Analysis of Sequences and Loci of *p44* Homologs Expressed by *Anaplasma phagocytophila* in Acutely Infected Patients

Quan Lin,[1] Ning Zhi,[1] Norio Ohashi,[1] Harold W. Horowitz,[2] Maria E. Aguero-Rosenfeld,[3] John Raffalli,[2] Gary P. Wormser,[2] and Yasuko Rikihisa[1]*

Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio 43210,[1] and Division of Infectious Diseases, Department of Medicine,[2] and Department of Pathology,[3] New York Medical College, Valhalla, New York 10595

Anaplasma phagocytophila **is an obligatory intragranulocytic bacterium that causes human granulocytic ehrlichiosis. Immunodominant 44-kDa outer membrane proteins of *A. phagocytophila* are encoded by a *p44* multigene family. In the present study, expression profiles of *p44* genes in the blood of acutely infected patients in the year 2000 were characterized. A single *p44* gene was predominantly expressed in peripheral blood leukocytes from one patient, while up to 17 different *p44* genes were transcribed without a single majority in the other two patients. The cDNA sequences of the central hypervariable region of several *p44* genes were identical among the isolates from the three patients and a 1995 *A. phagocytophila* isolate. *A. phagocytophila* was isolated by cell culture from all of the three 2000 patients. Genomic Southern blot analysis of the three 2000 and two 1995 *A. phagocytophila* isolates with probes specific to the most dominant *p44* transcript in each patient showed that the *p44* loci in the *A. phagocytophila* genome were conserved. Analysis of the predicted amino acid sequences of 43 different *p44* genes including 19 new sequences found in the present study, revealed that five amino acids were absolutely conserved. The hypervariable region was subdivided into five domains, including three extremely hypervariable central domains. These results suggest that variations in the sequences of *p44* are not random but are restricted. Furthermore, several *p44* genes are not hypermutatable in nature, based on the conservation of gene sequences and loci among isolates obtained 5 years apart.**

Human granulocytic ehrlichiosis (HGE) is an emerging tick-borne zoonosis in the United States and Europe (2, 6, 16, 20). HGE is characterized by a nonspecific febrile illness, accompanied by malaise, chills, headache, myalgias, and arthralgias. Laboratory findings include leukopenia, thrombocytopenia, and elevated serum transaminases (2). HGE can be severe enough to require hospitalization and can be fatal, particularly when treatment is delayed due to misdiagnosis or in the elderly or immunocompromised patients (2, 9). The etiologic agent of HGE is *Anaplasma phagocytophila,* a gram-negative, obligatory intracellular bacterium that infects granulocytes of various mammalian species (6, 7).

The 44-kDa immunodominant outer membrane proteins (P44s) of *A. phagocytophila* are encoded by a multigene family consisting of more than 20 different paralogous genes dispersed throughout the genome (22, 23). Our previous studies suggested that P44s may play a role in the pathogenesis of HGE by inducing proinflammatory cytokines (13) and that antibodies to P44s neutralize *A. phagocytophila* infection (12). The protein structure of P44s is characterized by a single central hypervariable region consisting of approximately 94 amino acids flanked by N- and C-terminal conserved regions (22).

Studies on expression of *p44* paralogs by *A. phagocytophila* in cell culture and in tick transmission models with mice and horses revealed that selected sets of *p44* genes are differentially expressed in different host environments (23, 24). However,

which *p44* paralogs are transcribed by *A. phagocytophila* in HGE patients is unknown. Therefore, the objectives of the present study were to (i) analyze profiles of *p44* expression by *A. phagocytophila* among HGE patients and (ii) determine the nature of variation of *p44* gene sequences and *p44* loci among human *A. phagocytophila* isolates.

MATERIALS AND METHODS

Patient specimens. Serum samples and buffy coat specimens were prepared from the blood of three patients (NY31, NY36, and NY37) suspected of having HGE based on their clinical presentation at the Westchester Medical Center in June and July of 2000 (Table 1). Parts of the buffy coat specimens were stored in RNALater (Ambion, Inc. Austin, Tex.) until used for transcription analysis of the *p44* multigene family.

Culture isolation and IFA testing. The culture isolation of *A. phagocytophila* from the buffy coat specimens and indirect fluorescent antibody (IFA) testing with local human isolate NY-13 as an antigen were performed at New York Medical College as previously described (1, 11).

