Genetic and Antigenic Diversity of Major Immunoreactive Proteins in
Globally Distributed *Ehrlichia canis* Strains

Xiaofeng Zhang¹, Tian Luo¹, Avi Keysary⁵, Gad Baneth⁶, Simone Miyashiro¹,
Carmela Strenger⁵, Trevor Waner⁵, and Jere W. McBride¹-⁴

¹Departments of Pathology and ²Microbiology and Immunology, ³Center for Biodefense and
Emerging Infectious Diseases, ⁴Institute for Human Infections and Immunity
University of Texas Medical Branch, Galveston, Texas
⁵Israel Institute for Biological Research, Ness Ziona, Israel
⁶School of Veterinary Medicine, Hebrew University, Rehovot, Israel

**Running title:** *E. canis* diversity

Corresponding Author:
Jere W. McBride, Ph.D.
Department of Pathology
Center for Biodefense and Emerging Infectious Diseases
University of Texas Medical Branch
Galveston, Texas 77555-0609
Tel: 409-747-2498
Fax: 409-747-2455
Email: jemcbrid@utmb.edu
Abstract

The extent of knowledge regarding diversity of globally distributed *E. canis* has been limited to information gained from a few evolutionarily conserved genes. In this study, *E. canis* strains from the United States (US, Jake), Brazil (BR, São Paulo) and Israel (IS, 611 and IS-R, Ranana) were used to examine the antigenic and genetic diversity of four well characterized major immunoreactive protein genes/proteins. The gp36 and gp200 were the most divergent genes, and nucleotide substitutions in the gp36 tandem repeat region of the IS strain, but not IS-R strain, resulted in two amino acid differences (S → P; P → T) in each nine amino acid repeat (epitope-containing region). DNA sequences of gp19 and gp140 were completely conserved in the US and BR strains, but differences were found in the IS strains, including two fewer tandem repeats in the gp140 and a single amino acid substitution in the IS gp19. *E. canis* whole cell lysates from each isolate were examined by Western immunoblotting using sera from naturally infected dogs from each country, and four major immunoreactive proteins (gp19, gp36, gp140 and gp200) were identified in each strain using protein-specific antisera. The US and BR strains exhibited highly conserved immunoreactive protein profiles, while some differences were identified in the IS strain. Sera from naturally infected Israeli dogs confirmed gene sequencing information, which demonstrated two distinct *E. canis* strains, defined by the gp36 gene. Conversely, the gp19 was strongly reactive and present in all *E. canis* isolates. The gp140 and gp200 were also present in all strains, although the gp140 in the IS strain had two fewer tandem repeats and exhibited a smaller mass.
*Ehrlichia canis* is a globally distributed tick-transmitted obligately intracellular bacterium that is the primary etiologic agent of canine monocytic ehrlichiosis (CME), and has been identified as the cause of human ehrlichiosis in patients from Venezuela (38,39). Rickettsiosis in dogs caused by *E. canis* was first reported in 1935 in Algeria, and later reported in southern India and other parts of Africa the 1940’s (9,31). Subsequently, *E. canis* was relatively unrecognized until it was associated with outbreaks of canine tropical pancytopenia (TCP) in Singapore and Malaysia from 1963 to 1968 (51), and was identified as the cause of an epizootic of TCP in US military dogs stationed in Vietnam in late 1968 (17,36). *E. canis* infections have since been well documented in the United States, Israel, Brazil and Vietnam (1,3,12,16,20-22,36,49), and serologic and/or molecular evidence of infection has also been reported in temperate regions where *R. sanguineus* is commonly found including Central and South America, the Caribbean, parts of Africa, southern Europe and southeast Asia (2,5-8,15,18,19,23,32,33,41,42,44,50).

