

## Sequence Analysis of the *msp4* Gene of *Anaplasma phagocytophilum* Strains

José de la Fuente,<sup>1,2\*</sup> Robert F. Massung,<sup>3</sup> Susan J. Wong,<sup>4</sup> Frederick K. Chu,<sup>5</sup> Hans Lutz,<sup>6</sup> Marina Meli,<sup>6</sup> Friederike D. von Loewenich,<sup>7</sup> Anna Grzeszczuk,<sup>8</sup> Alessandra Torina,<sup>9</sup> Santo Caracappa,<sup>9</sup> Atilio J. Mangold,<sup>10</sup> Victoria Naranjo,<sup>2</sup> Snorre Stuen,<sup>11</sup> and Katherine M. Kocan<sup>1</sup>

Department of Veterinary Pathobiology, College of Veterinary Medicine, Oklahoma State University, Stillwater, Oklahoma<sup>1</sup>; Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia<sup>3</sup>; Diagnostic Immunology Laboratory, Wadsworth Center,<sup>4</sup> and Wadsworth Center for Laboratories and Research,<sup>5</sup> New York State Department of Health, Albany, New York; Instituto de Investigación en Recursos Cinegéticos IREC (CSIC-UCLM-JCCM), Ciudad Real, Spain<sup>2</sup>; Clinical Laboratory, Faculty of Veterinary Medicine, University of Zurich, Zurich, Switzerland<sup>6</sup>; Institute of Medical Microbiology and Hygiene, Department of Medical Microbiology and Hygiene, University Clinic of Freiburg, Freiburg, Germany<sup>7</sup>; Department of Infectious Diseases, Medical University of Białystok, Białystok, Poland<sup>8</sup>; Istituto Zooprofilattico Sperimentale della Sicilia, Palermo, Italy<sup>9</sup>; Instituto Nacional de Tecnología Agropecuaria, Estación Experimental Agropecuaria Rafaela, Rafaela, Santa Fe, Argentina<sup>10</sup>; and Department of Production Animal Clinical Sciences, Norwegian School of Veterinary Science, Sandnes, Norway<sup>11</sup>

Received 22 January 2004/Returned for modification 1 July 2004/Accepted 7 November 2004

The causative agent of human granulocytic ehrlichiosis was recently reclassified as *Anaplasma phagocytophilum*, unifying previously described bacteria that cause disease in humans, horses, dogs, and ruminants. For the characterization of genetic heterogeneity in this species, the homologue of *Anaplasma marginale* major surface protein 4 gene (*msp4*) was identified, and the coding region was PCR amplified and sequenced from a variety of sources, including 50 samples from the United States, Germany, Poland, Norway, Italy, and Switzerland and 4 samples of *A. phagocytophilum*-like organisms obtained from white-tailed deer in the United States. Sequence variation between strains of *A. phagocytophilum* (90 to 100% identity at the nucleotide level and 92 to 100% similarity at the protein level) was higher than in *A. marginale*. Phylogenetic analyses of *msp4* sequences did not provide phylogeographic information but did differentiate strains of *A. phagocytophilum* obtained from ruminants from those obtained from humans, dogs, and horses. The sequence analysis of the recently discovered *A. phagocytophilum msp2* gene corroborated these results. The results reported here suggest that although *A. phagocytophilum*-like organisms from white-tailed deer may be closely related to *A. phagocytophilum*, they could be more diverse. These results suggest that *A. phagocytophilum* strains from ruminants could share some common characteristics, including reservoirs and pathogenicity, which may be different from strains that infect humans.

*Anaplasma phagocytophilum* (Rickettsiales: Anaplasmataceae), also known as the human granulocytic ehrlichiosis (HGE) agent, *Ehrlichia equi*, and *Ehrlichia phagocytophila*, is the causative agent of human granulocytic anaplasmosis, tick-borne fever of ruminants, and equine and canine granulocytic anaplasmosis (16). HGE was first described in 1994; since then, it has been reported in Europe, South America, and North America, where it is now the predominant form of anaplasmosis and among the most common tick-borne infections in the United States (28). In humans, the disease is characterized by fever, headache, myalgia, and malaise, leukopenia, thrombocytopenia, and evidence of hepatic injury (6). Tick-borne fever is reported as a febrile disease of goats, sheep, and cattle, with

clinical signs varying from undetectable illness to severe febrile disease associated with opportunistic infections, hemorrhage, and abortions (16). Equine and canine diseases are characterized by fever, depression, anorexia, leukopenia, and thrombocytopenia; equine infection also frequently results in limb edema and ataxia and may lead to opportunistic infections (16).

The genera *Anaplasma* and *Ehrlichia* contain obligate intracellular bacteria that multiply within parasitophorous vacuoles in the cytoplasm. Within these vacuoles, the organisms undergo a developmental cycle involving reticulated and dense forms (6). Members of the *Ixodes persulcatus* complex transmit *A. phagocytophilum*, which is maintained, at least in part, in small and medium-sized mammals such as the white-footed mouse (*Peromyscus leucopus*), the raccoon (*Procyon lotor*), and the gray squirrel (*Sciurus carolinensis*) (21, 44). Some evidence suggests that subclinical persistent infections occur in domestic and wild ruminants, including deer (16). In cell culture, *A.*

\* Corresponding author. Mailing address: Department of Veterinary Pathobiology, College of Veterinary Medicine, Oklahoma State University, Stillwater, OK 74078. Phone: (405) 744-0372. Fax: (405) 744-5275. E-mail: jose\_delafuente@yahoo.com.

TABLE 1. Samples used in the study<sup>a</sup>

Sample (n)	Origin (location, date)	Reference(s)
HZ	Human (New York)	Rikihisa et al. (38)
NY18	Human (New York); human HL-60 cells for this study	Asanovich et al. (2)
HGE2	Human (Minnesota); human HL-60 and tick ISE6 cells for this study	Goodman et al. (18)
Blood samples (21)	Human (New York, 1996–2002)	Barbet et al. (3) S. J. Wong and F. K. Chu, unpublished
WI-HI,-WI-H4 (4)	Human (Wisconsin)	Massung et al. (29)
WI-CI	Canine (Wisconsin)	Massung et al. (29)
MRK	Equine (California; previously <i>E. equi</i> ); human HL-60 for this study	Asanovich et al. (2)
Webster	Human (Wisconsin); human HL-60 for this study	Asanovich et al. (2)
WTD-OK1, WTD-OK4	<i>A. phagocytophilum</i> -like organisms; white-tailed deer (Oklahoma)	Dawson et al. (10)
WTD-UGA1, WTD-UGA2	<i>A. phagocytophilum</i> -like organisms; white-tailed deer (Georgia)	Dawson et al. (10)
AP-V1	White-tailed deer; <i>A. phagocytophilum</i> variant 1 (Maryland)	Massung et al. (25, 27, 29)
Elsa and Olga	Bovine (Switzerland, 2003)	H. Lutz and M. Meli, unpublished
Donkey 136, 137, 138 (3)	Donkey (Sicily and Italy, 2003)	J. de la Fuente and A. Torina, unpublished
Roe deer 2, 6, 8, 9 (4)	Roe deer (Germany, 2003)	F. von Loewenich, unpublished
Bison 7, 12, 15, 16, 18, 21 (6)	European bison (Poland, 2004)	F. von Loewenich and A. Grzeszczuk, unpublished
Horse 31	Equine (Germany, 1999)	von Loewenich et al. (48)
Ovine 5 and 7	Ovine (Norway, 2003)	Stuen et al. (41)
LL	Human (New York)	Lin et al. (22)
MN-2	Human (Minnesota)	Lin et al. (22)
OS	Ovine (Scotland, United Kingdom)	Lin et al. (22)

<sup>a</sup> The sequence of the *msp4* gene was characterized in all strains of *A. phagocytophilum* except for strains LL, MN-2, and OS, for which information was obtained from Lin et al. (22) for the analysis of *msp2* genes, together with the sequences derived from HZ, MRK, WI-H1-WI-H4, WI-C1, Webster, HGE2, NY18, Elsa, Olga, and Ovine 5 and 7 strains and New York blood samples.

