INFECTION AND IMMUNITY, Feb. 2004, p. 659–666 Vol. 72, No. 2
0019-9567/04/$08.00+0 DOI: 10.1128/IAI.72.2.659–666.2004
Copyright © 2004, American Society for Microbiology. All Rights Reserved.

Sequence Analysis of *p44* Homologs Expressed by *Anaplasma phagocytophilum* in Infected Ticks Feeding on Naive Hosts and in Mice Infected by Tick Attachment

Suleyman Felek,[1] Sam Telford III,[2] Richard C. Falco,[3] and Yasuko Rikihisa[1]*

Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio 43210-1093[1]; Division of Infectious Diseases, School of Veterinary Medicine, Tufts University, North Grafton, Massachusetts 01536[2]; and Vector Ecology Laboratory, Louis Calder Center, Fordham University, Armonk, New York 10504[3]

Received 17 July 2003/Returned for modification 22 September 2003/Accepted 11 November 2003

The 44-kDa immunodominant outer membrane proteins (P44 proteins) of *Anaplasma phagocytophilum* are encoded by the *p44* polymorphic multigene family. The present study examined *p44* expression and analyzed the cDNA sequences of various *p44* transcripts from the spleens and blood of mice infected by the bites of ticks infected with the *A. phagocytophilum* NTN-1 strain or of naturally infected nymphal ticks and in the salivary glands and midgut tissues of these ticks. A total of 300 *p44* cDNAs were subjected to sequence analysis. Of these, 40 distinct *p44* species were found, and all of these had orthologs in the *A. phagocytophilum* HZ strain genome that shared 95 to 100% base sequence identity. The number of unique *p44* species expressed in mouse blood was greater than that for mouse spleens. Higher numbers of different *p44* transcripts were also expressed in the salivary glands of ticks than in the midgut tissues. Variations in the sequences of the same *p44* cDNA species within a single *A. phagocytophilum* strain and among different strains were concentrated in the conserved regions flanking the central hypervariable region of *p44* genes. No mosaic sequences derived from two or more *p44* species were found within the *p44* hypervariable region. The conservation of the hypervariable region of each *p44* cDNA species of *A. phagocytophilum* in naturally infected ticks and in different geographic isolates suggests that each *A. phagocytophilum* genome carries a set of *p44* paralogs to be expressed. Thus, a large but restricted repertoire of *p44* hypervariable sequences exists in *A. phagocytophilum* strains in the Northeastern United States.

Human granulocytic ehrlichiosis (HGE) is a tick-borne zoonosis that was first reported in the United States in 1994. This disease has been documented in North America and in Europe and was designated in 1998 as a nationally reportable disease in the United States (22). *Anaplasma phagocytophilum*, the etiologic agent of HGE, is a gram-negative, obligatory intracellular bacterium that primarily infects neutrophils (10). In the Eastern and Midwestern United States, *A. phagocytophilum* is transmitted by *Ixodes scapularis* ticks (8, 33). *Ixodes pacificus* and *Ixodes ricinus* have been identified as vectors in California (30) and Europe (26), respectively. The major reservoir of *A. phagocytophilum* in the Northeastern and Midwestern United States is the white-footed mouse (*Peromycus leucopus*), which may be infected either transiently or persistently (29, 34). Although various wild and domestic animal species have been reported to be infected with *A. phagocytophilum* (10), to date the 16S rRNA gene sequence corresponding to *A. phagocytophilum* that infects humans (the HGE agent) has been identified only in white-footed mice (29, 34, 37), horses (16a, 27, 36), and dogs (16a, 28).

The incubation period of HGE is 1 to 2 weeks after the tick bite occurs (2). The most common symptoms and signs of the disease include fever, myalgia, rigors, headache, and malaise (2, 11). Most patients also exhibit leukopenia, thrombocytopenia, and elevated levels of hepatic transaminases (2, 11). The fatality rate is approximately 2 to 5% (22), and deaths have occurred primarily in patients that acquire opportunistic infections after contracting HGE (11).

Outer membrane proteins of 44 to 49 kDa have been shown to act as the major antigens recognized by human sera (1, 15, 35, 38–40). The 44-kDa major outer membrane proteins (P44 proteins) of *A. phagocytophilum* are encoded by the *p44* polymorphic multigene family. The *p44* genes are dispersed in the genome (38, 40). Based on the preliminary *A. phagocytophilum* HZ genome sequence (http://www.tigr.org), the total number of *p44* paralogs is >80. *p44* genes consist of a central hypervariable region, of approximately 280 bp, which is flanked by conserved sequences. Many, but not all, *p44* genes lack start codons (38, 40, 41). Interleukin-1β, tumor necrosis factor-α, and interleukin-6 are generated when human peripheral blood leukocytes are exposed to recombinant P44-1 protein (18). This release of proinflammatory cytokines may explain the clinical signs and hematological abnormalities associated with HGE and suggests that P44 proteins play a role in the pathogenesis of this disease (18). Members of our laboratory previously reported that passive immunization of mice with monoclonal antibodies against P44 proteins partially protects them against *A. phagocytophilum* infection (17). Thus, P44 proteins may represent potential vaccine candidates for HGE.

Expression studies showed that *p44* genes are differentially expressed in cultured cells and ticks and in the blood of mice, horses, and human patients (4, 14, 20, 40, 41). *p44* is homologous to *msp2* of *Anaplasma marginale*, which is a bovine in-

* Corresponding author. Mailing address: Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093. Phone: (614) 292-9677. Fax: (614) 292-6473. E-mail: rikihisa.1@osu.edu.

traerythrocytic agent (24, 38, 40). It has been reported that the segmental gene conversion of multiple *msp2* genes at a single expression locus in *A. marginale* allows unlimited variation of the expressed *msp2* chimera or mosaic (3, 6, 7). Recently, Barbet et al. proposed that the same mechanism controls the expression of variable *p44* genes in *A. phagocytophilum* (4). If this is true, the hypervariable region sequences of *p44* in the expression locus will be chimeras or mosaics of two or more donor *p44* genes elsewhere in the same genome. However, this contrasts with previous data which suggested that the hypervariable region sequences in several *p44* paralogs cloned from different genomic loci are identical in the corresponding regions of *p44* cDNA (40, 41). For the present study, we investigated *p44* expression in the blood and spleens of strain DBA/2 mice infected by attaching *I. scapularis* ticks to simulate natural infection and in salivary glands and midgut tissues of infected ticks feeding on naive hosts (transmission-fed ticks). These ticks were infected either naturally or experimentally with the NTN-1 strain of *A. phagocytophilum*. The NTN-1 strain was maintained by alternating tick-mouse passages (33). P44 cDNA sequences from field ticks and the NTN-1 strain were compared with *p44* DNA sequences of the HZ strain. In addition, individual cDNA sequences of the same *p44* species, both within the same strain and among different strains, were compared to identify sequence differences and variable regions.