Cultivation and purification. Three *A. phagocytophila* strains (isolates NY-31, NY-36, and NY-37) that were newly isolated from patients NY31, NY36, and NY37, respectively, in 2000 and two human strains, HZ and LL (19, 21), were cultured in HL-60 cells with RPMI 1640 medium (Invitrogen, Carlsbad, Calif.) supplemented with 10% heat-inactivated fetal bovine serum (Atlanta Biologicals, Norcross, Ga.) at 37°C in 5% CO_2–95% air. The infectivity was checked daily by Diff-Quik staining of cytocentrifuged cells (Baxter Scientific Products, Obetz, Ohio). The cells were harvested when the infectivity reached 80%. *A. phagocytophila* organisms were purified by Sephacryl S-1000 (Amersham Pharmacia Biotech, Piscataway, N.J.) column chromatography as previously described (17).

Western blot analysis. The procedure for Western blotting was essentially that described previously (18). Briefly, proteins of *A. phagocytophila* from three isolates were separated by 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane. The membrane was incubated with the serum samples diluted to 1:500. As a positive control, a mouse anti-P44-specific monoclonal antibody, 5C11 (12), was used at a dilution of

---

* Corresponding author. Mailing address: Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, 1925 Coffey Road, Columbus, OH 43210-1093. Phone: (614) 292-5661. Fax: (614) 292-6473. E-mail: rikihisa.1@osu.edu.

TABLE 1. Clinical signs and IFA titers of patients with HGE

Patient	Clinical signs	Date (mo/day/yr)	IFA		Date of blood collection for expts[a] (mo/day/yr)
			Titer	Date (mo/day/yr)	
NY31	Fever, rigors, dizziness, sore throat, diarrhea	6/15/2000	<80	6/21/2000	6/21/2000
			≥2,560	6/28/2000	
			1,280	7/12/2000	
			640	10/31/2000	
			640	2/6/2001	
NY36	Fever, chills, fatigue, anorexia, nausea, myalgias	7/7/2000	<80	7/10/2000	7/10/2000
			≥2,560	7/17/2000	
			≥2,560	7/31/2000	
			≥2,560	10/3/2000	
			320	2/13/2001	
			320	5/15/2001	
NY37	Fever, fatigue, muscle aches, stiff-neck, joint aches	7/18/2000	<80	7/19/2000	7/19/2000
			<80	7/21/2000	
			320	8/9/2000	
			≥80	11/14/2000	
			<80	2/20/2001	

[a] Used for mRNA analysis, Western blotting, and isolation of *A. phagocytophila*.

1:1,000. The peroxidase-conjugated affinity-purified anti-human or anti-mouse immunoglobulin G secondary antibody (Kirkegaard & Perry Laboratories Inc., Gaithersburg, Md.) was diluted to 1:2,000. The peroxidase reaction was carried out in 70 mM sodium acetate buffer (pH 6.2) containing 0.3% diaminobenzidine tetrahydrochloride (Nakarai Tesque, Inc., Kyoto, Japan) and 0.03% H_2O_2, and the reaction was stopped by washing the membrane in 0.1 M H_2SO_4.

RT-PCR and cDNA cloning. Total RNA was extracted from the buffy coat with Trizol reagent (Invitrogen, Carlsbad, Calif.). The RNA was further purified with the RNeasy mini kit (Qiagen, Valencia, Calif.). For cDNA synthesis, the isolated RNA (0.3 to 0.5 μg) was heated at 70°C for 10 min and reverse transcribed in a 20-μl reaction mixture containing 0.5 mM each of the four deoxynucleoside triphosphates, 200 U of SuperScript II reverse transcriptase (Invitrogen), 200 ng of random hexamer, and 3 mM $MgCl_2$ at 42°C for 50 min. A pair of degenerate oligonucleotides, p3761 (5′-CTGCTCTKGCCAARACCTC-3′, where K is T or G and R is A or G) and p4183 (5′-CAATAGTYTTAGCTAGTAACC-3′, where Y is C or T), were designed based on the conserved regions among cDNA sequences of *p44*-homologous genes for amplification of multiple *p44* transcripts.

PCR was performed in a 50-μl reaction mixture containing 4 μl of the cDNA product, 10 pmol each of the p3761 and p4183 primers, 0.2 mM each of the four deoxynucleoside triphosphates, 5 U of *Taq* DNA polymerase, and 1.5 mM $MgCl_2$, with 3 min of denaturation at 94°C, followed by 40 cycles of 1 min of denaturation at 94°C, 1 min of annealing at 55°C, and 2 min of extension at 72°C. To rule out DNA contamination in the RNA preparation, reverse transcription (RT)-PCR without reverse transcriptase was carried out (negative control). Amplified RT-PCR products were cloned into the pCRII vector (Invitrogen) according to the manufacturer's instruction, and 49 or 50 cDNA clones that were randomly selected from the transformants of each patient specimen were sequenced by the dideoxy termination method with the ABI Prism BigDye terminator cycle sequencing reaction kit and an ABI 373XL Stretch DNA sequencer.