Development of globally useful serologic and molecular based diagnostics as well as effective vaccines for CME is dependent on understanding the genetic diversity of *E. canis*, particularly with respect to the major immunoreactive proteins. Molecular characterization of evolutionarily conserved genes, such as 16S rRNA have provided little information on strain diversity and suggest a high level of conservation (39,40,43,47,48). Similarly, the immunoreactive major outer membrane proteins p28/p30 in US and Venezuelan strains of *E. canis* appear to be highly conserved (13,29,30,46), an observation that was extended to *E. canis* characterized from six human patients from Venezuela (38). Other genes such as the thio-oxidoreductase gene (*dsb*) and *gltA* were also found to be conserved in geographically dispersed strains (23,32).
The genome of *E. canis* has been sequenced, and a small group of acidic tandem repeat and ankyrin repeat containing proteins associated with host-pathogen interactions were identified (24). Several of these proteins were previously considered major immunoreactive proteins and have been well studied, including the gp200, gp140, gp36 and gp19 (11,25,26,28,53). The *E. canis* gp36 is an acidic serine-rich protein that contains a major antibody epitope in the tandem repeat region (11). Examination of the gp36 gene in US, Brazilian and Cameroonian strains of *E. canis* identified variations in the number of tandem repeats and nucleic acid changes that resulted in four amino acid substitutions (10). However, the diversity of other major immunoreactive *E. canis* proteins in globally dispersed strains is not known. A homogenous pattern of proteins reacting with *E. canis* dog sera by immunoblot from the US, France, Israel and the Virgin Islands was previously reported (14). However, differences in protein reactivity were noted with sera collected from dogs from Italy and Zimbabwe, suggesting potential for diversity in the antigenic composition of *E. canis* strains in these countries (14).

The objective of this study was to determine the genetic and antigenic diversity of proteins subject to immune pressure in globally dispersed strains of *E. canis*. Four major immunoreactive protein genes (gp200, gp140, gp36 and gp19) were sequenced from each strain, and immunoblotting profiles to *E. canis* whole cell lysates compared. Strains from the United States and Brazil exhibited homogeneous immunoblotting patterns compared to the strain from Israel. Sequencing of four major immunoreactive protein genes demonstrated that US and BR strains were highly similar and strains from Israel were the more divergent.
Materials and Methods

Ehrlichia canis strains and propagation. *Ehrlichia canis* strains used in this study originated from the United States (Jake strain; US), Israel (611 strain; IS) and Brazil (São Paulo strain; BR). DNA was also obtained from an Israeli dog (Ranana; IS-R) naturally infected with *E. canis* for comparison. *E. canis* strains (US, IS and BR) were propagated in DH82 cells (canine histiocyte) with MEM (Gibco, Grand Island, NY) supplemented with 5% fetal bovine serum (Hyclone, Logan, Utah), 1% HEPES (Sigma Chemical Co, St. Louis, Mo), 1% sodium pyruvate (Sigma) and 1% non-essential amino acids (Sigma). The IS strain was provided to our laboratory in J774 cells (murine); however, cell free ehrlichiae from these cultures was used to infect DH82 cells for the antigen used in this study. Infected cells were harvested when morulae were observed in all cells. Cells were pelleted (5000 x g; 15 min) and resuspended in phosphate buffered saline (PBS) and sonicated twice (40 Hz) for 10 s and large cell debris pelleted by centrifugation (1500 x g; 10 min) at 4°C. The supernatant containing cell-free ehrlichiae was centrifuged (10,000 x g; 15 min) at 4°C. The pellet was then washed once in PBS and pelleted (10,000 x g; 15 min) at 4°C and resuspended in PBS. The suspension containing bacteria was frozen at -80°C and utilized as an antigen and DNA source. Protein concentration of purified *E. canis* antigen was determined using the BCA protein assay (Pierce Biotechnology, Rockford, IL).

PCR amplification and cloning of major immunoreactive protein genes. *E. canis* genomic DNA was extracted from purified antigen using a commercial kit according the manufacturer’s protocol (MasterPure Complete DNA and RNA purification kit; Epicentre, Madison, Wis.). The primers used for the amplification of *E. canis* genes (gp19, gp36, gp140, gp200) (Table 1) were
designed using primer design software (PrimerSelect; DNASTAR, Madison, Wis.) and *E. canis* genome sequence information (Integrated Microbial Genomes system; United States Department of Energy, Joint Genome Institute, Walnut Creek, Calif.).