MATERIALS AND METHODS

Tick attachment. One hundred *I. scapularis* nymphs were collected in Westchester County, N.Y., by drag sampling (12). Ticks were kept in an incubator at 18°C under 95 to 100% relative humidity with a 12-h photoperiod for more than 10 days, until they were attached to male DBA/2 mice (5 to 6 weeks old) (Harlan Sprague-Dawley, Indianapolis, Ind.).

The *A. phagocytophilum* NTN-1 strain used was in approximately the 14th passage (tick-mouse cycle). The original isolate was obtained by inoculating a C3H mouse with Feulgen-positive salivary glands from ticks collected from the yard of the index patient in Nantucket, Mass. (33). This mouse served as the host for lab-reared ticks. The strain is maintained by allowing infected ticks to feed on naive mice and then placing naive larvae on the mice. Two hundred *I. scapularis* larvae from a laboratory colony were infected by attaching them to ICR strain SCID mice infected with the NTN-1 strain of *A. phagocytophilum* (Taconic Farm Inc., Germantown, N.Y.). Ticks were incubated at 22°C and 95% relative humidity, with a 12-h photoperiod, until they molted. Ticks were then stored at 10°C until their use.

Since field-collected ticks may be infected with various strains of *A. phagocytophilum*, one potentially infected free-living nymph from Westchester County, N.Y., was attached to each of 16 naive mice to prevent mixing of strains. Since NTN-1 is a strain, five nymphs removed from mice experimentally infected with NTN-1 were attached to each of 10 naive mice to increase the chance of transmission. The backs of the mice were shaved with a razor. A tick chamber was prepared from 0.5-ml PCR tubes by removing most of the tube and making a hole in the cap. This ring-like structure (5 mm in diameter) was attached to a fine nylon mesh. Either one or five ticks were placed in each chamber. The open end of the cap and the mesh of the chambers were then fixed to the skin of mice with a water-based adhesive (3 M, St. Paul, Minn.). Each mouse was restrained in a wire cage for 3 days to protect the ticks from host grooming. Food and water were freely accessible to the mice. Engorged ticks were collected from the chambers after 3 days. Most of the ticks were detached by that time; those that were still attached were removed with a fine forceps.

Tick and mouse samples. Immediately after collection, ticks were soaked for 10 s in a 10% bleach solution, rinsed with distilled deionized water, and dried with Kimwipes (Kimberly-Clark, Roswell, Ga.). Ticks were cut in half on the median line with a sterile razor blade. The salivary glands and midgut tissues were then excised under a dissecting microscope. As much blood as possible was removed from the midgut by washing with Dulbecco's minimal essential medium (GIBCO-BRL, Grand Island, N.Y.). Specimens were stored in RNA-Later RNA preserving reagent prior to analysis (Qiagen, Valencia, Calif.) according to the manufacturer's recommended protocol.

Mice were killed 10 days after tick attachment under isoflurane anesthesia. A heart puncture was used to obtain heparinized blood, which was then centrifuged to obtain the buffy coat. One-third of the buffy coat was immediately used for DNA purification, and the remainder was stored in RNA-Later (Qiagen). The spleen was removed aseptically, minced, and stored at −20°C in RNA-Later until further analysis.

DNA isolation from ticks and blood samples and nested PCR. For identification of infected ticks and mice, DNA was isolated from individual ticks and from the buffy coat of individual mice. A QIAmp tissue kit (Qiagen) was used to extract DNA from ticks and a QIAmp blood kit (Qiagen) was used to extract DNA from buffy coat specimens according to the manufacturer's instructions. The DNA was used as a template for nested PCR amplification of the *A. phagocytophilum p44* gene. DNA from *A. phagocytophilum* HZ cultured in HL-60 cells (31) was used as a positive control for the template, and double-distilled water was included in place of template DNA as a negative control. In the first PCR, 10 μl of template DNA was amplified in a 50-μl reaction mixture containing 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 5 mM $MgCl_2$, a 0.2 mM concentration of each deoxynucleoside triphosphate, 1.5 U of *Taq* DNA polymerase (Invitrogen-Life Technologies, San Diego, Calif.), and 5 pmol of the *p44* gene-specific primer pair p3708-p4257 (40). Amplification was performed in a GeneAmp PCR System 9700 thermal cycler (Perkin-Elmer Applied Biosystems, Norwalk, Conn.) with a three-step program (5 min of denaturation at 94°C; 35 cycles of 1 min of denaturation at 94°C, 1 min of annealing at 52°C, and 1 min of extension at 72°C; and a final extension of 7 min). For the nested PCR, 1 μl of the first PCR product was amplified in a second 50-μl reaction mixture prepared as described above, except that primers p3761 and p4183 were used (20). The same temperature cycle was used as described for the first PCR, except that the annealing temperature was 57°C.

RT-PCR and cDNA cloning. The total RNA was purified from blood and spleen specimens from individual mice and from salivary gland and midgut specimens from pooled or single ticks by use of the RNeasy Protect mini kit (Qiagen) according to the manufacturer's instructions. Samples containing 1 to 2 μg of RNA were treated with 2 U of amplification-grade DNase I (Invitrogen) at 37°C for 20 min. DNase I was then inactivated by heating at 65°C for 10 min in the presence of 2.5 mM EDTA. One half of the total RNA was then subjected to reverse transcription (RT) in a 20-μl reaction mixture containing 10 mM random hexamers, a 0.5 mM concentration of each deoxynucleoside triphosphate, 1 U of RNase inhibitor (Invitrogen), and 200 U of SuperScript II reverse transcriptase (Invitrogen) at 42°C for 50 min. The reaction was terminated by heating the mixture to 70°C for 15 min. The other half of the total RNA was subjected to the identical procedure in the absence of reverse transcriptase. PCR was carried out on both reaction mixtures to exclude the possibility of contamination of the RNA preparation by DNA. Samples containing 2 μl of cDNA template were subjected to PCR using the *p44* gene-specific primer pair p3708-p4257 (40). The resulting PCR products were cloned into the PCRII TA cloning vector (Invitrogen), and both strands of the inserted DNA were sequenced.

Sequence analysis. DNA sequences were analyzed with the program DNASTAR (DNASTAR Inc., Madison, Wis.). DNA and amino acid sequences were aligned and nucleotide sequence identities were determined by using the CLUSTAL method in the DNASTAR program. Phylogenetic analysis was performed with the PHYLIP software package (version 3.5).