*phagocytophilum* infects and replicates in undifferentiated HL-60 human promyelocytic cells; in HL-60 cells differentiated into neutrophil-like cells, potential precursors of the myelomonocytic lineage (18); and in two tick cell lines (IDE8 and ISE6) that were derived originally from embryos of *Ixodes scapularis* (30–32).

A major protein antigen(s) is expressed on the outer membrane of *A. phagocytophilum*, and some of the immunodominant major surface proteins (MSPs) share sequence similarity with *Anaplasma marginale* MSP2 and MSP4 (22) and *Ehrlichia ruminantium* MAP1 (46). Although the biological function of *A. marginale* MSP4 is unknown, this MSP is probably involved in host-pathogen interactions and may evolve more rapidly than other nuclear gene proteins because of selective pressures exerted by host immune systems. Furthermore, the analysis of *msp4* sequences provided phylogeographic patterns for *A. marginale* strains (11–14).

Although *A. phagocytophilum* has a broad geographic distribution, all strains identified thus far appear to have considerable serological cross-reactivity and a minor degree of variation in the nucleotide sequences of the 16S rRNA, *groESL*, *gltA*, *ank*, and *msp2* genes, with the exception of some *ank* sequences from infected German ticks that are different from other *ank* sequences of human and animal strains (16, 22, 25, 26, 42, 47). However, the clinical and host tropism diversity of *A. phagocytophilum* suggests the presence of genetic differences among these bacteria that have not been characterized. The present study was undertaken to characterize variations in the sequence of an outer membrane protein gene, homologous

to *A. marginale msp4*, in *A. phagocytophilum* strains obtained from different host species in the United States and Europe.

#### MATERIALS AND METHODS

***A. phagocytophilum* samples.** The *A. phagocytophilum* samples obtained from naturally infected humans, dogs, horses, donkeys, white-tailed deer (WTD), roe deer, sheep, European bison, and cows and from infected cultured human HL-60 and tick ISE6 cells (isolates) used in this study for the characterization of *msp4* sequences are listed in Table 1. The geographic distribution of the samples covered the western and eastern United States and Germany, Poland, Norway, Italy, and Switzerland in Europe. Four samples of *A. phagocytophilum*-like organisms from WTD were also included in the study (Table 1). *msp4* sequences for *Anaplasma centrale* (GenBank accession number AF428090), *Anaplasma ovis* (AF393742), and *A. marginale* (GenBank accession number AY012929) (37), the *E. ruminantium map1* (GenBank accession number AY028378) (1) and the *Wolbachia* spp. MSP (GenBank accession number AF020068) (5) sequences were also included in the analysis together with *Anaplasma* spp. *msp4* sequences.

***msp4*, *msp2*, and 16S rRNA PCR and sequencing.** A sequence of 283 amino acids with 58% identity (BLAST score, E value =  $1e^{-99}$ ) to *A. marginale* (Oklahoma isolate) MSP4 (GenBank accession number AAG34718) was identified in silico in the *A. phagocytophilum* (HZ isolate) genome (TIGR 212042\_contig:366) by TBLASTN search of the preliminary sequence data obtained from The Institute for Genomic Research (<http://www.tigr.org>) at the National Center for Biotechnology Information (GenBank accession number NC\_004351). Oligonucleotides (MAP4AP5 [5'-ATGAATTACAGAGAATTGCTTGTAGG-3'] and MSP4AP3[5'-TTAATTGAAAGCAAATCTTGCTCCTATG-3'], for upstream

TABLE 2. Percent identity among nucleotide (above diagonal formed by boldface entries) and percent similarity among deduced amino acid (below diagonal) sequences between *Ehrlichia* and *Anaplasma* species<sup>a</sup>

	Wol spp.	<i>E. r.</i>	<i>E. ch.</i>	<i>A. c.</i>	<i>A. o.</i>	<i>A. marginale</i>		<i>A. phagocytophilum</i>									
						Wetumka, Okla.	Mexico	HZ	WI-H	MRK	AP-V1	Elsa	Bison 12	Bison 15	Horse 31	Roe deer	Ovine
<i>Wolbachia</i> spp.	<b>100</b>	36	38	35	34	35	35	39	39	39	39	38	38	38	39	38	39
<i>E. ruminantium</i>	20	<b>100</b>	58	41	42	46	46	46	46	46	46	46	46	46	46	46	46
<i>E. chaffeensis</i>	27	44	<b>100</b>	47	49	49	48	52	52	52	52	52	52	52	52	53	53
<i>A. centrale</i>	26	33	32	<b>100</b>	85	84	85	64	64	64	63	63	63	63	63	63	63
<i>A. ovis</i>	26	32	32	89	<b>100</b>	92	92	66	66	66	66	66	66	66	66	64	65
<i>A. marginale</i> (Wetumka, Okla.)	26	32	32	89	97	<b>100</b>	99	66	65	66	65	65	65	65	65	64	65
<i>A. marginale</i> (Mexico) <sup>b</sup>	26	32	32	88	96	99	<b>100</b>	65	65	65	65	64	65	64	65	63	64
<i>A. phagocytophilum</i>																	
HZ <sup>c</sup>	26	28	31	62	62	62	62	<b>100</b>	100	100	99	98	99	98	99	90	97
WI-H <sup>d</sup>	25	28	31	61	62	61	61	100	<b>100</b>	100	99	98	98	98	98	90	97
MRK <sup>e</sup>	26	28	31	61	62	61	61	100	99	<b>100</b>	99	98	98	98	98	90	97
AP-V1	26	28	31	62	62	62	62	99	99	99	<b>100</b>	98	98	98	98	90	97
Elsa <sup>f</sup>	25	27	30	62	62	62	62	99	99	99	99	<b>100</b>	99	100	99	90	97
Bison 12 <sup>g</sup>	25	27	30	62	62	62	62	100	99	99	99	100	<b>100</b>	99	99	90	97
Bison 15 <sup>h</sup>	25	27	30	62	62	62	62	99	99	99	99	100	100	<b>100</b>	99	90	97
Horse 31	26	28	31	62	62	62	62	100	100	100	99	99	100	99	<b>100</b>	90	97
Roe deer <sup>i</sup>	25	28	30	61	61	61	61	93	92	92	93	92	92	92	93	<b>100</b>	91
Ovine <sup>j</sup>	26	28	31	62	62	62	62	99	99	99	99	98	99	98	99	93	<b>100</b>

<sup>a</sup> Sequences were aligned and percent identity and similarity were determined with the AlignX program. Abbreviations: *Wol*, *Wolbachia*; *E. r.*, *E. ruminantium*; *E. ch.*, *E. chaffeensis*; *A. c.*, *A. centrale*; *A. o.*, *A. ovis*.