The *E. canis* gp19 gene (Ecaj_0113) was amplified with primers (Table 1), which target intergenic regions (~50-bp upstream; ~9-bp downstream) flanking the gene. The gp19 gene was amplified by PCR using Hot Master Mix (Eppendorf, Westbury, NY) with the following thermal cycling protocol: 95°C for 4 min; 30 cycles at 95°C for 30 s, 47°C for 30 s, and 72°C for 1 min, followed by a 72°C extension for 7 min. The *E. canis* gp36 gene (Ecaj_0109) was amplified using primers (Table 1) that targeted highly conserved genes (upstream, Ecaj_0108, MerR transcriptional regulator; downstream, Ecaj_0110, tryptophanyl tRNA synthase) flanking the gp36 gene. PCR was performed as described for the gp19 gene, except Platinum Taq DNA Polymerase High Fidelity (Invitrogen; Carlsbad, Calif.) was used with an annealing temperature of 55°C and extension at 72°C for 3 min. The *E. canis* gp140 (Ecaj_0017) was amplified using primers (Table 1) located within the open reading frame (forward, bases 1 to 29; reverse, bases 2034 to 2061). PCR was performed with conditions described for the gp36 gene, except the annealing temperature of 57°C for 30 s and extension at 72°C for 1.5 min were used. The gp200 gene (Ecaj_0365) was amplified using primers (Table 1) targeting the intergenic region (~250-bp upstream; ~20-bp downstream) flanking the gene. The gene was amplified using conditions described for the gp36 gene, except an annealing temperature of 61°C and an extension step of 72°C for 5 min were used.

PCR amplicons for all four genes were separated and visualized by agarose gel electrophoresis (1.2% FlashGel DNA System; Lonza, Walkersville, Md.). The gp19 amplicon was purified using a purification kit (ExoSAP-IT, USB Corp, Cleveland, OH) sequenced directly.
using the same primers. All other PCR amplicons were cloned directly into universal TOPO TA sequencing vectors (Invitrogen), plasmids purified using a plasmid purification kit (High Pure Plasmid Isolation Kit, Roche, Indianapolis, IN) and sequenced using primers supplied with the vector.

**DNA sequencing.** PCR amplicons and plasmids were sequenced with an ABI Prism 377XL DNA sequencer (PerkinElmer Applied Biosystems, Foster City, CA) at the University of Texas Medical Branch Protein Chemistry Core Laboratory.

**Cloning and expression of recombinant IS and US gp36.** The IS-gp36 gene (810-bp) was amplified by PCR using forward primer (5’-ATG CTA TTT ATA CTA ATG GGT TAT TG-3’) and reverse primer (5’-CAG GGT AAG CTG AGT ATA TAA ATC-3’) with IS strain DNA as the template with the following thermal cycling protocol: 94°C for 30s, 55°C for 30s, and extension at 72°C for 1 min. The PCR product was cloned into pBAD/Thio Fusion Vector (Invitrogen) and recombinant proteins (IS and US) expressed and purified as previously described (11).

**Gel electrophoresis and Western immunoblotting.** Purified *E. canis* whole-cell lysates (5 µg/well) were solubilized in LDS sample buffer (NuPAGE; Invitrogen) and separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using Novex 4-12% Bis-Tris gels (15 well, NuPAGE; Invitrogen) and morpholinepropanesulfonic acid (MOPS) running buffer (NuPAGE, Invitrogen). Separated lysates were transferred to a pure nitrocellulose (Protran BA85, 0.45 µm pore size; Whatman, Florham Park, NJ) using a semi-dry transfer
apparatus (BioRad, Hercules, Calif). Anti- *E. canis* sera (US dog #02160, 1:2000; 04283, 1:1000; 20699, 1:1000; ZOC, 1:1000; Israeli dogs #53, #18, #17 and #1, 1:1000; Brazilian dogs #157, 1:500; #37, 1:1000; #42, 1:800; #45, 1:800) used for Western immunoblots were obtained from dogs naturally infected with *E. canis*. Rabbit anti-recombinant protein sera (gp19, 1:5000; gp36, 1:500; gp140, 1:200; and gp200, 1:100) were used to identify native *E. canis* immunoreactive proteins in whole cell lysates from each strain. Western immunoblotting was performed as previously described (25).