Genomic Southern blot analysis. Genomic DNAs from three isolates (isolates NY-31, NY-36, and NY-37) and strains HZ and LL were purified by lysis with sodium dodecyl sulfate, pronase digestion, phenol-chloroform extraction, and ethanol precipitation as described elsewhere (15). For Southern blot analysis, the genomic DNAs were digested with three restriction enzymes, *Eco*RI, *Pst*I, and *Xba*I. DNA probes specific to transcripts *p44-23* and *p44-28*, which were predominantly transcribed in the samples from patients NY36 and NY37, respectively, were prepared by PCR as described elsewhere (23). Briefly, primer pairs were designed to amplify the central hypervariable region of *p44-23*, pP44-23-5 (5′-AATGGGGATCATGCTGCTAT-3′) and pP44-23-3 (5′-AGGGTCACCTT CTATGTTCTT-3′), and the central hypervariable region of *p44-28*, pP44-28-5

(5′-TTTGTAGGACGAAGCGGAAG-3′) and pP44-28-3 (5′-AGGTTGAGGC TTTGTGTTAGT-3′).

Amplicons were cloned into a pCRII vector (Invitrogen). The DNA insert excised from each recombinant plasmid was used as a template to prepare the probe that was labeled with [α-^{32}P]dATP with the random primer kit (Amersham Pharmacia Biotech). Hybridization was performed in rapid hybridization buffer (Amersham Pharmacia Biotech) at 65°C as described elsewhere (15). The membrane was exposed to a Hyperfilm (Amersham Pharmacia Biotech). The densities of hybridized bands were analyzed with a PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.).

Nucleotide sequence accession numbers. Multiple amino acid sequences were aligned by the Clustal V method in the MegAlign program (DNAStar, Madison, Wis.). Phylogenetic analysis was performed with the PHYLIP software package (version 3.5).

Published sequences of the P44 multigene family used for construction of amino acid alignments were: *hge-1*, AF356507 (14); *hge-44*, AF037599 (10); *p44-1*, AF059181; *p44-2*, AF135254; *p44-12*, AF135255; *p44-15*, AF135256; *p44-18*, AF135257; *p44-19*, AF135263; (23) *p44-3*, AF412818; *p44-4*, AF412819; *p44-5*, AF412820; *p44-6*, AF412821; *p44-7*, AF412822; *p44-8*, AF412823; *p44-9*, AF412824; *p44-10*, AF412825; *p44-11*, AF412826; *p44-13*, AF412827; *p44-14*, AF412828; *p44-16*, AF412829; *p44-17*, AF412830; *p44-21*, AF412831; and *p44-20*, AF414591 (24).

Sequences of newly identified *p44* transcripts have been assigned GenBank accession numbers: *p44-23*, AY064512; *p44-24*, AY064513; *p44-25*, AY064514; *p44-26*, AY064515; *p44-27*, AY064516; *p44-28*, AY064517; *p44-29*, AY064518; *p44-30*, AY064519; *p44-31*, AY064520; *p44-32*, AY064521; *p44-33*, AY064522; *p44-34*, AY064523; *p44-35*, AY064524; *p44-36*, AY064525; *p44-37*, AY064526; *p44-38*, AY064527; *p44-39*, AY064528; *p44-40*, AY064529; and *p44-41*, AY064530.