<sup>b</sup> The *A. marginale* Mexico isolate was selected because its *msp4* sequence is among the most divergent from the Wetumka, Okla., isolate.

<sup>c</sup> HZ also represents HGE2, NY18, Webster, and WI-C1 and New York human blood samples.

<sup>d</sup> WI-H includes strains WI-H1 to WI-H4.

<sup>e</sup> MRK also represents strains Donkey 136, 137, and 138.

<sup>f</sup> Elsa also represents strains Olga and Bison 7.

<sup>g</sup> Bison 12 also represents strain Bison 18 and Bison 21.

<sup>h</sup> Bison 15 also represents strain Bison 16.

<sup>i</sup> Roe deer includes strains Roe deer 2, 6, 8, and 9.

<sup>j</sup> Ovine includes strains Ovine 5 and Ovine 7.

and downstream primers, respectively) were designed for amplification of the *A. phagocytophilum msp4* 849-bp coding region with the sequence of the HZ isolate. DNA was extracted as reported previously (3, 8, 13, 26). The *msp4* gene was amplified from 1 µl (0.1 to 10 ng) of DNA by PCR with 10 pmol of each primer (MSP4AP5 and MSP4AP3) in a 50-µl volume (1.5 mM MgSO<sub>4</sub>, 0.2 mM deoxynucleoside triphosphate, 1× AMV-Tfl 5× reaction buffer, 5 U of *Tfl* DNA polymerase) with the Access RT-PCR system (Promega, Madison, Wis.). Reactions were performed in an automated DNA thermal cycler (Mastecycler personal; Eppendorf, Westbury, N.Y.) for 35 cycles. After an initial denaturation step of 30 s at 94°C, each cycle consisted of a denaturing step of 30 s at 94°C and an annealing-extension step of 1 min at 68°C. The program ended by storing the reaction mixtures at 4°C. PCR products were electrophoresed on 1% agarose-ethidium bromide gels to check the size of amplified fragments in comparison with DNA size markers. Amplified fragments were resin purified (Wizard; Promega) and cloned into pGEM-T vector (Promega) or used directly for sequencing both strands by double-stranded dye-termination cycle sequencing (Core Sequencing Facility, Department of Biochemistry and Molecular Biology, Noble Research Center, Oklahoma State University). When cloned, at least two clones were sequenced from each PCR.

Amplification and sequencing of *msp2* genes was done as described above for *msp4* genes but with the oligonucleotide primers described by Lin et al. (22) for the *msp2* full-length DNA PCR (MSP25 [5'-TTATGATTAGGCCTTTGGGCATG-3'] and MSP23 [5'-TCAGAAAGATACACGTGCC-3']) and amplification cycles consisting of a denaturing step of 30 s at 94°C, an annealing step of 30 s at 60°C, and an extension step of 1.5 min at 68°C.

A nested PCR for the amplification of a 546-bp amplicon of the 16S rRNA gene was done as described by Massung et al. (29).

**Sequence alignment and phylogenetic analysis.** The *msp4* and *msp2* gene coding regions and the MSP4 protein sequences were used for sequence alignment and phylogenetic analysis. Multiple sequence alignment was performed with the program AlignX (Vector NTI Suite, version 8.0; InforMax, North Bethesda, Md.) with an engine based on the CLUSTAL W algorithm (45). Nucleotides were coded as unordered, discrete characters with five possible character states: A, C, G, T, or N. Gaps were coded as missing data. Maximum parsimony (MP) analyses were conducted with equal weights for all characters and substitutions and heuristic

searches with 10 random additions of input taxa. To examine the effect of the method of analysis on the resulting phylogeny, a phylogenetic tree was constructed based on the sequence distance method using the neighbor-joining (NJ) algorithm of Saitou and Nei (39) with the Kimura two-parameter model or Poisson corrections for nucleotide and amino acid sequences, respectively. Phylogenetic analyses were conducted with MEGA 2 (20). The stability or accuracy of the inferred topology(ies) was assessed via bootstrap analysis (17) of 1,000 iterations. Character state changes for *A. phagocytophilum msp4* were polarized by designating *A. marginale*, *A. centrale*, and *A. ovis* as outgroups.

**Nucleotide sequence accession numbers.** The GenBank accession numbers for *msp4* sequences of *A. phagocytophilum* HZ, WI-H, MRK, AP-V1, Elsa, Donkey, Horse 31, Roe deer, Ovine, and Bison strains are AY530194 to AY530198, AY702925, and AY706386 to AY706391; for *msp2* sequences of Elsa and Ovine strains, the accession numbers are AY706392 and A706393.

## RESULTS

The analysis of sequence variation in the *msp4* coding region of *A. phagocytophilum* showed some heterogeneity in this locus at the nucleotide (90 to 100% identity) and protein (92 to 100% similarity) levels (Table 2). The 21 samples from infected human blood analyzed from New York during 1996 to 2000 had the same *msp4* sequence, identical to the sequence of the HZ and NY18 New York isolates and to isolates HGE2 Webster (originally isolated from humans) and WI-C1 (originally isolated from a dog) (Table 1). The four strains from humans in Wisconsin (WI-H1 to WI-H4) (WI-H in Tables 2 and 3; Fig. 1 to 3) had identical *msp4* sequences, differing only in one nucleotide at position 687 introducing an M-to-I amino acid change at position 229 with respect to the HZ reference sequence (Table 3). The MRK isolate, originally isolated from

TABLE 3. Differences in the *msp4* nucleotide and amino acid sequences among *A. phagocytophilum* strains<sup>a</sup>

Strain	Nucleotide positions																																
	98	106	118	191	208	250	259	262	271	272	286	291	333	340	411	419	427	428	463	464	466	467	505	610	619	648	687	736	763	764	781		
HZ	A, N	G, A	A, S	C, A	A, B	A, T	G, V	T, S	G, A	C, A	C, H	T, D	C, D	A, T	A, E	A, Y	G, V	T	G, A	C, A	A, T	C, T	A, S	A, N	G, D	A, R	A, I	G, A	A, N	A, N	G, A		
WI-H	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
MRK	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
AP-V1	*	*	*	A, D	*	*	*	*	*	*	*	*	*	*	*	*	*	T, I	*	*	*	*	*	*	*	T, S	*	*	*	*	*	*	*
Elsa	*	*	*	*	*	*	*	*	*	*	*	A, E	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Bison 12	*	*	*	*	*	*	*	*	*	*	*	A, E	*	*	*	*	A, I	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Bison 15	*	*	*	*	*	*	*	*	*	*	*	A, E	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Horse 31	*	*	*	*	*	*	*	*	*	*	*	A, E	*	*	*	*	A, I	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Roe deer	G, S	T, S	G, G	A, D	G, G	G, A	A, I	G, A	A, N	A, N	A, N	A, E	G, A	C, D	T, F	*	*	C, A	A, R	G, R	G, A	*	G, G	A, N	*	C	A, T	G, G	G, G	T, S			
Ovine	*	*	*	A, D	*	*	*	*	A, I	T, I	A, N	*	*	*	G	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