**Major immunoreactive protein-specific antisera.** Anti-sera specific for the gp200, gp140, gp36 and gp19 were produced in rabbits as previously described (11,27,28,33).

**Sequence analysis.** Nucleic acid and amino acid alignments (using Clustal W algorithm), percent identity, and phylogenetic relationships were performed with MegAlign (Lasergene v5.08; DNASTar, Madison, Wis).

**Nucleotide sequence accession numbers.** Gene sequences for genes sequenced in this study for *E. canis* (US, Jake strain) were previously available in Genbank (DQ085427, gp36; DQ858221, gp19; AF252298, gp200; AF112369, gp140). The *E. canis* (BR, São Paulo strain) gp36 gene was also available in Genbank (DQ146154). Accession numbers were assigned to genes sequenced in this study: *E. canis* BR, gp19 (EU118960), gp140 (EU118964) and gp200 (EF636664). The *E. canis* (IS, 611 strain) genes were assigned the following accession numbers: gp19 (EU118959), gp36 (EF636663), gp140 (EU118963), gp200 (EF636665). Gene sequences
amplified from the Israeli dog (IS-R) naturally infected with *E. canis* were assigned the following accession numbers: gp19 (EU118958), gp36 (EU118961) and gp140 (EU118962).

**Results**

**Diverse immunoreactive proteins (gp36 and gp200).** Major immunoreactive protein genes, gp36 and gp200 were amplified and sequenced from the three isolates (US, BR and IS) and one blood sample (gp36 only; gp200 could not be amplified from blood) from an Israeli dog (IS-R) naturally infected with *E. canis* (Figs. 1 and 2). The gp36 was the most divergent with amino acid identities ranging from 81.5 to 91.7. The *E. canis* gp36 (US strain) had 12 tandem repeats, but six additional repeats were found in the Brazilian strain, which decreased the overall percent identity (81.5%). However, only eight nucleotide differences (8 amino acid changes) were found between US and BR gp36 genes, four of those were located in the C-terminus (last 15 amino acids). Conversely, substantial divergence in IS gp36 gene was found. Two nucleotides that encode the nine amino acid repeat differed and resulted in two amino acid substitutions were noted only in the IS strain. The IS strain had 11 tandem repeats and the IS-R strain had 10. The repeat region of US, BR and IS-R were identical in sequence, but differed in the number of repeats (12, 18 and 10, respectively). The domain with the highest level of divergence in gp36 was found in the C-terminal region (last 18 amino acids). The IS strain exhibited the least identity (33%) with the US strain in this region, while BR and IS-R strains had higher identity with the US strain.

The gp200 from the US and BR strains exhibited a high amino acid identity (99.6%), and the IS strain had lower identity (94.3%). There were four amino acid differences between the US
and BR gp200, including one insertion (position 1360). The IS strain exhibited numerous amino acid changes with a higher frequency of changes (47%) found in a 325 amino acid stretch located in the C-terminal region (amino acids 950 to 1275) of the protein.

**Conserved immunoreactive protein genes (gp19 and gp140).** Major immunoreactive protein genes, gp19 and gp140, were amplified and sequenced from the three isolates (US, BR and IS) and one blood sample from an Israeli dog (IS-R) naturally infected with *E. canis*. The US and BR strains had identical gp19 and gp140 gene sequences. The gp19 gene of IS strain had one nucleotide substitution (position 104) that resulted in a single amino acid change (Glu to Gly at position 35) in the epitope-containing region that was previously reported (28). Interestingly, the gp19 gene amplified from the Israeli dog naturally infected with *E. canis* (IS-R) was identical to the US and BR strains. A high degree of overall nucleic acid and amino acid conservation was observed in the gp19 (99.99 to 100 identity) in geographically dispersed strains.