RESULTS

Clinical signs, laboratory findings, IFA titers, and isolation of strains of *A. phagocytophila* from patients (year 2000). All three patients had clinical evidence of HGE, including fever, headache, malaise, and myalgia and laboratory findings of leukopenia and/or thrombocytopenia (Table 1). IFA titers of all patients were less than 1:80 when specimens were initially collected, rising to 1:2,560 after 7 days in patients NY31 and
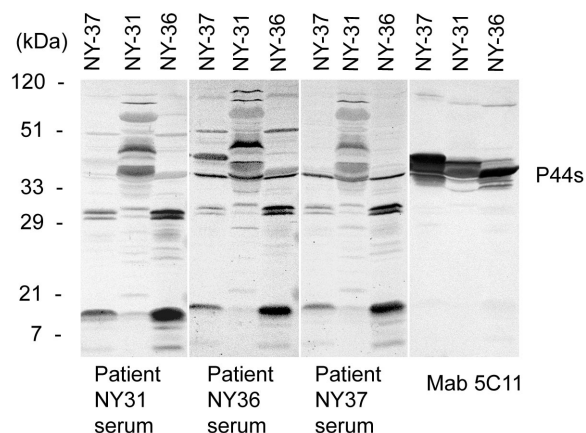
FIG. 1. Western blot analysis of serum samples and monoclonal antibody 5C11 against new isolates of *A. phagocytophila*. Three purified isolates (NY-31, NY-36, and NY-37) were used as the antigens. Multiple 44-kDa proteins (P44s) that reacted with serum samples from patients NY31, NY36, and NY37 and monoclonal antibody (MAb) 5C11 are indicated by the bracket on the right. Numbers on the left show molecular sizes.

FIG. 2. Analysis of expression of *p44* multigene family in buffy coats from three HGE patients by RT-PCR. DNA+, with 0.1 ng of genomic DNA from the purified *A. phagocytophila* strain HZ as a positive control; DNA−, without DNA as a negative control; RT+, with reverse transcriptase; RT−, without reverse transcriptase; lane M, molecular size markers. Numbers on the left show molecular sizes (in base pairs).

NY36 and to 1:320 after 21 days in patient NY37 (Table 1). *A. phagocytophila* was recovered from the blood of all three patients at the time of initial blood sample collection. These new isolates from patients NY31, NY36, and NY37 were designated isolates NY-31, NY-36, and NY-37, respectively. Morulae (microcolonies of *A. phagocytophila*) were seen in granulocytes only in the peripheral blood smear from patient NY36.

Reactivity of serum samples with three new isolates of *A. phagocytophila*. By Western immunoblot analysis, the reactivity of each patient's serum samples with the three new isolates (isolates NY-31, NY-36, and NY-37) was examined. Although these serum specimens had IFA titers of <1:80 (Table 1), by Western blot analysis we could see distinct positive reactions to each isolate. As a positive control, a P44-specific monoclonal antibody (5C11) was used to detect all P44 proteins in the new isolates. The proteins were observed as multiple bands of similar sizes. All three patients' serum samples reacted with P44 proteins, indicating that these proteins were expressed by *A. phagocytophila* in the patients (Fig. 1).

Differences in the reaction pattern were seen among the three isolates used as antigens. Whereas the reaction patterns of the serum samples from the three patients were similar regardless of which isolate was used as the antigen (Fig. 1), fewer P44 proteins of isolates NY-37 and NY-36 were recognized by the three patients' serum samples than by monoclonal antibody 5C11, suggesting that some of the P44s produced in isolates NY-37 and NY-36 in culture were not produced or recognized in all three patients. On the other hand, the P44s produced in isolate NY-31 in culture were well recognized by all patients' serum samples.

Characterization of *p44* transcripts expressed by *A. phagocytophila* in three patients. Since Western blot analysis cannot distinguish among P44 homologs expressed by *A. phagocytophila* in patients, transcripts of *p44* homologs in buffy coat specimens from acute-phase patients were analyzed by the combination of RT-PCR, cloning, and sequencing of the products. Our previous study showed that the RT-PCR test could
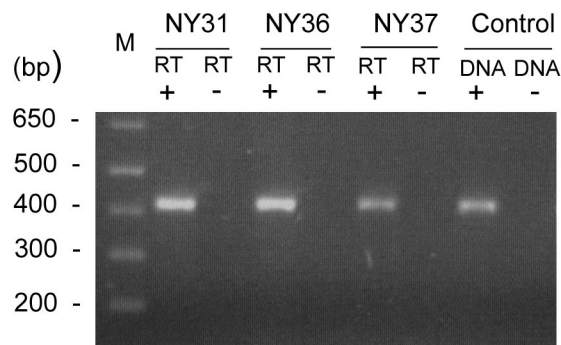
detect approximately one *p44* transcript per ehrlichial organism and that the random selection of transcripts of PCR products was not biased toward any particular cDNA sequence (24). By RT-PCR, approximately 420-bp cDNA fragments of *p44* transcripts were amplified from the total RNAs that were prepared from buffy coats of the initial blood specimens from the three HGE patients (Fig. 2). The RT-PCR product was observed as a broad band in the gel stained with ethidium bromide due to the expression of multiple *p44* paralogs with hypervariable regions of different lengths and the fact that the primer pair used in the RT-PCR anneals with the 5′- and 3′-end conserved regions. No amplicon was detected without reverse transcriptase, indicating the absence of contamination by genomic DNA in the RNA preparation (Fig. 2).

