger surface-exposed regions on *E. canis* P28 were located in the N-terminal region of the protein. Surface-exposed hydrophilic regions were found in all four of the variable regions of *E. canis* P28. Ten T-cell motifs in *E. canis* P28 were predicted by using the Rothbard-Taylor algorithm (24), and high antigenicity of P28 was predicted by the Jameson-Wolf antigenicity algorithm (Fig. 6) (13). Similarities in antigenicity
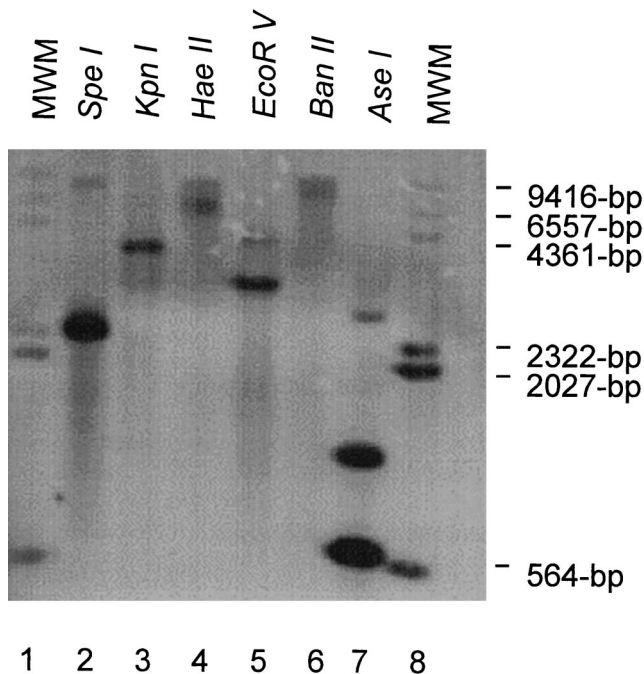

FIG. 5. Southern blot analysis of *E. canis* genomic DNA completely digested with six individual restriction enzymes and hybridized with a P28 DIG-labeled probe (lanes 2 to 7). Lanes 1 and 8, DIG-labeled molecular size markers (MWM).
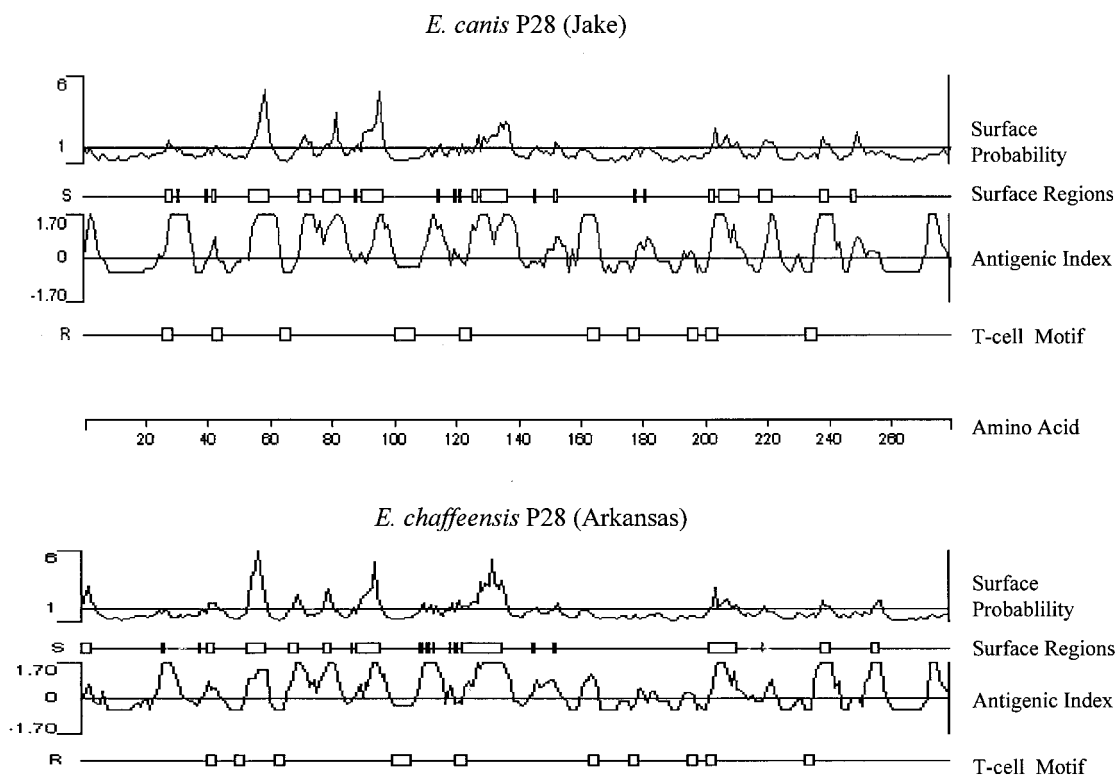
FIG. 6. Comparison of predicted protein characteristics of *E. canis* P28 (Jake strain) and *E. chaffeensis* P28 (Arkansas strain). Surface probability predicts the surface residues by using a window of hexapeptide. A surface residue is any residue with >2.0 nm^2 of water-accessible surface area. A hexapeptide with a value higher than 1 was considered a surface region. The antigenic index predicts potential antigenic determinants. The regions with values above zero are potential antigenic determinants. T-cell motif locates the potential T-cell antigenic determinants by using a motif of five amino acids, with residue 1 glycine or polar, residue 2 hydrophobic, residue 3 hydrophobic, residue 4 hydrophobic or proline, and residue 5 polar or glycine. The scale indicates amino acid positions.