The gp140 gene of the IS strain had two fewer tandem repeats (12; 14 in US strain), but had nine nucleotide substitutions that resulted in nine amino acid changes, seven that were localized to two of the 12 tandem repeats (Fig. 3). The IS-R strain had a same number of tandem repeats as the IS strain, and had identical amino acid changes in seven locations. However, eight additional amino acid changes (in repeat region) unique to IS-R compared to IS were noted (Fig. 4). A high degree of nucleic acid and amino acid conservation in the gp140 (99.8 to 100% identity) was observed in geographically dispersed strains.

**Western immunoblotting.** *E. canis* whole cell lysates (US, BR, and IS) were reacted with sera obtained from dogs naturally infected with *E. canis* in each respective country (Fig. 4A). The
reactivity of all sera with each lysate (US, BR, IS) was relatively homogeneous, and most consistent in the masses >45 kDa and <100 kDa. The most notable differences were observed in known major immunoreactive proteins including gp36 and gp140. The Brazilian gp36 was substantially larger than the US strain due to six additional repeat units (Figs. 1 and 4A). Two types of immunoreactivity were consistently observed with Israeli dog sera. One type crossreacted with the US and Brazilian gp36 (Fig. 4A) and the other type reacted with a protein present in the IS strain that was of similar in size to the US strain gp36, and these sera did not cross react with US and BR gp36 (Fig. 4A). The gp140 of the Israeli strain was smaller (2 repeats less) than the US and BR strains (Fig. 3) and was strongly recognized in all strains with the Israeli sera, but the IS gp140 was weakly recognized by sera from Brazil and the USA (Fig. 4A). The gp19 exhibited the same mass in all strains and was strongly recognized by dog sera from each country (Fig. 4A). The gp200 exhibited a similar mass in all strains and immunoreactivity of the gp200 was more prominent in the BR and US whole cell lysates probed with USA sera (Fig. 4A). Another notable differences were an unknown protein (~90-95 kDa) in all strains that strongly reacted with sera from Israel among all strains, but the IS protein was weakly recognized by Brazilian and USA sera (Fig. 4A).

Anti-US strain gp36 antibody reacted strongly with BR and US strain gp36, but no reactivity was observed with the IS strain. Anti-US strain gp200 and gp19 antibodies reacted with a proteins of similar size in all strains (Fig 4B), and anti-US strain gp140 reacted with all strains although the gp140 in the IS strain was smaller in mass (Fig. 4B).

Three sera from dogs naturally infected with *E. canis* from Israel, Brazil and the USA were reacted with the respective *E. canis* whole cell lysates (IS, BR and US) to evaluate consistency of antigen recognition by different dogs from the same location. A similar antigen
reactivity pattern of the sera with each respective *E. canis* lysate was observed (Fig. 4C). Some
minor differences in antigen recognition were noted among dogs from the same region, but
immunodominant proteins were consistently recognized by all dogs. One exception worth
noting was observed with the sera from Israel. As noted above a protein similar in size to the
gp36 was strongly reactive with two of the Israeli dog sera (Israel Ab1), and two Israeli dog sera
(Islaw Ab2) reacted with the US and BR gp36 as shown in Figs. 4A and 5.