<sup>a</sup> Only variable nucleotide positions that result in amino acid changes are shown, with nucleotides identical to the reference HZ isolate sequence marked with asterisks. Encoded amino acids are shown in italics. Adenine in the translation initiation codon is position 1. Strains of *A. phagocytophilum* are described in Tables 1 and 2.

a horse in California and identified as *E. equi*, and the Italian strains from donkeys had identical *msp4* sequences with a mutation at nucleotide position 648, resulting in an S-to-R amino acid change at position 216 with respect to the HZ sequence (Table 3). A strain from a horse in Germany had a sequence differing from the *msp4* sequence of the MRK isolate, with 12 silent mutations with respect to the HZ sequence (Table 3). The most divergent strains of *A. phagocytophilum* were those derived from ruminants (Tables 2 and 3). The sequence of *A. phagocytophilum* variant 1 (AP-V1) from WTD differed in six nucleotide positions, producing two amino acid changes of D to A and I to T at positions 64 and 156, respectively, when compared to the sequence of the HZ isolate (Table 3). Swiss bovine Elsa and Olga samples and the Polish strain Bison 7 had 17 changes in the nucleotide sequence, resulting in two amino acid changes of E to D and I to V at positions 97 and 143, respectively, compared to the HZ sequence (Table 3). Polish Bison strains 12, 18, and 21 had 14 nucleotide changes, resulting in only one amino acid change identical to the E to D mutation at position 97 in Elsa, Olga, and Bison 7 strains when compared to the HZ sequence (Table 3). Polish Bison strains 15 and 16 had 18 changes at the nucleotide level when compared to the HZ sequence, resulting in the same two amino acid changes found in Elsa, Olga, and Bison 7 strains (Table 3). The Ovine strains from Norway had identical *msp4* sequences, with 27 nucleotide mutations resulting in three amino acid changes when compared to the HZ sequence (Table 3). The sequence of the *msp4* gene was identical in four *A. phagocytophilum* strains from German Roe deer strains, with 86 nucleotide mutations resulting in 23 amino acid changes, compared to the sequence of the HZ isolate (Table 3). The DNA samples from *A. phagocytophilum*-like organisms obtained from WTD in the United States (WTD-OK1, WTD-OK4, WTD-UGA1, and WTD-UGA2) did not produce any detectable signal in the *msp4* PCR, but a PCR targeting a 546-bp amplicon of the 16S rRNA gene produced a positive signal (data not shown).

The HGE2 isolate was analyzed following growth in both human HL-60 and tick ISE6 cells to evaluate if the *msp4* sequence was conserved in bacteria grown in mammalian and tick cells; growth resulted in identical *msp4* sequences.

Phylogenetic analysis of *Anaplasma* spp. *msp4* gene and protein sequences did not provide phylogeographic information (Fig. 1). However, differences were observed in NJ and MP analyses of *msp4* gene and protein sequences according to the original host of the strain, particularly differentiating the *A. phagocytophilum* strains obtained from bovine (Elsa and Olga), WTD (AP-V1), roe deer (Roe deer 2, 6, 8, and 9), European bison (Bison 7, 12, 15, 16, 18, and 21) and ovine (Ovine 5 and 7) ruminants (Fig. 1 and data not shown).

Phylogenetic analysis of *A. phagocytophilum msp2* sequences confirmed this result (Fig. 2). The 21 samples from infected human blood analyzed from New York had the same *msp2* sequence, identical to the sequence of the HZ and NY18 New York isolates. The *msp2* sequences of isolates HGE2, WI-H1 to WI-H4, and WI-C1 were identical to the sequence of strain MN-2. The MRK and Webster isolates had identical *msp2* sequences. The strains Elsa and Olga had identical *msp2* sequences, different from the sequence of strains Ovine 2 and 5, which also had identical sequences.