and T-cell motifs were observed between *E. canis* P28 and *E. chaffeensis* P28.

Homology of *p28* gene sequences from different *E. canis* isolates. The *p28* genes from seven *E. canis* isolates, four from North Carolina and one each from Florida, Oklahoma, and Louisiana, were amplified by PCR with primers EC28OM-F and EC28OM-R and sequenced directly with the same primers. Alignment of the *p28* gene nucleic acid sequences revealed that the *p28* genes from these isolates were identical.

Serodiagnosis. Sera from six clinical cases of canine ehrlichiosis were incubated with the recombinant protein in immunoblots at a 1:100 dilution. Antibodies in the sera of five dogs (83%) reacted with the P28 recombinant protein (Fig. 7).

DISCUSSION

Proteins with similar molecular masses have been identified and cloned from multiple rickettsial agents, including *E. canis*, *E. chaffeensis*, and *C. ruminantium* (14, 19, 21). In this report, we demonstrated the cloning, expression, and characterization of a gene encoding a mature 28-kDa protein of *E. canis* that is homologous to the *omp-1* multiple-gene family of *E. chaffeensis* and the *C. ruminantium map-1* gene. The *E. canis p28* gene is also homologous, but different from the previously reported *E. canis* 28-kDa-protein 1 gene (complete) and 28-kDa-protein 2 gene (partial) (21). Previous studies have identified a 30-kDa protein of *E. canis* that reacts with convalescent-phase antisera against *E. chaffeensis*, but this protein was believed to be antigenically distinct (22). Our findings, based on comparison of amino acid substitutions in four variable regions of *E. canis*

P28, support this possibility. Together these findings also suggest that the amino acids responsible for the antigenic differences between *E. canis* and *E. chaffeensis* P28 are located in these variable regions and are readily accessible to the immune system. Reddy et al. (21) reported that immunoreactive peptides were located in the variable regions of the 28-kDa proteins of *C. ruminantium*, *E. chaffeensis*, and *E. canis*. Analysis of *E. canis* P28 and *E. chaffeensis* P28 revealed that all of the variable regions have predicted surface-exposed amino acids. A study with dogs demonstrated a lack of cross-protection between *E. canis* and *E. chaffeensis* (7). This observation may be related to antigenic differences in the variable regions of P28 as well as in other immunologically important antigens of these ehrlichial species. Another study found that convalescent-phase human antisera from *E. chaffeensis*-infected patients recognized a 29- or 28-kDa protein(s) of *E. chaffeensis* and also reacted with homologous proteins of *E. canis* (5). Homologous and cross-reactive epitopes on *E. canis* P28 and *E. chaffeensis* P28 appear to be recognized by the immune system.