**Immunoreactivity of recombinant IS and US gp36.** The IS (611 strain) and US (Jake strain)
recombinant gp36 proteins were expressed and reacted with *E. canis* infected dog sera from
Israel, Brazil and the USA. The expressed recombinant IS-gp36 exhibited a molecular mass
slightly smaller (~53 kDa, including 16 kDa N- and C-terminal fusion tags) than the recombinant
US gp36. The recombinant IS gp36 reacted strongly with sera from an Israeli dog naturally
infected with *E. canis*, but not with the recombinant US-gp36. The US-gp36 reacted strongly
with sera from the USA and Brazil. Sera from Israeli dogs that recognized IS-gp36 did not cross
react with US-gp36. In contrast, Israeli sera that recognized US-gp36 did not recognize IS-gp36
(Fig. 5).
E. canis is the most widely dispersed Ehrlichia species, yet little information is available regarding the antigenic variability of the organism. Previous studies have indicated that some genes including 16 rRNA, dsb, and p28/p30 exhibit a high level of conservation in geographically dispersed strains (13,32,38-40,43,46). However, one serologic comparison has also suggested there is antigenic variability in geographically dispersed E. canis strains (14). We have examined, for the first time, the molecular diversity among four major immunoreactive proteins (gp200, gp140, gp36 and gp19) and compared the reactivity of three (US, IS, BR) E. canis whole cell lysates against homologous and heterologous sera from three continents. The major immunoreactive proteins of E. canis elicit a strong immune response, and thus the genes encoding these proteins may exhibit higher diversity as a result of increased selective immune pressure. Furthermore, two of these genes contain tandem repeats, and variations in the number and sequence of Ehrlichia tandem repeat containing proteins is well established. This is the first study in which three globally dispersed strains of E. canis were propagated in a single laboratory in order to closely compare the antigenic profile under the same conditions and to examine genetic differences in four newly characterized major immunoreactive protein genes in these strains. Previously, sera from various locations were reacted with a single E. canis antigen preparation (US) to gain some information regarding antigenic diversity (14). Future studies with a larger number of strains from more locations may provide additional insight into the diversity of E. canis, this study does provide important information with regard to E. canis in the three countries included in this investigation. Western immunoblotting of native E. canis lysates reacted with homologous and heterologous sera revealed that the immunoblot pattern of immunoreactive proteins is consistently homogeneous with regard to
protein mass and immunoreactivity, suggesting that most of these proteins are conserved among geographically separated strains. These findings are in agreement with a previous study which reported relatively homogenous immunoblot patterns using sera from different geographic locations (14). However, the former study was limited in that it compared antigen profiles (proteins <110 kDa) using a single *E. canis* (Florida strain) whole cell lysate preparation. Furthermore, consistent and strong recognition of *E. canis* antigens (>80 kDa) reactive by immunoblot in this study, were not consistently identified in the previous study using sera originating from the United States (14). This difference is likely due to protein blotting conditions resulting in more efficient transfer of high molecular mass proteins in our study. Furthermore, some notable differences in the molecular masses of immunodominant antigens were also identified in this study. The gp19 and gp36 were the most immunodominant antigens in both the US and Brazilian immunoblots, while proteins in the 28-30 kDa range, consistent with the mass of the major outer membrane protein (p28/p30) were present, but were less dominant. The immunodominant proteins identified this study are consistent with those identified in our previous study using the same US antigen, but with sera from experimentally infected dogs (25).

Antigens that were most visibly different among strains (molecular mass) were some of the proteins specifically examined in this study. The most divergent of the four examined in this study was the gp36, a secreted protein that elicits an early antibody response directed at the tandem repeat region, and differentially expressed on dense-cored ehrlichiae (11). Previous studies had reported differences in the number of gp36 tandem repeats in *E. canis* strains (10) as well as the ortholog in *E. ruminantium* (Erum1110) (4), a finding confirmed in this study. Interestingly, there were two types of sera from naturally infected dogs from Israel that were
identified based on reactivity to the gp36. The first type reacted strongly with the US and BR
gp36 and the second type reacted with a protein in the IS strain with similar size to the US gp36
and consistent with the size of the IS gp36 that was sequenced in this study. This difference can
be explained using the gp36 gene sequence information from the IS and IS-R strains. Some dogs
appear to be infected with a IS-R strain type, which has a gp36 identical in the antibody epitope
region as the US gp36 (11). In contrast, the IS strain, which was propagated in the laboratory
and used in the immunoblots has a divergent gp36, which has two amino acid substitutions (S →
P; P → T) in the epitope containing repeat region. Thus, the serologic response to gp36 in the IS
antigen preparation is dependent on the strain of *E. canis* infecting the dogs. Evaluation of four
Israeli sera from naturally infected dogs found half were specific for the IS gp36 and half
contained antibodies to the US gp36 (IS-R strain). Serologic and molecular evidence indicate
that there are at least two distinct *E. canis* strains circulating in Israel. This is in contrast to the
US and Brazil, which appear to be more conserved as has been previously reported
(10,11,29,30,53). Another interesting region of diversity among all strains was in the gp36
carboxy-terminal region immediately downstream of the repeat region. Antibody epitopes have
not been identified in this region (11), suggesting that this diversity is not a direct result of
humoral immune selection pressure. The divergence of gp36 among the strains examined in this
study supports the conclusion that this gene useful for molecular genotyping of *E. canis* strains.