Statistical analysis. The Mood median test (23) was used to compare diversities of the various species of *p44* transcripts found in tissues. Minitab statistical software, version 13.31, was used for statistical analysis.

GenBank accession numbers. The GenBank accession numbers for the new p44 sequences described here are as follows: *p44-51* (NTN-1), AY234863; *p44-52* (NTN-1), AY234864; *p44-53* (NTN-1), AY234865; *p44-54* (NTN-1), AY234866; *p44-55* (NTN-1), AY234867; *p44-56* (NTN-1), AY236466; and *p44-57* (Westchester), AY234868. The GenBank accession number for the unfinished whole-genome sequence of the *A. phagocytophilum* HZ strain is NC_004351, and the sequence is also available online (http://www.tigr.org).

RESULTS

Infection rates of ticks and mice, as determined by PCR. For determination of the rate of *A. phagocytophilum* infection in ticks for the experiment, DNAs were extracted from 10 individual naturally infected (Westchester strain) nymphal ticks and from 5 individual experimentally infected nymphal ticks

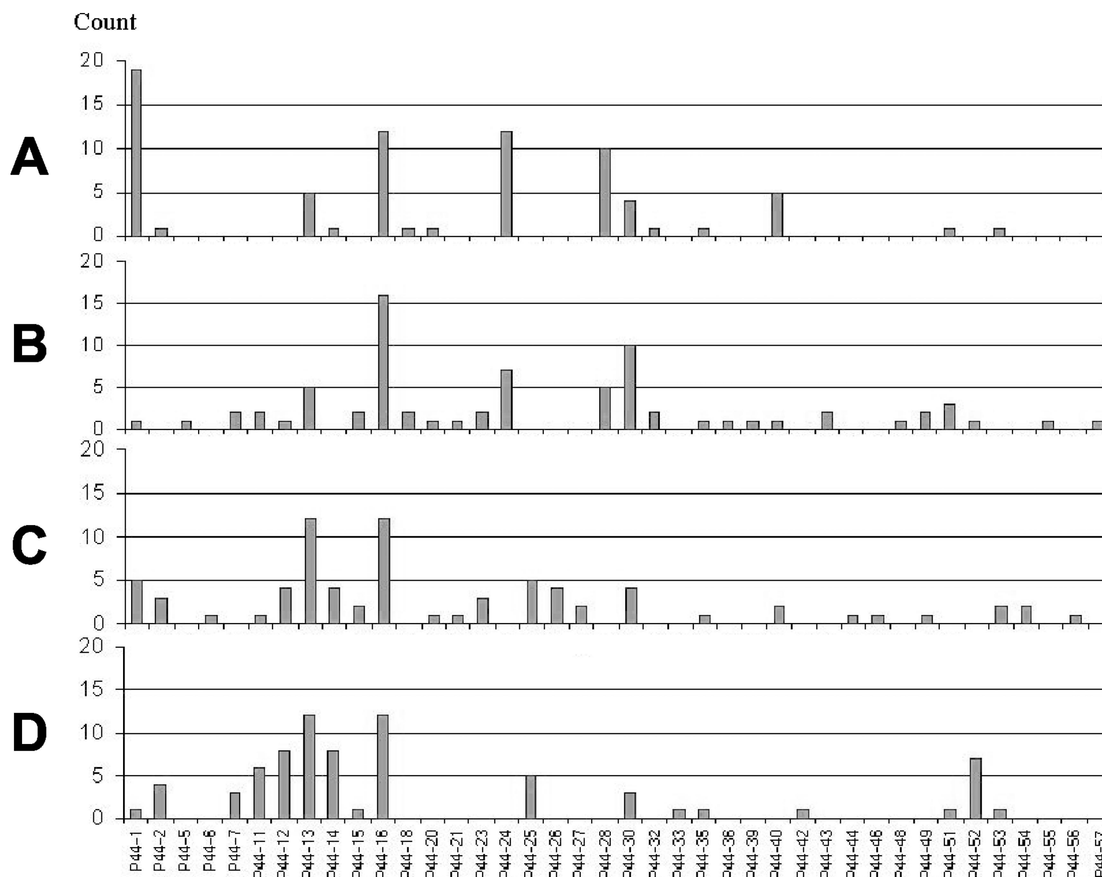

FIG. 1. Numbers of each *p44* transcript species found expressed by *A. phagocytophilum* NTN-1 or Westchester in various tissues taken from three infected mice and from infected ticks attached to the mice. (A) Mouse spleens; (B) mouse blood; (C) tick salivary glands; (D) tick midgut specimens. The vertical axis shows numbers of cDNA clones of each *p44* species detected among a total of 300 cDNA clones sequenced. The horizontal axis shows the various *p44* cDNA clone species.

(NTN-1 strain) and were subjected to nested PCR with *p44* primers. One of the 10 naturally infected ticks and 1 of the 5 experimentally infected ticks were found to be positive. Of DNAs extracted from blood samples derived from 26 individual mice, specimens from 1 of 16 mice hosting naturally infected ticks (one tick per mouse) and from 9 of 10 mice hosting experimentally infected ticks (five ticks per mouse) were found to be positive for *A. phagocytophilum* infection by nested PCR.

TABLE 1. Numbers of different *p44* cDNA species identified in *A. phagocytophilum*-infected mouse and tick tissues

Mouse no./tick ID[c]	No. of different *p44* cDNA species in indicated tissue			
	Mice		Ticks	
	Spleen	Blood	Salivary gland	Midgut
4/Westchester	4	16	12	11
26/NTN-1	8	14	12	4
27/NTN-1	7	10	14	7
Average ± SD	6.3 ± 2.1[a]	13.3 ± 3.1	12.7 ± 1.2[b]	7.3 ± 3.6

[a] $P = 0.014$ (when compared to the number of *p44* species expressed in the blood).
[b] $P = 0.014$ (when compared to the number of *p44* species expressed in the midgut).
[c] The Westchester tick was one naturally infected tick collected in Westchester County, N.Y. The NTN-1 ticks identification included five ticks experimentally infected with the NTN-1 strain of *A. phagocytophilum.*