Several reports have demonstrated that the 30-kDa antigen of *E. canis* exhibits strong immunoreactivity (22, 23). Antibodies in convalescent-phase antisera from humans and dogs have consistently reacted with proteins in this size range from *E. chaffeensis* and *E. canis*, suggesting that they may be important serodiagnostic as well as immunoprotective antigens (5, 6, 22). In addition, antibodies to 30-, 24-, and 21-kDa proteins develop early in the immune response to *E. canis* (22, 23), suggesting that these proteins may be especially important in the immune response during the acute stage of disease and thus
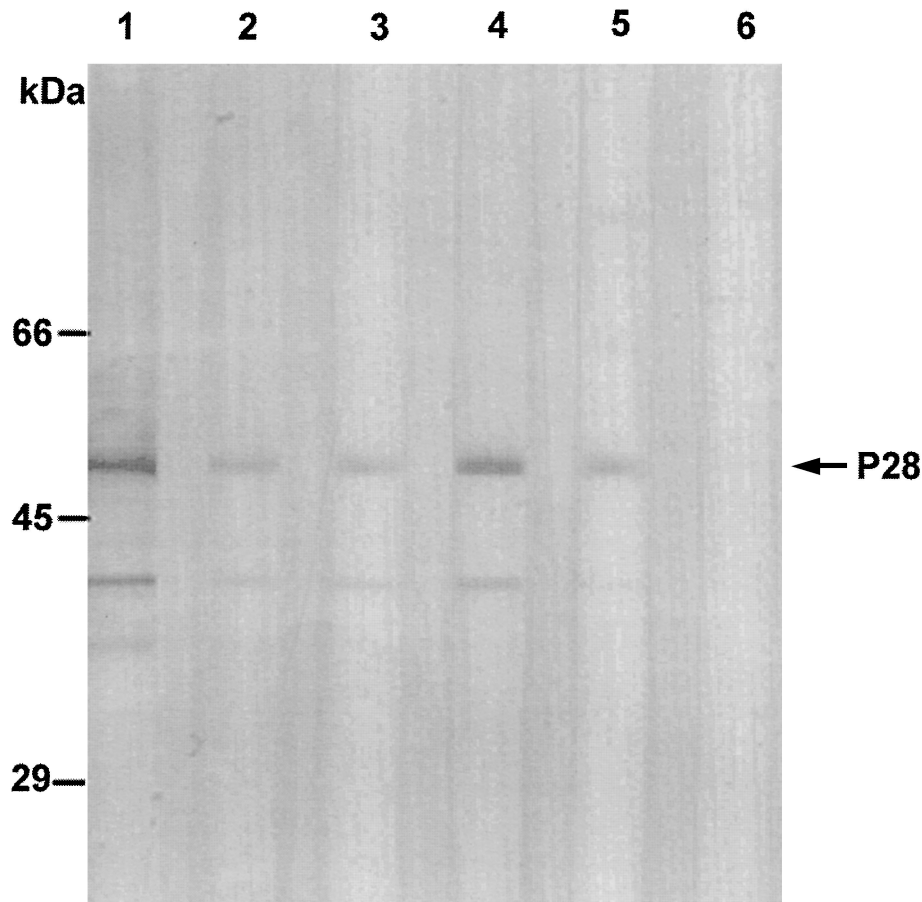
FIG. 7. Immunoblot analysis of sera from six clinical cases of canine ehrlichiosis (lanes 1 to 6) incubated with the 50-kDa recombinant P28 fusion protein (arrow). Smaller reactive bands, which may be degradative products of the P28 fusion protein, are also visible.