The amplified cDNAs from different *p44* transcripts could be distinguished by the sequences of their central hypervariable regions (23). Forty-nine or 50 cDNA clones derived by cloning the amplicon in each blood specimen from the three patients were randomly selected and sequenced. By this analysis, 10, 17, and 6 different *p44* transcripts were expressed by *A. phagocytophila* in the buffy coat from patients NY31, NY36, and NY37, respectively (Fig. 3). Because several cDNA clones from these three patients had identical sequences, a total of 26 different *p44* transcripts composed of seven sequences (P44-1, P44-7, P44-10, P44-12, P44-13, P44-14, and P44-15) that were previously identified in a 1995 isolate (23) and 19 new sequences (P44-23 to P44-41) were found in the present study. Among these 26 P44 proteins, the highest and the lowest amino acid sequence identities were 70.2 and 33.6%, between the P44-13 and the P44-33 and between the P44-24 and the P44-35, respectively.

The results of the expression analysis for individual patients are summarized in Fig. 4. We define dominant expression as representing greater than 10% of the 49 or 50 cDNA clones examined for each patient's sample. Five of 10 transcripts (P44-23, P44-15, P44-1, P44-7, and P44-14) were dominantly expressed in patient NY31. In patient NY36, three of 17 *p44* transcripts (P44-28, P44-29, and P44-30) were dominantly expressed. One of six transcripts (P44-28) was extremely domi-
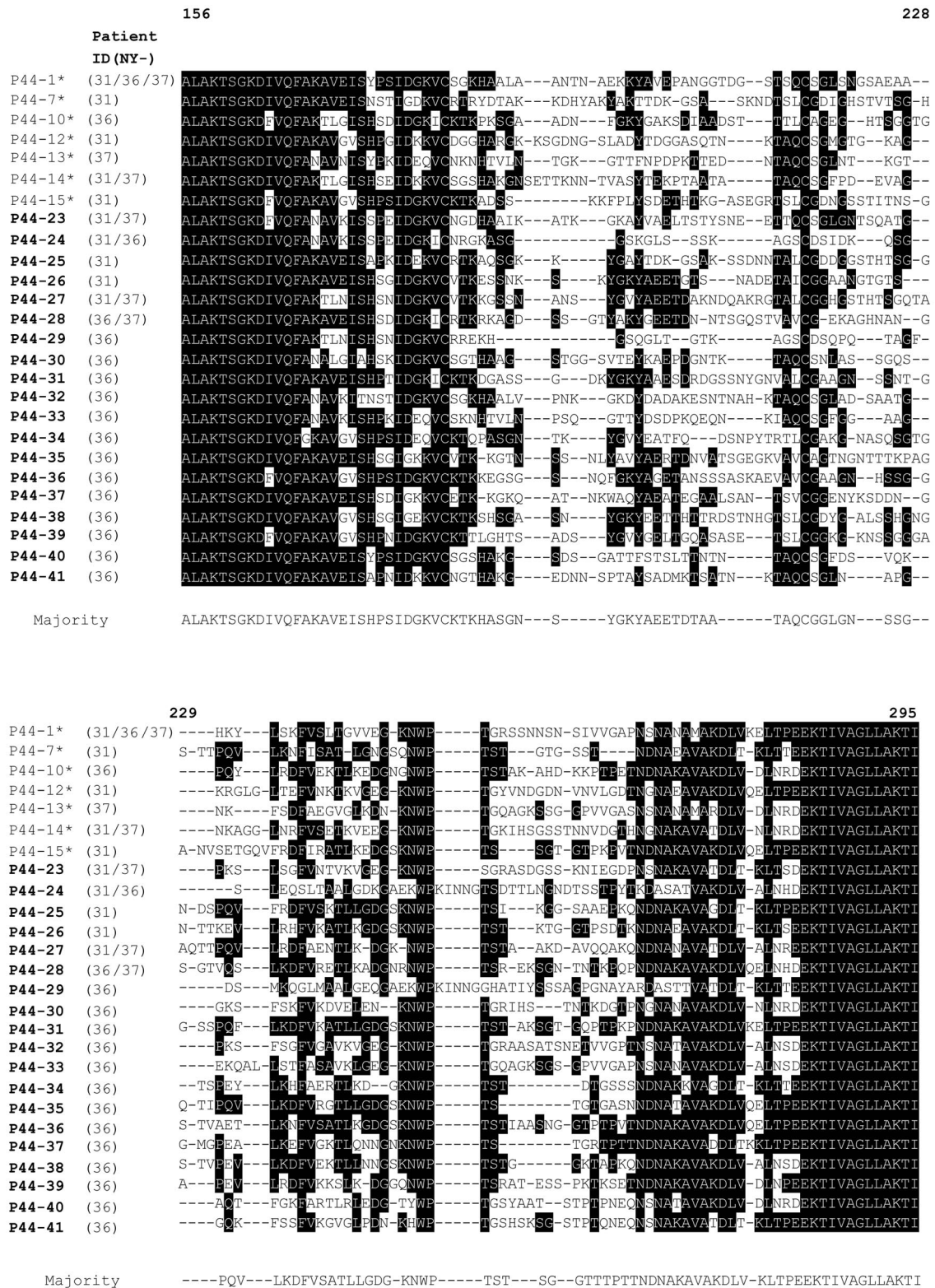
FIG. 3. Alignment of deduced amino acid sequences of *p44* cDNA clones in peripheral blood leukocytes from three HGE patients. Aligned positions of identical amino acids are indicated in white against a black background, and the amino acid found at that position in the majority of the clones is indicated below. Gaps are indicated by dashed lines. The numbers at the top of the sequence indicate the corresponding position of the amino acid residues in the P44-1 protein. The P44 proteins in boldface and with asterisks are sequences identified in the current and previous (23, 24) studies, respectively.
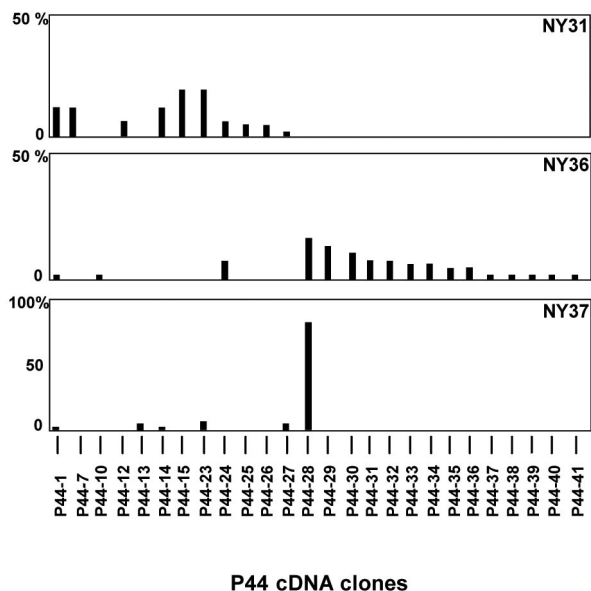
FIG. 4. Relative frequencies of *p44* transcripts in *A. phagocytophila* in three HGE patients. The vertical axis shows relative frequencies of each *p44* cDNA clone detected among 49 (NY31) or 50 (NY36 and NY37) cDNA clones. The horizontal axis indicates the *p44* cDNA clones.