Differential expression of *p44* in various tissues. For identification of the specific *p44* transcripts expressed in mice and ticks, samples were subjected to *p44* RT-PCR. The resulting *p44* cDNA PCR products were sequenced on both strands. Twenty-five cDNA clones from blood samples and 25 cDNA clones from spleen samples of each of the three infected mice (one with the Westchester strain and two with the NTN-1 strain) were sequenced. cDNA clones derived from salivary gland and midgut samples from one Westchester strain-infected tick or two pools of five ticks infected with the NTN-1 strain removed from each of these three mice were also analyzed, such that a total of 300 *p44* cDNA clones were sequenced. Overall, a total of 40 different species of *p44* were found to be transcribed in all specimens, with 27 different *p44* species found in all blood samples alone (Fig. 1). Higher numbers of *p44* species were found in mouse blood (13.3 ± 3.1 [mean ± standard deviation; [median, 14; $N = 3$) than in mouse spleens (6.3 ± 2.1; median, 7; $N = 3$) ($P = 0.014$). Higher numbers of *p44* species were found in tick salivary glands (12.6 ± 1.2; median, 12; $N = 3$) than in tick midgut specimens (7.3 ± 3.6; median, 7; $N = 3$) ($P = 0.014$) (Table 1). Some sequences (*p44-18* and *p44-28*) were found in three or

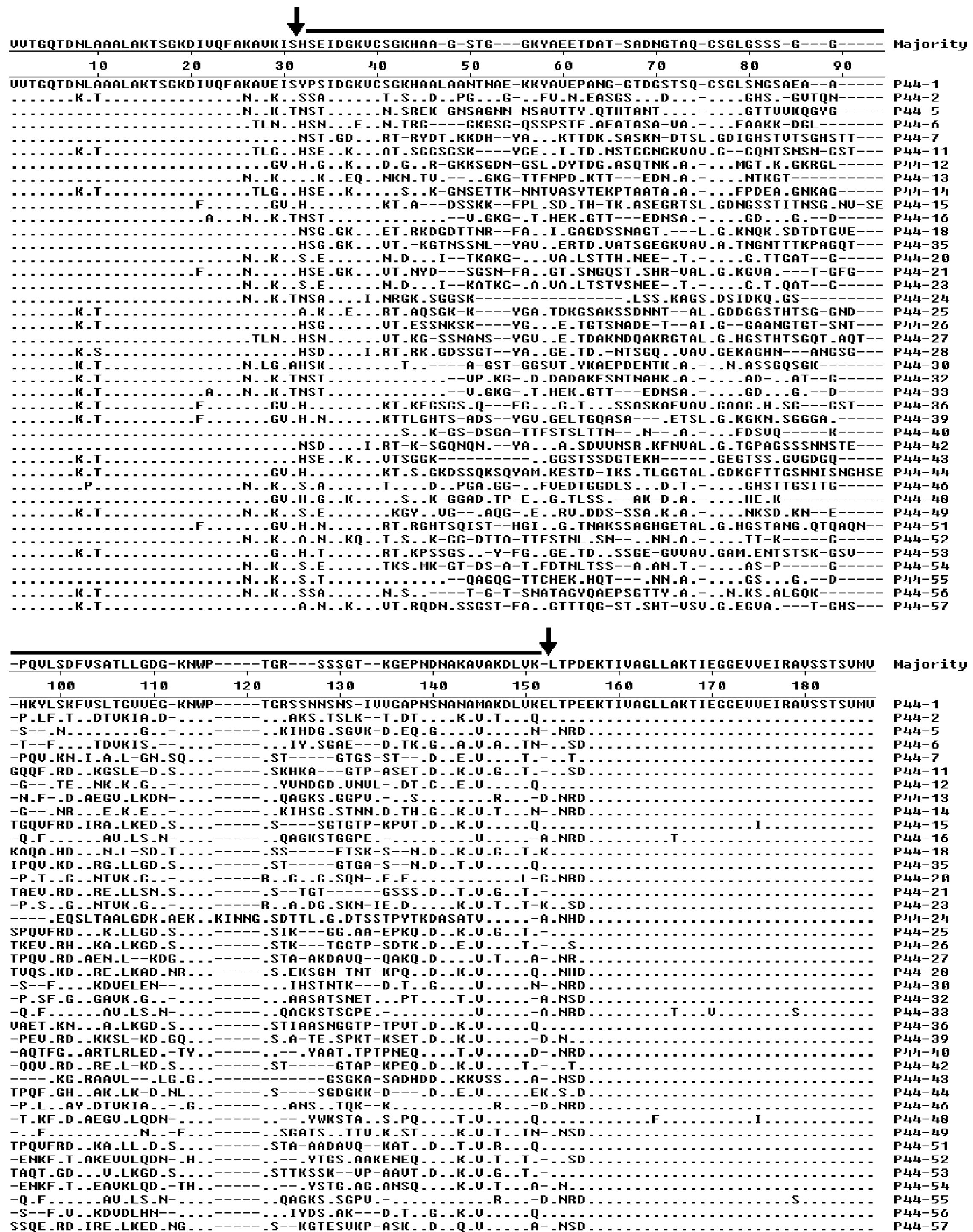

FIG. 2. Alignment of representative deduced amino acid sequences of 40 *p44* species found to be expressed in the present study. Aligned positions of identical amino acids are indicated by dashes. Gaps are indicated by dashed lines. The hypervariable region (40) is marked with a thick line (between two vertical arrows). Sequences P44-51 to P44-57 were identified for the first time in the present study.

more of the six mouse specimens analyzed, but not in any of the tick specimens. Other *p44* species, such as *p44-25*, were found in more than three tick specimens, but not in mouse specimens, suggesting that some *p44* species are preferentially expressed in either mammals or ticks (Fig. 1). We identified seven new *p44* cDNA sequences (*p44-51* to *-57*) that shared 61.3 to 85.2% nucleotide sequence identities with the closest known *p44* cDNA or DNA sequences. These sequences have not yet been found to be expressed by the *A. phagocytophilum* HZ, NY-31, NY-36, or NY-37 strains. An alignment of the deduced amino acid sequences of all 40 expressed *p44* species showed sequence identities ranging from 44 to 86%, as shown

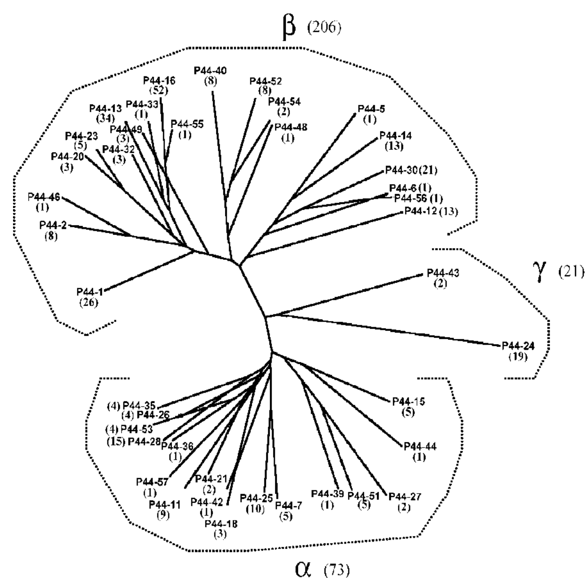

FIG. 3. Phylogenetic analysis of deduced amino acid sequences of 40 different *p44* transcript species found expressed in mice and/or ticks. Amino acid sequences were aligned with the Clustal V program, and the tree was constructed by the neighbor-joining method. cDNA clone frequencies (of a total of 300 clones sequenced) for each *p44* species are indicated in parentheses.

in Fig. 2. This finding suggests the copresence of diverse P44 antigenic phenospecies.