The phylogenetic analysis of *Ehrlichia* spp. and *Anaplasma*



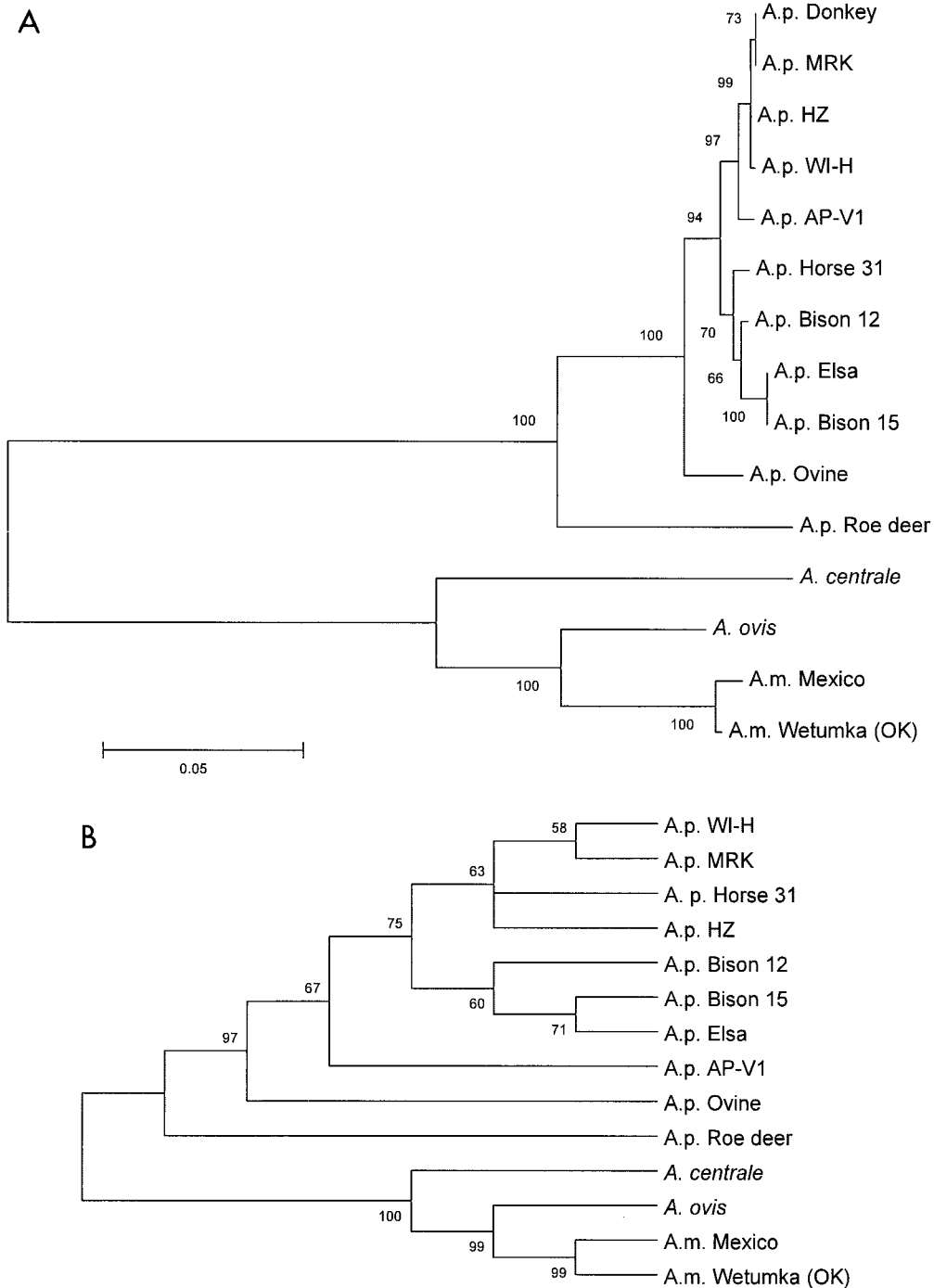


FIG. 1. Phylogenetic analysis of *A. phagocytophilum* strains based on the *msp4* gene (A) and protein (B) sequences with NJ with Kimura two-parameter analysis or Poisson corrections and bootstrap analysis of 1,000 replicates. Numbers on the branches indicate percent support for each clade. Strains of *A. phagocytophilum* are described in Tables 1 and 2. Abbreviations: A.p., *A. phagocytophilum*; A.m., *A. marginale*.

spp. was done using the *msp4* data for *A. phagocytophilum*, *A. marginale*, *A. centrale*, and *A. ovis*; the *E. chaffeensis* 28-kDa MSP; and the *E. ruminantium map1* and the *Wolbachia* spp. MSP sequences (Table 2). *Anaplasma* spp. were monophyletic with respect to *Wolbachia* and *Ehrlichia* spp. (Fig. 3). *E. ruminantium* and *E. chaffeensis* grouped together in a clade separated from *Wolbachia* and *Anaplasma* spp. (Fig. 3).

### DISCUSSION

Phylogenies of *A. phagocytophilum* have been largely inferred using ribosomal and housekeeping genes (16). The *groESL* genes have been used to differentiate North American from European strains and to compare *A. phagocytophilum* strains from roe deer and red deer in Slovenia (34, 42). Al-

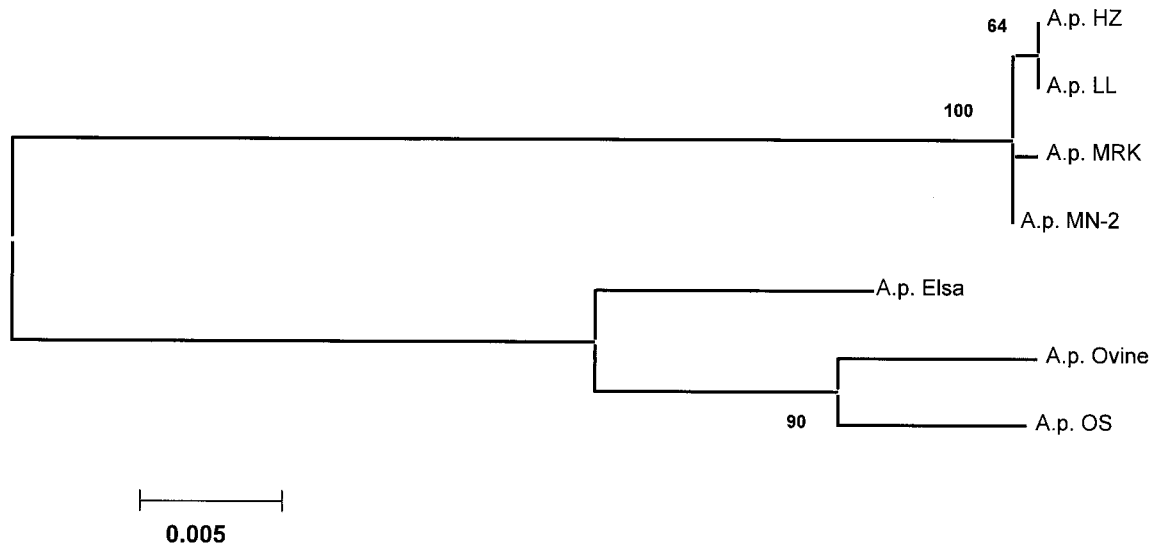


FIG. 2. Phylogenetic analysis of *A. phagocytophilum* strains based on the *msp2* gene sequences with NJ with Kimura two-parameters correction and bootstrap analysis of 1,000 replicates. Numbers on the branches indicate percent support for each clade. Strains of *A. phagocytophilum* are described in Tables 1 and 2. A.p., *A. phagocytophilum*.

though ribosomal and housekeeping genes are powerful tools for determining relatedness at the genus and species level, their strong conservation gives little resolution of closely related species or strains. For this analysis, genes encoding structural or surface proteins like those involved in pathogen-host interactions are more appropriate. Massung et al. (26) used the *ank* gene encoding an *A. phagocytophilum* structural protein for phylogenetic analysis and were able to separate the North American strains into two clades, upper Midwest and Northeast, which were both distinct from the European strains. However, von Loewenich et al. (47) described *ank* variant sequences from infected ticks that are different from all other *ank* sequences and whose geographic distribution is still unknown. The homologue of *A. marginale* MSP4, which has been shown to provide phylogenetic and phylogeographic information about geographic isolates of *A. marginale* (11–14), was identified in silico in the *A. phagocytophilum* genome sequence, thus providing a potentially more informative genetic marker for the characterization of *A. phagocytophilum* strains.

The best characterized major antigenic protein of *A. phagocytophilum*, p44, is highly polymorphic in different populations of the organism derived from tick or mammalian host cells (3). For phylogenetic analyses, a genetic marker which is conserved through the life cycle of the bacterium is preferable. The *msp4* sequence was stable in *A. phagocytophilum* (isolate HGE2) grown in human HL-60 and tick ISE6 cells, corroborating the results obtained for *A. marginale* MSPs in vitro and in vivo (4; unpublished results). The recently discovered *msp2* gene, different from the antigenically variable p44 gene family of *A. phagocytophilum* (22), was also shown to be a candidate gene for phylogenetic studies of *A. phagocytophilum* strains. Future studies addressing the genetic stability of *A. phagocytophilum* *msp2* and *msp4* genes during the life cycle of the bacterium in vivo may further confirm the validity of the use of these genes in phylogenetic studies.