nant in patient NY37, and the percentage of the cDNA clones (42 of 50) was 84%. Thus, the transcript species in patients NY31 and NY36 were more diverse than those in patient NY37.

We analyzed the phylogenetic relationship among the 26 different *p44* transcripts expressed in these three HGE patients
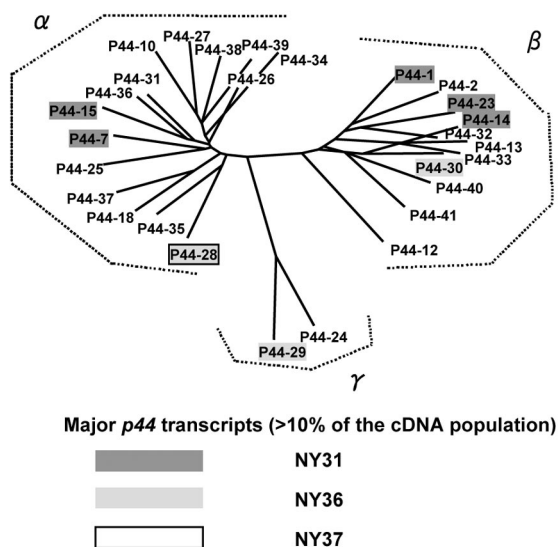

FIG. 5. Phylogenetic analysis of deduced amino acid sequences of 26 different *p44* transcripts found in three patients and the corresponding regions of P44-2 and P44-18. Amino acid sequences were aligned with the Clustal V program, and the tree was constructed with the neighbor-joining method. Boldfacing and shading are used to mark the transcripts that were predominantly expressed in each HGE patient. P44-28 was detected in both patients NY36 and NY37.