Phylogenetic relationship among expressed *p44* paralogs. A phylogenetic analysis of the deduced amino acid sequences of all 40 of the *p44* cDNA species showed that they were distributed into three groups, namely α, β, and γ, as defined previously (20). Twenty-one, or 52.5%, of these belonged to the α group; 17, or 42.5%, belonged to the β group; and 2, or 5%, belonged to the γ group of *p44* species. Of all 300 *p44* cDNA clones that were sequenced, 206, or 68.7%, belonged to the α group; 73, or 24.3%, belonged to the β group; and 21, or 7%, belonged to the γ group (Fig. 3). Average of 9.3 ± 12 ($n = 21$) and 4.3 ± 3.8 ($n = 17$) cDNA clones of each *p44* species were found in the β and α groups, respectively, and these numbers were not significantly different ($P = 0.083$). Of the 73 cDNA clones associated with the α group, 35, or 47.9%, were found in mice, and of the 206 cDNA clones associated with the β group, 94, or 45.6%, were found in mice. *p44* genes in the γ group included *p44-24* (found in 19 cDNA clones from four of the six mouse specimens analyzed) and *p44-43* (found in 2 cDNA clones from one of six mouse specimens analyzed). These species (*p44-24* and *p44-43*) were found to be expressed only in mouse specimens (Fig. 1 and 3). *p44-18* (α group) was rarely expressed in the mice 10 days after tick placement. The seven most frequently expressed *p44* genes from all 300 cDNA clones identified in mice and ticks were *p44-16* (52 clones, or 17.3%), *p44-13* (34 clones, or 11.3%), *p44-1* (26 clones, or 8.7%), *p44-23* (24 clones, or 8.0%), *p44-30* (21 clones, or 7.0%), *p44-24* (19 clones, or 6.3%), and *p44-28* (15 clones, or 5.0%) (Fig. 3).

Aligned base sequence comparison. cDNA sequences (~500 bp after the removal of primer sequences) belonging to the same *p44* species exhibited minor sequence variations within the same strain as well as among strains. These were unlikely to result from errors of *Taq* polymerase, since the same bases differed in both strands and often among several cDNA clones. When *p44* cDNA sequences from the Westchester strain (from naturally infected ticks) and the NTN-1 strain (from experimentally infected ticks) were compared with the preliminary *A. phagocytophilum* HZ strain genome sequence, orthologs were found for all 40 cDNA species. The levels of identity between the HZ strain genome sequence and other cDNA base sequences were 95% for *p44-54*, 97% for *p44-13*, *p44-20*, and *p44-56*, and 98 to 100% for the remaining 36 different *p44* sequences. To find regions of sequence variations within the same *p44* cDNA species, we aligned the cDNA sequences within the same *p44* species and strains and combined the data. Finally, the alignment data for each *p44* species from a total of eight *p44* species were combined. For this analysis, the eight *p44* species that showed the most frequent sequence differences were selected, as shown in Fig. 4A. To find regions of sequence variations in the same *p44* cDNA species among different strains, we aligned one representative majority *p44* species cDNA sequence from each strain with the respective *p44* ortholog DNA sequence for the HZ strain (35 NTN-1 strain cDNAs, 24 Westchester strain cDNAs, and 40 HZ strain DNA sequences, for a total of 99 sequences), as shown in Fig. 4B. The results of these cDNA sequence alignments showed that sequence differences among the same *p44* cDNA species were found primarily in the 5′- and 3′-conserved regions flanking the central hypervariable regions of *p44* genes (Fig. 4). The central hypervariable region base sequences (40) were highly conserved among the same *p44* cDNA species across different strains as well as within each strain (Fig. 4). The DNA sequence variations were primarily found in the third base of the codon. Thus, these variations rarely changed the predicted amino acid sequences and did not alter the *p44* cDNA species.

DISCUSSION

Antigenic changes in the bacterial surface play an important role in avoiding immunosurveillance and destruction by the host immune system. Polymorphic and immunodominant P44 proteins of *A. phagocytophilum* are the primary candidates for such a role. Antigenic variation is known to be used by several tick-borne bacterial pathogens to help them persist for long periods in the immunocompetent mammalian host, since tick feeding is seasonal (5). Furthermore, in regions where HGE is endemic, wild mammalian reservoirs frequently develop antibodies to the pathogen (25, 32, 37). Thus, the bacterial expression of diverse surface antigens likely improves the bacterium's odds of survival during tick transmission. In the present study, we identified several *p44* species that were simultaneously expressed in tick and mouse tissues. Mouse blood and the tick salivary gland expressed a more diverse array of *p44* species than the mouse spleen and the tick midgut. Blood and salivary glands are two tissues that may play a role in the transmission of ehrlichiae between mammalian and tick hosts, supporting the hypothesis that the diverse expression of various *p44* species may be important for the transmission of *A. phagocytophilum*. Within a single tick, as many *p44* species were transcribed as were found in pooled ticks. Differential expression of *p44* genes diversifies the outer membrane protein properties of *A.*