Strains of *A. phagocytophilum* generally showed a high de-

gree of identity at the *msp4* locus, similar to previous results using the 16S rRNA and *groESL* genes and confirmed by genome-wide analysis by pulsed-field electrophoresis (15, 29, 36, 42). However, the strain from European roe deer was unique in that its *msp4* sequence was much more diverse than that for any other strain examined. Further studies to characterize additional genes and gene products of the roe deer strain are needed to determine if the diversity noted for the *msp4* gene reflects genomic divergence or the specific evolution of the *msp4* gene is under some selective pressure. With the exception of the roe deer strain, the *msp4* sequences had a lower degree of diversity than *ank* gene sequences; but unlike results with the *ank* sequences, phylogenetic analysis of the *msp4* sequences differentiated between strains of *A. phagocytophilum* from humans, dogs, and horses from those obtained from ruminants (26). As shown before for *A. marginale* strains in tick pathogen phylogenetic studies (13), the possibility of differentiating between *A. phagocytophilum* samples according to the mammalian host of origin suggested that *msp4* sequences could be used for coevolutionary studies. The degree of *msp4* sequence heterogeneity among strains of *A. phagocytophilum* was higher than among geographic isolates of *A. marginale*, with 99 to 100% identity (Table 2) (14). This higher degree of heterogeneity could be a result of the large number of reservoir hosts in which *A. phagocytophilum* strains have been identified, compared to the relatively limited host range of *A. marginale*, and may provide a higher level of resolution in coevolutionary studies involving the *msp4* gene of *A. phagocytophilum*.

The DNA samples of *A. phagocytophilum*-like organisms obtained from WTD, WTD-OK1, WTD-OK4, WTD-UGA1, and WTD-UGA2, which did not produce any detectable signal in the *msp4* PCR, were suitable for PCR analysis as revealed by the positive signal in the 16S rRNA targeting reaction. These results suggested that the sequence of the *msp4* gene in these samples is more diverse than among strains of *A. phagocyto-*

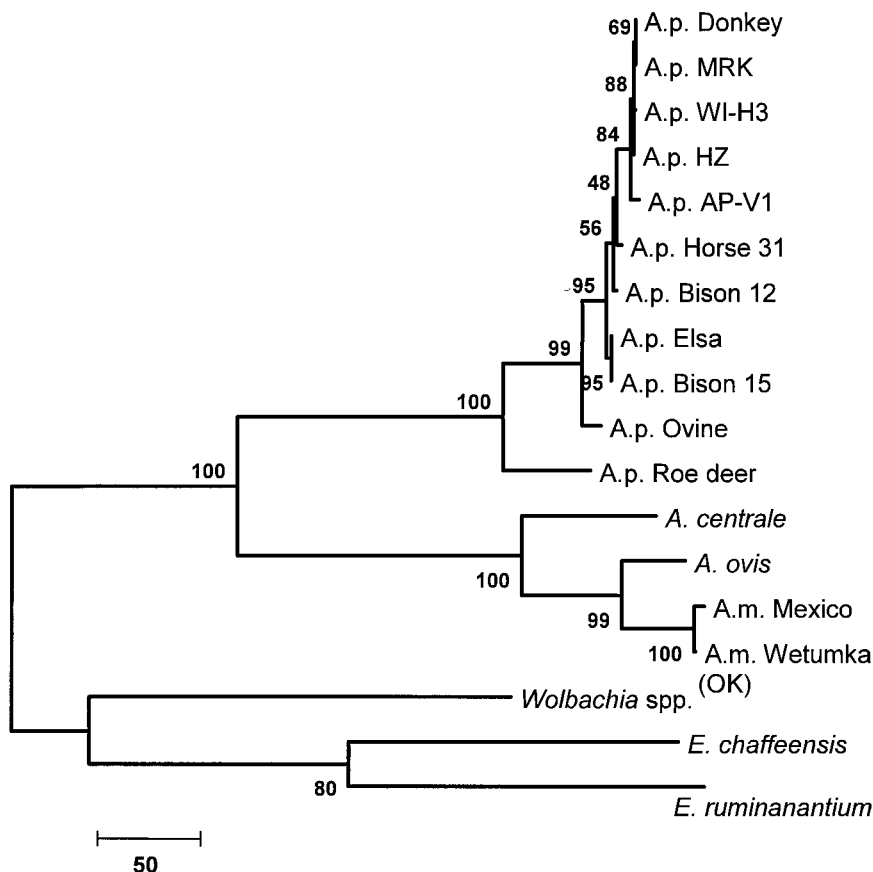


FIG. 3. Phylogenetic analysis of *Anaplasma* and *Ehrlichia* species based on MP analysis of *msp4* sequence data and bootstrap analysis with 1,000 iterations. Numbers on the branches indicate percent support for each clade. Strains of *A. phagocytophilum* are described in Tables 1 and 2. Abbreviations: A.p., *A. phagocytophilum*; A.m., *A. marginale*.

*philum*. Alternatively, the sensitivity of the *msp4* PCR was below the detection level for these samples (five copies of *msp4* per nanogram of DNA). Nevertheless, although *A. phagocytophilum*-like organisms isolated from WTD may be closely related to *A. phagocytophilum* (33), they could prove to be a separate species after further analysis.

The *A. phagocytophilum* variant 1 (sample AP-V1) coexists with and is closely related to the human strains of *A. phagocytophilum* but has never been associated with human infection (9). AP-V1 has been found in *I. scapularis* ticks and WTD in Rhode Island, Connecticut, Maryland, Pennsylvania, and Wisconsin in the United States (9). Strains Elsa and Olga were obtained from infected cattle in Switzerland, where *I. ricinus* is the tick vector; these strains have not been associated with human infection (19, 36). Ovine strains of *A. phagocytophilum* produce a tick-borne fever in sheep with variable pathogenicity but have never been recovered from humans (40). *A. phagocytophilum* strains associated with human disease in Europe have been obtained from *I. ricinus* ticks but not from wild animals such as roe deer (23, 24, 34, 35, 47). These results suggested that *A. phagocytophilum* strains from ruminants could share some common characteristics, including reservoirs and pathogenicity, which may be different from strains that infect humans (27).

The organisms in the order Rickettsiales were recently reclassified based on biological characteristics and genetic analyses

of 16S rRNA, *groESL*, and *gltA* genes (16). These phylogenetic analyses consistently supported the formation of four distinct genera within the family Anaplasmataceae: *Anaplasma*, *Ehrlichia*, *Wolbachia*, and *Neorickettsia*. On the basis of these analyses, Dumler et al. (16) proposed that *E. equi*, *E. phagocytophila*, and the HGE agent should be recognized as synonymous and be renamed as *A. phagocytophilum*. This reclassification has been supported by the analysis of other genetic markers (7, 26, 43). The results reported here using the surface protein gene *msp4* also support this classification, although strains with different reservoirs and pathogenicity may exist.

#### ACKNOWLEDGMENTS

This research was supported by project no. 1669 of the Oklahoma Agricultural Experiment Station; the Endowed Chair for Food Animal Research (K.M.K.), College of Veterinary Medicine, Oklahoma State University; the Forschungskommission of the Albert-Ludwigs-Universität Freiburg, Germany (grant to F.V.L.); and the NIH Centers for Biomedical Research Excellence through a subcontract to J.D.L.F. from the Oklahoma Medical Research Foundation and the Oklahoma Center for the Advancement of Science and Technology, Applied Research Grants AR00(1)-001 and ARO21-37.