The gp200, which is the largest major immunoreactive protein identified in *E. canis* had
high conservation between US and BR strains, but substantial diversity was found in the IS
strain. The gp200 is a secreted nuclear translocated ankyrin repeat containing protein that has
five major species-specific epitopes located primarily in terminal acidic domains (34,35). The
amino acid changes in the IS gp200 were distributed throughout the protein, but a higher
frequency of amino acid substitutions were noted in a carboxy-terminal 325 amino acid domain
of the protein. Amino acid substitutions were identified in known gp200 epitopes (35). The
carboxy-terminal and internal epitopes had at least two amino acid substitutions; however, only
one of the two known amino terminal epitopes had a single amino acid substitution. The IS
gp200 appeared to be less reactive with heterologous sera than with homologous sera, and these
substitutions in known epitopes are likely responsible for this difference. Conversely,
conservation of the N-terminal epitopes would result in recognition of the gp200 in all strains as
was demonstrated by immunoblot.

Two of the immunoreactive proteins (gp19 and gp140) examined in this study were
highly conserved. The gp140 has been previously shown to be identical in US strains (53);
however, the conservation of the gp140 outside the US has not been investigated. Consistent
with previous findings, the gp140 was found to be highly conserved among US, BR and IS
strains. The IS-R strain was the most divergent, with the most frequent substitution located in
the repeat region, where an asparagine was replaced by serine. This substitution was also
observed in the IS strain, but was more frequent in the IS-R strain. The fact that all of the
tandem repeats did not possess this substitution, suggests that is a point mutation that is
occurring as a result of immune pressure. The repeat region of the gp140 does contain a strong
antibody epitope in an area containing the substituted amino acid (McBride, unpublished data).
The most notable difference was found in the Israeli strains, which had two fewer tandem
repeats. This difference in tandem repeats was also evident by immunoblot, as the IS gp140
exhibited a smaller molecular mass. Variations in the number of tandem repeats has been
reported in *E. chaffeensis* gp120, the *E. canis* gp140 ortholog (45,52). However, *E. chaffeensis*
gp120 repeat variants have not been associated with differences in pathogenicity (37).
The *E. canis* gp19 is a recently characterized ortholog of *E. chaffeensis* VLPT that has a single major serine-rich epitope (28). In our previous study, we reported that the gp19 was highly conserved among strains from the US, Mexico, Brazil and Israel. However, single amino acid changes were noted in the Israeli and Mexican strains, both of which were located in the antibody epitope containing region (28). In this study, the IS-R strain was found to be identical to the US and BR strains, and the IS strain (611) was confirmed to have a single amino acid substitution at position 35. Although the IS strain has a single amino acid substitution in the epitope-containing region, it was still strongly recognized by homologous and heterologous antisera, suggesting that this change is not critical to epitope recognition. However, selective immune pressure may be responsible for these changes considering the location. The gp19 does elicit an early antibody response and is a dominant antigen on immunoblots. The conservation of this antigen in geographically dispersed *E. canis* strains suggests that this protein be useful for immunodiagnostics and vaccines that are widely applicable.