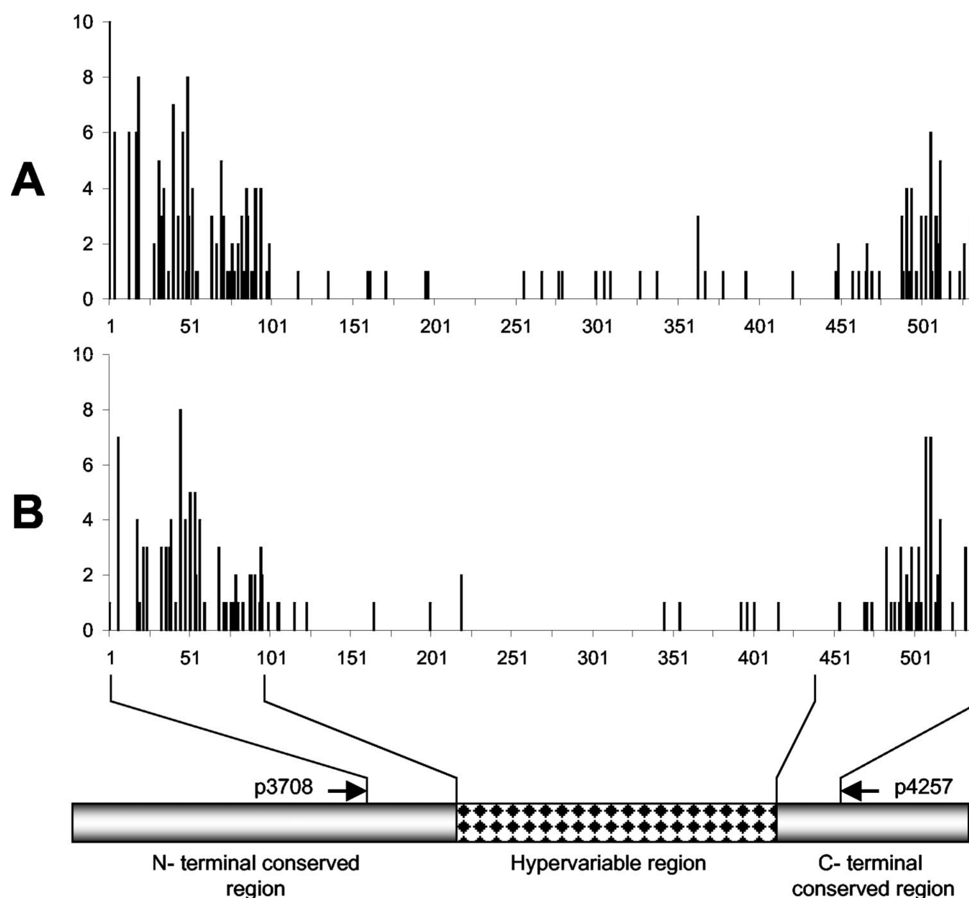

FIG. 4. Frequencies of base differences in cDNA (NTN-1 and Westchester strains) and DNA (HZ strain) sequences of individual *p44* genes among the same and different strains. (A) Differences within the same strain. The vertical axis shows combined frequencies of differences at each base position for each *p44* species within each strain. The eight *p44* species that showed the most frequent differences in sequences (*p44-12* [2 for the Westchester strain and 5 for the NTN-1 strain], *p44-14* [6 for the NTN-1strain], *p44-16* [4 for the Westchester strain and 12 for the NTN-1 strain], *p44-24* [6 for the NTN-1 strain], *p44-30* [2 for the Westchester strain and 12 for the NTN-1 strain], *p44-40* [4 for the NTN-1 strain], *p44-51* [2 for the Westchester strain and 3 for the NTN-1 strain], and *p44-53* [2 for the Westchester strain and 2 for the NTN-1 strain]) were aligned with the Clustal V program and compared within each *p44* species in each strain. (B) Differences among different strains. A single representative sequence was selected for each *p44* species in each strain. The cDNA and DNA sequences were aligned and base differences were compared with the majority at each nucleotide position. The vertical axis shows combined frequencies of differences at each base position for each *p44* species among strains (35 NTN-1 strain cDNAs, 24 Westchester strain cDNAs, and 40 HZ strain DNAs, for a total of 99 sequences). The horizontal axis indicates the base positions after alignment.

phagocytophilum. Thus, this may also help the bacteria to adapt to and avoid innate immune responses in ticks. The diverse antigenic variation of *p44* poses a challenge for the development of *p44*-based vaccines against *A. phagocytophilum*.

In the present study, we found 10 to 16 different *p44* species that were expressed in blood samples from each of three mice 10 days after tick placement. Ijdo et al. reported that six different *p44* species (of a total of 20 cDNA clones) were expressed in the blood specimens from one C3H/HeN mouse 10 days after intraperitoneal inoculation with SCID mouse blood infected with the NCH1 strain of *A. phagocytophilum* (14). In HGE patients, multiple *p44* species were cotranscribed (4, 20). With experimentally infected mammals, it was previously shown that two different *p44* species of the HZ strain were transcribed in the blood of horses 8 days after inoculation or tick attachment, two different *p44* species of the HZ strain were transcribed in mouse blood specimens 4 days after inoculation, and seven and two different *p44* species, respectively, were expressed in the midguts and salivary glands of transmission-fed female tick samples (41). Jauron et al. showed that P44 antigens were expressed at low levels in ISE6 tick cells but that they were strongly expressed in human HL-60 cells (16). Thus, coexpression of diverse *p44* species by *A. phagocytophilum* appears to be pervasive in various environments.

It was previously reported that *p44-18* was the major transcript species in the *A. phagocytophilum* HZ strain found in horse and mouse blood at 8 and 4 days postinfection, respectively, but that this transcript was rarely detected in ticks (41). Also, expression of the *p44-18* transcript was upregulated in *A. phagocytophilum* cultivated in HL-60 cells at 37°C compared to cells cultured at 24°C (41). In the present study, *p44-18* transcripts were minor and were found only in mouse samples, not in tick samples. Thus, it appears that *A. phagocytophilum* strains or other factors determine the major *p44* species expressed in mammals. Phylogenetic analysis of *p44* cDNA sequences revealed that the *p44* cDNAs are clustered into three

groups, the α, β, and γ groups. In the present study, *p44-24* and *p44-43*, which are members of the γ group, were found to be expressed only in mouse specimens. Because only two *p44* species were found for the γ group, we could not perform statistical analysis on this finding. Our sequence alignment showed that none of the six *p44* species of the NCH1 strain expressed in transmission-fed ticks described by Ijdo et al. (14) belong to the γ group.

A. marginale is the most prevalent tick-transmitted bacterium in cattle and is closely related to *A. phagocytophilum* (19). Antigenic variation in the major surface protein 2 (MSP2) of *A. marginale* causes persistent infections of red blood cells in cattle. *msp2* is a member of a multigene family. During infection in cattle, multiple MSP2 variants occur every 6 to 8 weeks, and such antigenic variation allows the bacteria to escape from preexisting immune responses (13). Nine *msp2* pseudogenes and a single complete gene have been reported in the St. Marines, Florida, and South Idaho strains of *A. marginale* (6). Recombination of multiple small segments of pseudogenes into the expression site by gene conversion generates a nearly infinite number of expressed *msp2* sequences (3, 7). The *p44* family of *A. phagocytophilum* shares a high degree of homology with the *A. marginale msp2* family (24, 38, 40). Recently, Barbet et al. proposed that similar segmental conversion acts to control the differential expression of *p44* genes by *A. phagocytophilum* (4). However, a recent study (21) suggested that differential *p44* gene expression in *A. phagocytophilum* human isolates occurs through nonreciprocal conversion of the entire (nonsegmental) *p44* hypervariable region, including parts of the conserved flanking region sequences, at the *p44* expression locus to a sequence copied from one of the conserved *p44* donor genomic loci. In the present study, we compared the cDNA sequences we obtained from the NTN-1 and Westchester tick strains with the *A. phagocytophilum* HZ strain genome sequence to investigate the possibility of combinatorial recombination such as that which occurs in *A. marginale*. We found that the HZ genome has *p44* genes at each distinct locus that are orthologous to every *p44* sequence (438 to 522 bp) expressed in the NTN-1 and Westchester tick strains. There were only between 0 and 5% differences among the same *p44* species sequences in the HZ, NTN-1, and Westchester tick strains, and 36 of 40, or 90%, of the sequences, shared 98 to 100% identity. This indicates that despite the geographical difference between Nantucket Island, Mass. (NTN-1 strain), and Westchester County, N.Y., or between HGE patient isolates and bacteria isolated from field-collected ticks, the strain divergence for each *p44* species was relatively low. Within a single strain or isolate, the identities of the same *p44* species cDNA sequences from four tissues of mice and ticks were 97 to 100%. This further suggests that combinatorial recombination is undetectable in *p44* genes of *A. phagocytophilum* and that the P44 antigenic repertoire at the hypervariable region is restricted.