and P44-18 and P44-2 based on the predicted amino acid sequence similarities of the central hypervariable regions. These P44 proteins could be divided into three groups (Fig. 5). Group α comprised 15 different P44 proteins, including P44-18, which is the mammalian P44 dominant in the HZ strain (22). Group β consisted of 11 proteins, including P44-2, which is constitutively expressed in the HZ strain in cell culture, mammals, and ticks (24). Group γ consisted of two proteins. Thus, the predominantly expressed P44s in each patient were not segregated by phylogenetic group.

Genetic analysis of *A. phagocytophila* isolates based on *p44* transcripts. The absence of particular transcripts in some patients may be due to the absence of the corresponding genes in the genome of the particular *A. phagocytophila* strain infecting those patients. To compare strain diversity with respect to *p44* multigenes, we analyzed the genetic relationship among five isolates of *A. phagocytophila* by genomic Southern blotting based on major *p44* transcripts detected by RT-PCR in the three patients' blood.

For preparation of *p44* cDNA-specific probes, we chose *p44-23*, which was predominantly expressed in patient NY31, and *p44-28*, which was predominantly expressed in patients NY36 and NY37, as representatives of major *p44* transcripts. Genomic DNAs were prepared from the three new *A. phagocytophila* isolates (NY-31, NY-36, and NY-37) and two *A. phagocytophila* strains (HZ and LL) isolated in 1995 (19, 21). Southern blot analysis revealed that overall patterns among these five *A. phagocytophila* strains were similar regardless of whether the *p44-23* or *p44-28* probe was used, suggesting that all *A. phagocytophila* isolates tested have these genes at identical genetic loci (Fig. 6).

With the *p44-23* probe, hybridization patterns were almost identical between isolate NY-31 and strain LL. The patterns were identical among isolate NY-36, isolate NY-37, and strain HZ: three bands were detected in three restriction digestions of the genomic DNA from isolate NY-31 and strain LL, but two bands were found in isolates NY-36 and NY-37 and strain HZ (Fig. 6A). With the *p44-28* probe, a single major band was detected in three restriction digestions of genomic DNA from four *A. phagocytophila* strains (isolates NY-31 and NY-37 and strains HZ and LL) (Fig. 6B). A single major band was also detected by *Eco*RI digestion of genomic DNA from isolate NY-36, but two bands were detected in both *Pst*I and *Xba*I digestions. The density of the band in the *Eco*RI digest of isolate NY-36 was approximately twofold greater than that of the bands in the *Pst*I and *Xba*I digests. This suggests that the *Eco*RI fragment contains two copies of the *p44-28* gene (Fig. 6B). Because the restriction enzymes used to digest *A. phagocytophila* genomic DNA do not cut within either of these *p44* genes, the multiple bands may represent the existence of additional copies of *p44-23* and *p44-28* paralogous genes in the genome.

Characterization of hypervariable regions of the *p44* multigene family. We previously identified 21 different sequences of *p44* multigenes by transcriptional analysis of strain HZ in mammals, ticks, and cell culture and by genomic DNA cloning (23, 24). In the present study, we found 19 new sequences of *p44* cDNAs from three HGE patients. A total of 43 different sequences, including three additional sequences found by others (10, 14), allowed us to characterize the hypervariable re-
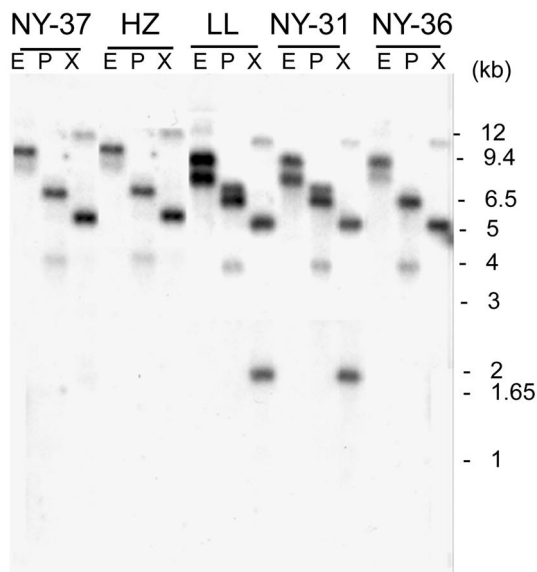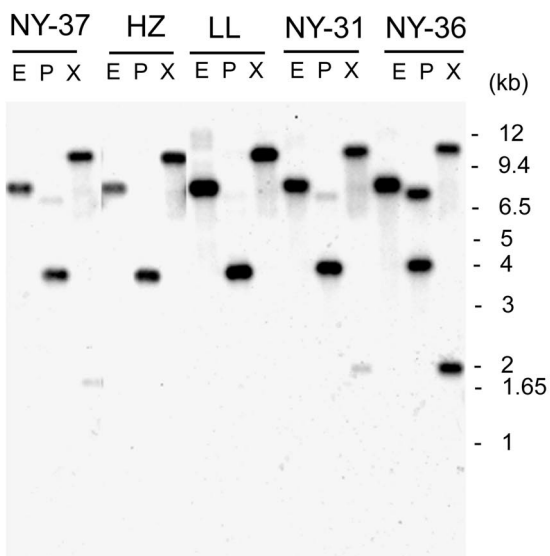
(A) *p44-23* probe