may be particularly useful for serodiagnosis. In addition, a family of homologous genes encoding outer membrane proteins with molecular masses of 28 kDa have been identified in *E. chaffeensis*, and mice immunized with recombinant *E. chaffeensis* P28 appeared to have developed immunity against homologous challenge (19). The P28 of *E. chaffeensis* has been demonstrated to be present in the outer membrane, and immunoelectron microscopy has localized the P28 on the surface of the organism, thus suggesting that it may serve as an adhesin (19). It is likely that the P28 of *E. canis* identified in this study has a similar location and function. The immunoprotective capacity of *E. canis* P28 is not known, but similar studies with *E. chaffeensis* P28 suggest that it may be a potential vaccine candidate.

There is evidence that the P28 from *E. canis* may be posttranslationally processed from an immature 30-kDa protein to a mature 28-kDa protein. Recently, a signal sequence was identified on *E. chaffeensis* P28 (31), and N-terminal amino acid sequencing has verified that the protein is posttranslationally processed, resulting in cleavage of the signal sequence to produce a mature protein (19). Sequences of OMP-1F and OMP-1E have also been proposed as leader signal peptides (19). Signal sequences identified on *E. chaffeensis* OMP-1F, OMP-1E, and P28 are homologous to the leader sequence of *E. canis* P28. However, two N-terminal signal sequences were identified on *E. canis* P28 within a 5-amino-acid region (SISFS). The first signal sequence produces a leader peptide

two amino acids shorter than that observed on the *E. chaffeensis* P28, due to a single amino acid substitution (serine) at position 21. The second signal sequence is identical to those on *E. chaffeensis* P28, OMP-1F, and OMP-1E and produces a leader peptide consisting of 25 amino acids. The homologies of the 25-amino-acid leader signal peptides of *E. chaffeensis* OMP-1F, OMP-1E, and P28 to *E. canis* P28 are 72, 68, and 64%, respectively. N-terminal amino acid sequencing could verify the cleavage site of the signal sequence of *E. canis* P28, but it is likely that the P28 *E. canis* protein that we have cloned is subject to posttranslational modification similar to that observed with *E. chaffeensis* P28.

Comparison of the *p28* genes from different strains of *E. canis* revealed that the gene is apparently completely conserved. Studies involving *E. chaffeensis* have demonstrated immunologic and molecular evidence of diversity in the *p28* gene. Patients infected with *E. chaffeensis* have variable immunoreactivity to the 29- and 28-kDa proteins, suggesting that there is antigenic diversity (5), and recent molecular evidence has been generated to support antigenic diversity in the *p28* gene from *E. chaffeensis* (31). However, differences in the host response to *E. chaffeensis* P28 may also explain some of the observed immunologic variability. A comparison of the *p28* genes of five *E. chaffeensis* isolates revealed that two isolates (Sapulpa and St. Vincent) were 100% identical but that three others (Arkansas, Jax, and 91HE17) were divergent by as much as 13.4% at the amino acid level. The conservation of *E. canis p28*

suggests that *E. canis* strains found in the United States may be genetically identical, and thus *E. canis p28* is an attractive vaccine candidate for canine ehrlichiosis in the United States. Further analysis of *E. canis* isolates outside the United States may provide information regarding the origin and evolution of *E. canis*. The documented immunoreactivity and conservation of the P28 protein suggests that it may be a reliable serodiagnostic antigen, and this proposal is further supported by the high rate of reactivity of clinical canine ehrlichiosis specimens with P28 in our study.