The novel finding of the present study is that in each *p44* cDNA species, conserved region base sequences flanking the central hypervariable (*p44* species-specific) region were paradoxically more variable than those of the central hypervariable region within the same strain as well as across different strains. This result is consistent with our first hypothesis, which proposed nonsegmental gene conversion of the *p44* hypervariable region (21). These findings also support our second hypothesis (21) that these divergent sequences in the flanking conserved regions result from heteroduplex formation between two nonidentical *p44* sequences at both regions and from mismatch repair, as previously demonstrated for homologous genes of *Escherichia coli* (9).

The present results show that *p44* expression patterns vary in different host tissues. Further studies are needed to understand the detailed mechanisms by which this diverse pattern of *p44* expression occurs and to determine what factors facilitate this varied expression in different environments. These new findings will be useful for the eventual design of effective vaccines and therapeutic strategies against *A. phagocytophilum*.

ACKNOWLEDGMENTS

We thank O. Ozturk for statistical help and T. Boccia for assistance with field work.

This research was supported by grant R01 AI47407 from the National Institutes of Health. The genome of *A. phagocytophilum* was sequenced at The Institute for Genomic Research. The sequencing project was supported by National Institutes of Health grant R01 AI47885 to Y.R.

REFERENCES

1. **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.
2. **Bakken, J. S.** 1998. The discovery of human granulocytotropic ehrlichiosis. J. Lab. Clin. Med. **132:**175–180.
3. **Barbet, A. F., A. Lundgren, J. Yi, F. R. Rurangirwa, and G. H. Palmer.** 2000. Antigenic variation of *Anaplasma marginale* by expression of MSP2 mosaics. Infect. Immun. **68:**6133–6138.
4. **Barbet, A. F., P. F. M. 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.
5. **Barbour, A. G., and B. I. Restrepo.** 2000. Antigenic variation in vector-borne pathogens. Emerg. Infect. Dis. **6:**449–457.
6. **Brayton, K. A., D. P. Knowles, T. C. McGuire, and G. H. Palmer.** 2001. Efficient use of a small genome to generate antigenic diversity in tick-borne ehrlichial pathogens. Proc. Natl. Acad. Sci. USA **98:**4130–4135.
7. **Brayton, K. A., G. H. Palmer, A. Lundgren, J. Yi, and A. F. Barbet.** 2002. Antigenic variation of *Anaplasma marginale msp2* occurs by combinatorial gene conversion. Mol. Microbiol. **43:**1151–1159.
8. **Daniels, T. J., R. C. Falco, I. Schwartz, S. Varde, and R. G. Robbins.** 1997. Deer ticks (*Ixodes scapularis*) and the agents of Lyme disease and human granulocytic ehrlichiosis in a New York City park. Emerg. Infect. Dis. **3:**353–355.
9. **DasGupta, C., and C. M. Radding.** 1982. Polar branch migration promoted by RecA protein: effect of mismatched base pairs. Proc. Natl. Acad. Sci. USA **79:**762–766.
10. **Dumler, J. S., A. F. Barbet, C. P. Bekker, G. A. Dasch, G. H. Palmer, S. C. Ray, Y. Rikihisa, and F. R. Rurangirwa.** 2001. Reorganization of 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.
11. **Dumler, J. S., and J. S. Bakken.** 1998. Human ehrlichioses: newly recognized infections transmitted by ticks. Annu. Rev. Med. **49:**201–213.
12. **Falco, R. C., D. F. McKenna, T. J. Daniels, R. B. Nadelman, J. Nowakowski, D. Fish, and G. P. Wormser.** 1999. Temporal relation between *Ixodes scapularis* abundance and risk for Lyme disease associated with erythema migrans. Am. J. Epidemiol. **149:**771–776.
13. **French, D. M., W. C. Brown, and G. H. Palmer.** 1999. Emergence of *Anaplasma marginale* antigenic variants during persistent rickettsemia. Infect. Immun. **67:**5834–5840.
14. **Ijdo, J. W., C. Wu, S. R. Telford III, and E. Fikrig.** 2002. Differential expression of the *p44* gene family in the agent of human granulocytic ehrlichiosis. Infect. Immun. **70:**5295–5298.
15. **Ijdo, J. W., Y. Zhang, E. Hodzic, L. A. Magnarelli, M. L. Wilson, S. R.**

Telford III, S. W. Barthold, and E. Fikrig. 1997. The early humoral response in human granulocytic ehrlichiosis. J. Infect. Dis. **176:**687–692.

16. **Jauron, S. D., C. M. Nelson, V. Fingerle, M. D. Ravyn, J. L. Goodman, R. C. Johnson, R. Lobentanzer, B. Wilske, and U. G. Munderloh.** 2001. Host cell-specific expression of a p44 epitope by the human granulocytic ehrlichiosis agent. J. Infect. Dis. **184:**1445–1450.

16a. **Johansson, K.-E., B. Pettersson, M. Uhlen, A. Gunnarsson, M. Malmzvist, and E. Olsson.** 1995. Identification of the causative agent of granulocytic ehrlichiosis in Swedish dogs and horses by direct solid phase sequencing of PCR products from the 16S rRNA gene. Res. Vet. Sci. **58:**109–112.

17. **Kim, H. Y., and Y. Rikihisa.** 1998. Characterization of monoclonal antibodies to the 44-kilodalton major outer membrane protein of the human granulocytic ehrlichiosis agent. J. Clin. Microbiol. **36:**3278–3284.