The development of reliable immunodiagnostics and vaccines for canine ehrlichiosis is dependent on understanding differences that may exist in geographically dispersed strains of *E. canis*, particularly with respect to these important genes. Further studies involving additional *E. canis* strains and more potentially important genes are needed to appreciate the full extent of global diversity of the organism and specific genes that have increased selection pressure. We specifically focused this study on genes with increased selection pressure, in order to provide additional insight into the diversity of *E. canis*. This information would also expand our knowledge with regard to the genetic variability in known targets of the host immune response and identify new and useful targets for genotyping *E. canis*. Based on the information generated in this study and others, *E. canis* appears to be more conserved than *E. chaffeensis* or *E.*
ruminantium, but variability in major immunoreactive proteins examined in this study indicates that substantial variability is present among E. canis strains. Furthermore, it is evident that variability in E. canis must be a consideration in developing widely applicable diagnostics and vaccines.
Acknowledgements

We thank Osnat Eyal for technical assistance and sequencing the IS-R gp19 gene and Xue-jie Yu and David H. Walker for critical review of the manuscript. This work was supported by the National Institutes of Health (AI 071145-01) and the Clayton Foundation for Research.
Reference List


Figure Legends

FIG. 1. Clustal alignment of the gp36 amino acid sequence of *E. canis* strains (US, BR, IS, and IS-R) from three continents. Boxed amino acids represent residues divergent from US sequence and a dash (-) represents a gap. Single tandem repeat unit is identified with a bar.

FIG. 2. Clustal alignment of the gp200 amino acid sequence of *E. canis* strains (US, BR, IS) from three continents. Boxed amino acids represent residues divergent from US sequence and a dash (-) represents a gap.

FIG. 3. Clustal alignment of the gp140 amino acid sequence of *E. canis* strains (US, BR, IS, and IS-R) from three continents. Boxed amino acids represent residues divergent from US sequence and a dash (-) represents a gap. Single tandem repeat unit is identified with a bar.

FIG. 4. (A) Western immunoblot of *E. canis* whole-cell lysates (BR, US, and IS) and uninfected DH82 cell lysates (Ctrl) probed with sera from dogs naturally infected with *E. canis* from each respective country (Israel #18 and #53, Brazil #157 and USA #02160); (B) *E. canis* whole-cell lysates (BR, IS and US) probed with rabbit anti-US gp200, gp140, gp36 and gp19 specific sera, respectively; (C) *E. canis* whole cell lysates from IS, BR and US reacted with three sera from dogs naturally infected with *E. canis* from USA (dog nos. 04283 [1], 20699 [2], and ZOC [3]), Brazil (dog nos. 37 [1], 42 [2], 45 [3]) and Israel (dog nos. 17 [1], 18 [2] and 1 [3]).
Fig. 5. Immunoreactivity of Israeli gp36 (lane 1) and US gp36 (lane 2) purified recombinant proteins probed with sera from dogs naturally infected with *E. canis* from Brazil, USA and Israel.
Table 1. Primers used for amplification of the US, Brazilian, and Israeli *E. canis* gp19, gp36, gp140 and gp200 genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer</th>
<th>Sequence (5’ to 3’)</th>
<th>Reverse primer</th>
<th>Sequence (5’ to 3’)</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p19</td>
<td>p19F</td>
<td>aagattgtgtgtgtatg</td>
<td>p19R</td>
<td>tttacgcttgcctgaat</td>
<td>492</td>
</tr>
<tr>
<td>p36</td>
<td>p36F-L</td>
<td>ggaatgattatataaagtttgc</td>
<td>p36R-L</td>
<td>gatcggtttgatggg</td>
<td>2080</td>
</tr>
<tr>
<td>p140</td>
<td>p140F</td>
<td>atggatatgtgaactataatgtgactac</td>
<td>p140R</td>
<td>tattaaatcaactgtttcttgtttagt</td>
<td>2061</td>
</tr>
<tr>
<td>p200</td>
<td>p200F</td>
<td>tttgccttcaaggaacatcg</td>
<td>p200R</td>
<td>tgcacetaatatcactagaaacac</td>
<td>4597</td>
</tr>
</tbody>
</table>