The presence of multiple polymorphic genes homologous to *E. canis* P28 corresponds to the presence of similar multiple-gene families in *E. chaffeensis* and *Anaplasma marginale* (1, 19). Six genes were found in the *omp-1* gene family of *E. chaffeensis*, and an *msp-3* multiple-gene family has been described for *A. marginale*. In our study, Southern blot hybridization of *E. canis* genomic DNA (Jake strain) digested with *Ase*I and hybridized with a DIG-labeled *p28* probe revealed the presence of at least three gene copies that were homologous to the *p28* gene. The restriction enzyme *Ase*I cuts within the *p28* gene; however, the *p28* probe was designed to be complementary with sequences internal to the *Ase*I restriction sites. In addition, *Ase*I cuts within the noncoding region found between the tandemly arranged *E. canis* 28-kDa-protein genes described previously (21). Thus, the three *p28* genes would be found on separate DNA fragments. The largest fragment from the *Ase*I digest (3 kb) that hybridized with the *p28* probe is at least three times larger than the *p28* gene. Therefore, the possibility of additional genes within this 3-kb fragment that are homologous to *p28*, and different from those already reported, cannot be eliminated. The hybridization pattern does suggest that all *p28* gene copies may be tandemly arranged along a single stretch of DNA. The role of multiple homologous genes is not known at this point; however, persistence of *E. canis* infections in dogs could conceivably be related to antigenic variation due to variable expression of homologous *p28* genes, thus enabling *E. canis* to evade immune surveillance. Variation of *msp-3* genes in *A. marginale* is partially responsible for variation in the MSP-3 protein, resulting in persistent infections (1). In addition, temperature-related gene expression resulting in phenotypic changes in *Borrelia hermsii* has also been reported (26). Studies to examine *p28* gene expression by *E. canis* in acutely and chronically infected dogs would provide insight into the role of the *p28* gene family in persistent *E. canis* infections.