18. **Kim, H. Y., and Y. Rikihisa.** 2000. Expression of interleukin-1beta, tumor necrosis factor alpha, and interleukin-6 in human peripheral blood leukocytes exposed to human granulocytic ehrlichiosis agent or recombinant major surface protein p44. Infect. Immun. **68:**3394–3402.

19. **Kocan, K. M., E. F. Blouin, and A. F. Barbet.** 2000. Anaplasmosis control. Past, present, and future. Ann. N. Y. Acad. Sci. **916:**501–509.

20. **Lin, Q., N. Zhi, N. Ohashi, H. W. Horowitz, M. E. Aguero-Rosenfeld, J. Raffalli, G. P. Wormser, and Y. Rikihisa.** 2002. Analysis of sequences and loci of *p44* homologs expressed by *Anaplasma phagocytophila* in acutely infected patients. J. Clin. Microbiol. **40:**2981–2988.

21. **Lin, Q., Y. Rikihisa, N. Ohashi, and N. Zhi.** 2003. Mechanisms of variable *p44* expression by *Anaplasma phagocytophilum.* Infect. Immun. **71:**5650–5661.

22. **McQuiston, J. H., C. D. Paddock, R. C. Holman, and J. E. Childs.** 1999. The human ehrlichioses in the United States. Emerg. Infect. Dis. **5:**635–642.

23. **Mood, A. M.** 1954. On the asymptotic efficiency of certain nonparametric two-sample test. Ann. Math. Stat. **25:**514–522.

24. **Murphy, C. I., J. R. Storey, J. Recchia, L. A. Doros-Richert, C. Gingrich-Baker, K. Munroe, J. S. Bakken, R. T. Coughlin, and G. A. Beltz.** 1998. Major antigenic proteins of the agent of human granulocytic ehrlichiosis are encoded by members of a multigene family. Infect. Immun. **66:**3711–3718.

25. **Nicholson, W. L., M. B. Castro, V. L. Kramer, J. W. Sumner, and J. E. Childs.** 1999. Dusky-footed wood rats (*Neotoma fuscipes*) as reservoirs of granulocytic ehrlichiae (*Rickettsiales*: *Ehrlichieae*) in northern California. J. Clin. Microbiol. **37:**3323–3327.

26. **Petrovec, M., J. W. Sumner, W. L. Nicholson, J. E. Childs, F. Strle, J. Barlic, S. Lotric-Furlan, and T. Avsic Zupanc.** 1999. Identity of ehrlichial DNA sequences derived from *Ixodes ricinus* ticks with those obtained from patients with human granulocytic ehrlichiosis in Slovenia. J. Clin. Microbiol. **37:**209–210.

27. **Pusterla, N., J. B. Huder, K. Feige, and H. Lutz.** 1998. Identification of a granulocytic *Ehrlichia* strain isolated from a horse in Switzerland and comparison with other rickettsiae of the *Ehrlichia phagocytophila* genogroup. J. Clin. Microbiol. **36:**2035–2037.

28. **Pusterla, N., J. Huder, C. Wolfensberger, B. Litschi, A. Parvis, and H. Lutz.** 1997. Granulocytic ehrlichiosis in two dogs in Switzerland. J. Clin. Microbiol. **35:**2307–2309.

29. **Ravyn, M. D., C. B. Kodner, S. E. Carter, J. L. Jarnefeld, and R. C. Johnson.** 2001. Isolation of the etiologic agent of human granulocytic ehrlichiosis from the white-footed mouse (*Peromyscus leucopus*). J. Clin. Microbiol. **39:**335–338.

30. **Reubel, G. H., R. B. Kimsey, J. E. Barlough, and J. E. Madigan.** 1998. Experimental transmission of *Ehrlichia equi* to horses through naturally infected ticks (*Ixodes pacificus*) from Northern California. J. Clin. Microbiol. **36:**2131–2134.

31. **Rikihisa, Y., N. Zhi, G. Wormser, B. Wen, H. W. Horowitz, and K. E. Hechemy.** 1997. Ultrastructural and antigenic characterization of granulocytic ehrlichiosis agent directly isolated and stably cultivated from a patient in New York state. J. Infect. Dis. **175:**210–213.

32. **Stafford, K. C. III, R. F. Massung, L. A. Magnarelli, J. W. Ijdo, and J. F. Anderson.** 1999. Infection with agents of human granulocytic ehrlichiosis, Lyme disease, and babesiosis in wild white-footed mice (*Peromyscus leucopus*) in Connecticut. J. Clin. Microbiol. **37:**2887–2892.

33. **Telford, S. R., III, T. J. Lepore, P. Snow, C. K. Warner, and J. E. Dawson.** 1995. Human granulocytic ehrlichiosis in Massachusetts. Ann. Intern. Med. **123:**277–279.

34. **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.

35. **Unver, A., S. Felek, C. D. Paddock, N. Zhi, H. W. Horowitz, G. P. Wormser, L. C. Cullman, and Y. Rikihisa.** 2001. Western blot analysis of sera reactive to human monocytic ehrlichiosis and human granulocytic ehrlichiosis agents. J. Clin. Microbiol. **39:**3982–3986.

36. **Von Loewenich, F. D., G. Stumpf, B. U. Baumgarten, M. Rollinghoff, 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.

37. **Walls, J. J., B. Greig, D. F. Neitzel, and J. S. Dumler.** 1997. Natural infection of small mammal species in Minnesota with the agent of human granulocytic ehrlichiosis. J. Clin. Microbiol. **35:**853–855.

38. **Zhi, N., N. Ohashi, Y. Rikihisa, H. W. Horowitz, G. P. Wormser, and K. Hechemy.** 1998. Cloning and expression of the 44-kilodalton major outer membrane protein gene of the human granulocytic ehrlichiosis agent and application of the recombinant protein to serodiagnosis. J. Clin. Microbiol. **36:**1666–1673.

39. **Zhi, N., Y. Rikihisa, H. Y. Kim, G. P. Wormser, and H. W. Horowitz.** 1997. Comparison of major antigenic proteins of six strains of the human granulocytic ehrlichiosis agent by Western immunoblot analysis. J. Clin. Microbiol. **35:**2606–2611.

40. **Zhi, N., N. Ohashi, and Y. Rikihisa.** 1999. Multiple *p44* genes encoding major outer membrane proteins are expressed in the human granulocytic ehrlichiosis agent. J. Biol. Chem. **274:**17828–17836.

41. **Zhi, N., N. Ohashi, T. Tajima, J. Mott, R. W. Stich, D. Grover, S. R. Telford III, Q. Lin, and Y. Rikihisa.** 2002. Transcript heterogeneity of the *p44* multigene family in a human granulocytic ehrlichiosis agent transmitted by ticks. Infect. Immun. **70:**1175–1184.

Editor: V. J. DiRita