The Institute for Genomic Research and Yasuko Rikihisa (Ohio State University) are acknowledged for supplying the preliminary sequence data of *A. phagocytophilum* HZ isolate through the website at <http://www.tigr.org>. Anthony F. Barbet (University of Florida) is acknowledged for supplying DNA from HGE2 isolate grown in HL-60

and ISE6 cells. Janet J. Rogers (Core Sequencing Facility, Department of Biochemistry and Molecular Biology, Noble Research Center, Oklahoma State University) is acknowledged for DNA sequencing.

## REFERENCES

- Allsopp, M. T., C. M. Dorfling, J. C. Maillard, A. Bensaid, D. T. Haydon, H. van Heerden, and B. A. Allsopp. 2001. *Ehrlichia ruminantium* major antigenic protein gene (*map1*) variants are not geographically constrained and show no evidence of having evolved under positive selection pressure. *J. Clin. Microbiol.* **39**:4200–4203.
- Asanovich, K. M., J. S. Bakken, J. E. Madigan, M. Aguero-Rosenfeld, G. P. Wormser, and J. S. Dumler. 1997. Antigenic diversity of granulocytic *Ehrlichia* isolates from humans in Wisconsin and New York and a horse in California. *J. Infect. Dis.* **176**:1029–1034.
- Barbet, A. F., P. F. Meeus, M. Belanger, M. V. Bowie, J. Yi, A. M. Lundgren, A. R. Alleman, S. J. Wong, F. K. Chu, U. G. Munderloh, and S. D. Jauron. 2003. Expression of multiple outer membrane protein sequence variants from a single genomic locus of *Anaplasma phagocytophilum*. *Infect. Immun.* **71**:1706–1718.
- Bowie, M. V., J. de la Fuente, K. M. Kocan, E. F. Blouin, and A. F. Barbet. 2002. Conservation of major surface protein 1 genes of the ehrlichial pathogen *Anaplasma marginale* during cyclic transmission between ticks and cattle. *Gene* **282**:95–102.
- Braig, H. R., W. Zhou, S. L. Dobson, and S. L. O'Neill. 1998. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipitensis*. *J. Bacteriol.* **180**:2373–2378.
- Carlyon, J. A., and E. Fikrig. 2003. Invasion and survival strategies of *Anaplasma phagocytophilum*. *Cell. Microbiol.* **5**:743–754.
- Chae, J. S., J. E. Foley, J. S. Dumler, and J. E. Madigan. 2000. Comparison of the nucleotide sequences of 16S rRNA, 444 *Ep-ank*, and *groESL* heat shock operon genes in naturally occurring *Ehrlichia equi* and human granulocytic ehrlichiosis agent isolates from Northern California. *J. Clin. Microbiol.* **38**:1364–1369.
- Chu, F. K. 1998. Rapid and sensitive PCR-based detection and differentiation of aetiologic agents of human granulocytotropic and monocytotropic ehrlichiosis. *Mol. Cell. Probes* **12**:93–99.
- Courtney, J. W., R. L. Dryden, J. Montgomery, B. S. Schneider, G. Smith, and R. F. Massung. 2003. Molecular characterization of *Anaplasma phagocytophilum* and *Borrelia burgdorferi* in *Ixodes scapularis* ticks from Pennsylvania. *J. Clin. Microbiol.* **41**:1569–1573.
- Dawson, J. E., C. K. Warner, V. Baker, S. A. Ewing, D. E. Stallknecht, W. R. Davidson, A. A. Kocan, J. M. Lockhart, and J. G. Olsen. 1996. Ehrlichia-like 16S rDNA sequence from wild white-tailed deer (*Odocoileus virginianus*). *J. Parasitol.* **82**:52–58.
- de la Fuente, J., E. J. Golsteyn Thomas, R. A. Van Den Bussche, R. G. Hamilton, E. E. Tanaka, S. E. Druhan, and K. M. Kocan. 2003. Characterization of *Anaplasma marginale* isolated from North American bison. *Appl. Environ. Microbiol.* **69**:5001–5005.
- de la Fuente, J., R. A. Van Den Bussche, J. C. Garcia-Garcia, S. D. Rodriguez, M. A. Garcia, A. A. Guglielmo, A. J. Mangold, L. M. Friche Passos, M. F. Barbosa Ribeiro, E. F. Blouin, and K. M. Kocan. 2002. Phylogeography of New World isolates of *Anaplasma marginale* (Rickettsiaceae: Anaplasmataceae) based on major surface protein sequences. *Vet. Microbiol.* **88**:275–285.
- de la Fuente, J., R. A. Van Den Bussche, and K. M. Kocan. 2001. Molecular phylogeny and biogeography of North American isolates of *Anaplasma marginale* (Rickettsiaceae: Ehrlichieae). *Vet. Parasitol.* **97**:65–76.
- de la Fuente, J., R. A. Van Den Bussche, T. Prado, and K. M. Kocan. 2003. *Anaplasma marginale* major surface protein 1 $\alpha$  genotypes evolved under positive selection pressure but are not markers for geographic isolates. *J. Clin. Microbiol.* **41**:1609–1616.
- Dumler, J. S., K. M. Asanovich, and J. S. Bakken. 2003. Analysis of genetic identity of North American *Anaplasma phagocytophilum* strains by pulsed-field gel electrophoresis. *J. Clin. Microbiol.* **41**:3392–3394.
- Dumler, J. S., A. F. Barbet, C. P. J. Bekker, G. A. Dasch, G. H. Palmer, S. C. Ray, Y. Rikihisa, and F. R. Rurangirwa. 2001. Reorganization of the genera in the families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. *Int. J. Syst. Evol. Microbiol.* **51**:2145–2165.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
- Goodman, J. L., C. Nelson, B. Vitale, J. E. Madigan, J. S. Dumler, T. J. Kurtti, and U. G. Munderloh. 1996. Direct cultivation of the causative agent of human granulocytic ehrlichiosis. *N. Engl. J. Med.* **334**:209–215.
- Hudson, J. R. 1950. The recognition of tick-borne fever as a disease of cattle. *Br. Vet. J.* **106**:3–17.
- Kumar, S., K. Tamura, and M. Nei. 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. *Comput. Appl. Biosci.* **10**:189–191.
- Levin, M. L., W. L. Nicholson, R. F. Massung, J. W. Sumner, and D. Fish. 2002. Comparison of the reservoir competence of medium-sized mammals and *Peromyscus leucopus* for *Anaplasma phagocytophilum* in Connecticut. *Vector Borne Zoonotic Dis.* **2**:125–136.
- Lin, Q., Y. Rikihisa, S. Felek, X. Wang, R. F. Massung, and Z. Woldehiwet. 2004. *Anaplasma phagocytophilum* has a functional *msp2* gene that is distinct from *p44*. *Infect. Immun.* **72**:3883–3889.