REFERENCES

1. **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.
2. **Anderson, B. E., J. E. Dawson, D. C. Jones, and K. H. Wilson.** 1991. *Ehrlichia chaffeensis*, a new species associated with human ehrlichiosis. J. Clin. Microbiol. **29:**2838–2842.
3. **Anderson, B. E., C. E. Greene, D. C. Jones, and J. E. Dawson.** 1992. *Ehrlichia ewingii* sp. nov., the etiologic agent of canine granulocytic ehrlichiosis. Int. J. Syst. Bacteriol. **42:**299–302.
4. **Brouqui, P., J. S. Dumler, D. Raoult, and D. H. Walker.** 1992. Antigenic characterization of ehrlichiae: protein immunoblotting of *Ehrlichia canis*, *Ehrlichia sennetsu*, and *Ehrlichia risticii*. J. Clin. Microbiol. **30:**1062–1066.
5. **Chen, S. M., L. C. Cullman, and D. H. Walker.** 1997. Western immunoblotting analysis of the antibody responses of patients with human monocytotropic ehrlichiosis to different strains of *Ehrlichia chaffeensis* and *Ehrlichia canis*. Clin. Diagn. Lab. Immunol. **4:**731–735.
6. **Chen, S. M., J. S. Dumler, H.-M. Feng, and D. H. Walker.** 1994. Identification of the antigenic constituents of *Ehrlichia chaffeensis*. Am. J. Trop. Med. Hyg. **50:**52–58.
7. **Dawson, J. E., and S. A. Ewing.** 1992. Susceptibility of dogs to infection with *Ehrlichia chaffeensis*, causative agent of human ehrlichiosis. Am. J. Vet. Res. **53:**1322–1327.
8. **Dawson, J. E., Y. Rikihisa, S. A. Ewing, and D. B. Fishbein.** 1991. Serologic diagnosis of human ehrlichiosis using two *Ehrlichia canis* isolates. J. Infect. Dis. **163:**564–567.
9. **Donatien, A., and F. Lestoquard.** 1935. Existance in Algerie d'une *Rickettsia* du chein. Bull. Soc. Pathol. Exot. **28:**418–419.
10. **Ewing, S. A.** 1963. Observations of leukocytic inclusion bodies from dogs infected with *Babesia canis*. J. Am. Vet. Med. Assoc. **143:**503–506.
11. **Groves, M. G., G. L. Dennis, H. L. Amyx, and D. L. Huxsoll.** 1975. Transmission of *Ehrlichia canis* to dogs by ticks (*Rhipicephalus sanguineus*). Am. J. Vet. Res. **36:**937–940.
12. **Harrus, S., T. Waner, I. Aizenberg, J. E. Foley, A. M. Poland, and H. Bark.** 1998. Amplification of ehrlichial DNA from dogs 34 months after infection with *Ehrlichia canis*. J. Clin. Microbiol. **36:**73–76.
13. **Jameson, B. A., and H. Wolf.** 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. CABIOS **4:**181–186.
14. **Jongejan, F., N. de Vries, J. Nieuwenhuijs, A. H. Van Vliet, and L. A. Wassink.** 1993. The immunodominant 32-kilodalton protein of *Cowdria ruminantium* is conserved within the genus *Ehrlichia*. Rev. Elev. Med. Vet. Pays. Trop. **46:**145–152.
15. **McBride, J. W., R. E. Corstvet, S. D. Gaunt, J. Chinsangaram, G. Y. Akita, and B. I. Osburn.** 1996. PCR detection of acute *Ehrlichia canis* infection in dogs. J. Vet. Diagn. Invest. **8:**441–447.
16. **McGeoch, D. J.** 1985. On the predictive recognition of signal peptide sequences. Virus Res. **3:**271–286.
16a. **National Center for Biotechnology Information.** 6 March 1998, posting date. Sequences. [Online.] http://www.ncbi.nlm.nih.gov/Entrez. [10 August 1998, last date accessed.]
17. **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.
18. **Nyindo, M. B., M. Ristic, D. L. Huxsoll, and A. R. Smith.** 1971. Tropical canine pancytopenia: in vitro cultivation of the causative agent—*Ehrlichia canis*. Am. J. Vet. Res. **32:**1651–1658.
19. **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.
20. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39:**648.
20a. **PSORT.** 5 June 1998, posting date. [Online.] PSORT program. http://www.imcb.osaka-u.ac.jp/nakai/form.htm. [16 July 1998, last date accessed.]
21. **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.
22. **Rikihisa, Y., S. A. Ewing, and J. C. Fox.** 1994. Western immunoblot analysis of *Ehrlichia chaffeensis*, *E. canis*, or *E. ewingii* infections in dogs and humans. J. Clin. Microbiol. **32:**2107–2112.
23. **Rikihisa, Y., S. A. Ewing, J. C. Fox, A. G. Siregar, F. H. Pasaribu, and M. B. Malole.** 1992. Analysis of *Ehrlichia canis* and a canine granulocytic *Ehrlichia* infection. J. Clin. Microbiol. **30:**143–148.
24. **Rothbard, J. B., and W. R. Taylor.** 1988. A sequence pattern common to T cell epitopes. EMBO J. **7:**93–100.
25. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed., p. 9.34–9.58. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. **Schwan, T. G., and J. Hinnebusch.** 1998. Bloodstream-versus tick-associated variants of a relapsing fever bacterium. Science **280:**1938–1940.
27. **Troy, G. C., and S. D. Forrester.** 1990. Canine ehrlichiosis, p. 404–414. *In* C. E. Green (ed.), Infectious diseases of the dog and cat. W. B. Saunders Co., Philadelphia, Pa.
28. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids. Res. **14:**4683–4690.
29. **Walker, J. S., J. D. Rundquist, R. Taylor, B. L. Wilson, M. R. Andrews, J. Barck, A. L. J. Hogge, D. L. Huxsoll, P. K. Hildebrandt, and R. M. Nims.** 1970. Clinical and clinicopathologic findings in tropical canine pancytopenia. J. Am. Vet. Med. Assoc. **157:**43–55.
30. **Weiss, E., J. C. Coolbaugh, and J. C. Williams.** 1975. Separation of viable *Rickettsia typhi* from yolk sac and L cell host components by renograffin density gradient centrifugation. Appl. Microbiol. **30:**456–463.
31. **Yu, X. J., J. W. McBride, and D. H. Walker.** 1999. Genetic diversity of the 28-kilodalton outer membrane protein gene in human isolates of *Ehrlichia chaffeensis*. J. Clin. Microbiol. **37:**1137–1143.