- Liz, J. S., L. Anderes, J. W. Sumner, R. F. Massung, L. Gern, B. Rutti, and M. Brossard. 2000. PCR detection of granulocytic ehrlichiae in *Ixodes ricinus* ticks and wild small mammals in western Switzerland. *J. Clin. Microbiol.* **38**:1002–1007.
- Liz, J. S., J. W. Sumner, K. Pfister, and M. Brossard. 2002. PCR detection and serological evidence of granulocytic ehrlichial infection in roe deer (*Capreolus capreolus*) and chamois (*Rupicapra rupicapra*). *J. Clin. Microbiol.* **40**:892–897.
- Massung, R. F., M. J. Mauer, J. H. Owens, N. Allan, J. W. Courtney, K. C. Stafford III, and T. N. Mather. 2002. Genetic variants of *Ehrlichia phagocytophila*, Rhode Island and Connecticut. *Emerg. Infect. Dis.* **8**:467–472.
- Massung, R. F., J. H. Owens, D. Ross, K. D. Reed, M. Petrovec, A. Bjersdorff, R. T. Coughlin, G. A. Beltz, and C. I. Murphy. 2000. Sequence analysis of the *ank* gene of granulocytic ehrlichiae. *J. Clin. Microbiol.* **38**:2917–2922.
- Massung, R. F., R. A. Priestley, N. J. Miller, T. N. Mather, and M. L. Levin. 2003. Inability of a variant strain of *Anaplasma phagocytophilum* to infect mice. *J. Infect. Dis.* **188**:1757–1763.
- Massung, R. F., and K. G. Slater. 2003. Comparison of PCR assays for detection of the agent of human granulocytic ehrlichiosis, *Anaplasma phagocytophilum*. *J. Clin. Microbiol.* **41**:717–722.
- Massung, R. F., K. Slater, J. H. Owens, W. L. Nicholson, T. N. Mather, V. B. Solberg, and J. G. Olson. 1998. Nested PCR assay for detection of granulocytic ehrlichiae. *J. Clin. Microbiol.* **36**:1090–1095.
- Munderloh, U. G., S. D. Jauron, V. Fingerle, L. Leitritz, S. F. Hayes, J. M. Hautman, C. M. Nelson, B. W. Huberty, T. J. Kurtti, G. G. Ahlstrand, B. Greig, M. A. Mellencamp, and J. L. Goodman. 1999. Invasion and intracellular development of the human granulocytic ehrlichiosis agent in tick cell culture. *J. Clin. Microbiol.* **37**:2518–2524.
- Munderloh, U. G., Y. Liu, M. Wang, C. Chen, and T. J. Kurtti. 1994. Establishment, maintenance and description of cell lines from the tick *Ixodes scapularis*. *J. Parasitol.* **80**:533–543.
- Munderloh, U. G., J. E. Madigan, J. S. Dumler, J. L. Goodman, S. F. Hayes, J. E. Barlough, C. M. Nelson, and T. J. Kurtti. 1996. Isolation of the equine granulocytic ehrlichiosis agent, *Ehrlichia equi*, in tick cell culture. *J. Clin. Microbiol.* **34**:664–670.
- Munderloh, U. G., C. M. Tate, M. J. Lynch, E. W. Howerth, T. J. Kurtti, and W. R. Davidson. 2003. Isolation of an *Anaplasma* sp. organism from white-tailed deer by tick cell culture. *J. Clin. Microbiol.* **41**:4328–4335.
- Petrovec, M., A. Bidovec, J. W. Sumner, W. L. Nicholson, J. E. Childs, and T. A. Zupanc. 2002. Infection with *Anaplasma phagocytophila* in cervids from Slovenia: evidence of two genotypic lineages. *Wien. Klin. Wochenschr.* **114**:641–647.
- Polin, H., P. Hufnagl, R. Haunschmid, F. Gruber, and G. Ladurner. 2004. Molecular evidence of *Anaplasma phagocytophilum* in *Ixodes ricinus* ticks and wild animals in Austria. *J. Clin. Microbiol.* **42**:2285–2286.
- Pusterla, N., J. B. Huder, H. Lutz, and U. Braun. 1998. Detection of *Ehrlichia phagocytophila* DNA in *Ixodes ricinus* ticks from areas in Switzerland where tick-borne fever is endemic. *J. Clin. Microbiol.* **36**:2735–2736.
- 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.
- Rikihisa, Y., N. Zhi, G. P. Wormser, B. Wen, H. W. Horowitz, and K. E. Hechemy. 1997. Ultrastructural and antigenic characterization of a granulocytic ehrlichiosis agent directly isolated and stably cultivated from a patient in New York State. *J. Infect. Dis.* **175**:210–213.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- Stuenkel, S., K. Bergstrom, M. Petrovec, I. Van de Pol, and L. M. Schouls. 2003. Differences in clinical manifestations and hematological and serological responses after experimental infection with genetic variants of *Anaplasma phagocytophilum* in sheep. *Clin. Diagn. Lab. Immunol.* **10**:692–695.
- Stuenkel, S., S. K. Whist, K. Bergström, and T. Mouton. Exclusion of genotypes may occur in *Anaplasma phagocytophilum* infected lambs. *Vet. Rec.*, in press.
- Sumner, J. W., W. L. Nicholson, and R. F. Massung. 1997. PCR amplification and comparison of nucleotide sequences from the *groESL* heat shock operon of *Ehrlichia* species. *J. Clin. Microbiol.* **35**:2087–2092.
- Taillardat-Bisch, A. V., D. Raoult, and M. Drancourt. 2003. RNA polymerase beta-subunit-based phylogeny of *Ehrlichia* spp., *Anaplasma* spp., *Neorickettsia* spp. and *Wolbachia pipitensis*. *Int. J. Syst. Evol. Microbiol.* **53**:455–458.
- Telford, S. R., III, J. E. Dawson, P. Katavolos, C. K. Warner, C. P. Kolbert, and D. H. Persing. 1996. Perpetuation of the agent of human granulocytic ehrlichiosis in a deer tick-rodent cycle. *Proc. Natl. Acad. Sci. USA* **93**:6209–6214.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W:



- improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Res.* **22**:4673–4680.
46. van Vliet, A. H., F. Jongejan, M. van Kleef, and B. A. van der Zeijst. 1994. Molecular cloning, sequence analysis, and expression of the gene encoding the immunodominant 32-kilodalton protein of *Cowdria ruminantium*. *Infect. Immun.* **62**:1451–1456.
47. von Loewenich, F. D., B. U. Baumgarten, K. Schröppel, W. Geißdörfer, M. Röllinghoff, and C. Bogdan. 2003. High diversity of *ankA* sequences of *Anaplasma phagocytophilum* among *Ixodes ricinus* ticks in Germany. *J. Clin. Microbiol.* **41**:5033–5040.
48. von Loewenich, F. D., G. Stumpf, B. U. Baumgarten, M. Röllinghoff, J. S. Dumler, and C. Bogdan. 2003. A case of equine granulocytic ehrlichiosis provides molecular evidence for the presence of pathogenic *Anaplasma phagocytophilum* (HGE agent) in Germany. *Eur. J. Clin. Microbiol. Infect. Dis.* **22**:303–305.