(B) *p44-28* probe

FIG. 6. Genomic Southern blot analyses of five *A. phagocytophila* strains with *p44* cDNA probes. Three isolates (NY-31, NY-36, and NY-37) and two strains (HZ and LL) are shown at the top of the panels. Panels A and B show hybridization patterns with the cDNA probes *p44-23* and *p44-28*, respectively. Numbers on the right indicate molecular sizes.

gions of the *p44* multigene family. The amino acids present at each position in 43 aligned sequences are shown in their order of frequency (Fig. 7). By this analysis, three major domains (D1, D2, and D3) with highly variable amino acid sequences were found within the central hypervariable region. These domains were partitioned by absolutely conserved amino acid residues (C, W, P, and A, as shown by arrowheads in Fig. 7). Computer analysis suggests that these three domains are highly

antigenic and that two domains (D1 and D3) are hydrophilic. Four regions with a high probability of being on the surface were identified, two of which are represented in the majority of amino acid sequences as KYAEETDTKGSNK in domain 1 and TGETN in domain 3 (Fig. 7). The remaining two predicted surface-exposed regions contain conserved amino acids and have their majority amino acid sequence GSKNWPT between D2 and D3 and TPEE at the 3′ end of the central hypervariable region (Fig. 7).

DISCUSSION

The present study revealed for the first time in situ *p44* gene expression of *A. phagocytophila* in patients with HGE. Two types of major surface antigen gene expression patterns were seen. The first is the expression of diverse *p44* paralogs with no single majority. The other is single dominant gene expression. The results suggest that the *A. phagocytophila* strain infecting patient NY37 was antigenically and phenotypically more homogeneous than the *A. phagocytophila* strain infecting patient NY31 or NY36. Patient NY37 had the lowest IFA antibody titer, 320, in contrast to a titer of >2,560 in the other two patients. This may be due to stimulation by a smaller variety of *p44hv*s, or perhaps more diverse *p44* gene expression allows *A. phagocytophila* to persist in patients, stimulating better host humoral immune responses. Large studies will be needed to understand the relationship between *A. phagocytophila* diversity and the immune response.

Seven transcripts previously identified in the 1995 isolate HZ were detected in buffy coats from these three patients. Southern blot analysis revealed that additional two *p44* genes found in the present study were also present in the two 1995 isolates. Although *p44-18*, which is dominantly expressed in the acute stages of experimentally infected mice and horses (24), was not detected in buffy coats from the three HGE patients, *p44-18* was detected by Southern blotting in all three new isolates at the same loci as in strain HZ (data not shown). These results and the existence of several common transcripts among the three patients suggest that *p44* genes are not hypermutatable and that the genetic loci are conserved (except for potentially duplicated genes) in nature. This suggests that in Westchester County, the population of *A. phagocytophila* that can establish infection and cause disease in humans might not have shifted drastically with respect to major surface antigen genes over the 5-year period between the summers of 1995 and 2000.

These results also indicate that *A. phagocytophila* maintains a larger repertoire of P44 proteins than was previously estimated based on Southern blot analysis with a probe specific to the N- and C-terminal conserved sequences (23), suggesting that one N- and C-terminal region can associate with more than two different central hypervariable regions. Analysis of the predicted amino acid sequences of 43 different P44 proteins (including the 19 new sequences found in the present study) revealed that five amino acids were absolutely conserved, subdividing the region into five domains, including three extremely hypervariable central domains. These three domains likely provide surface phenotype diversity. For example, one of these three domains was flanked by two absolutely
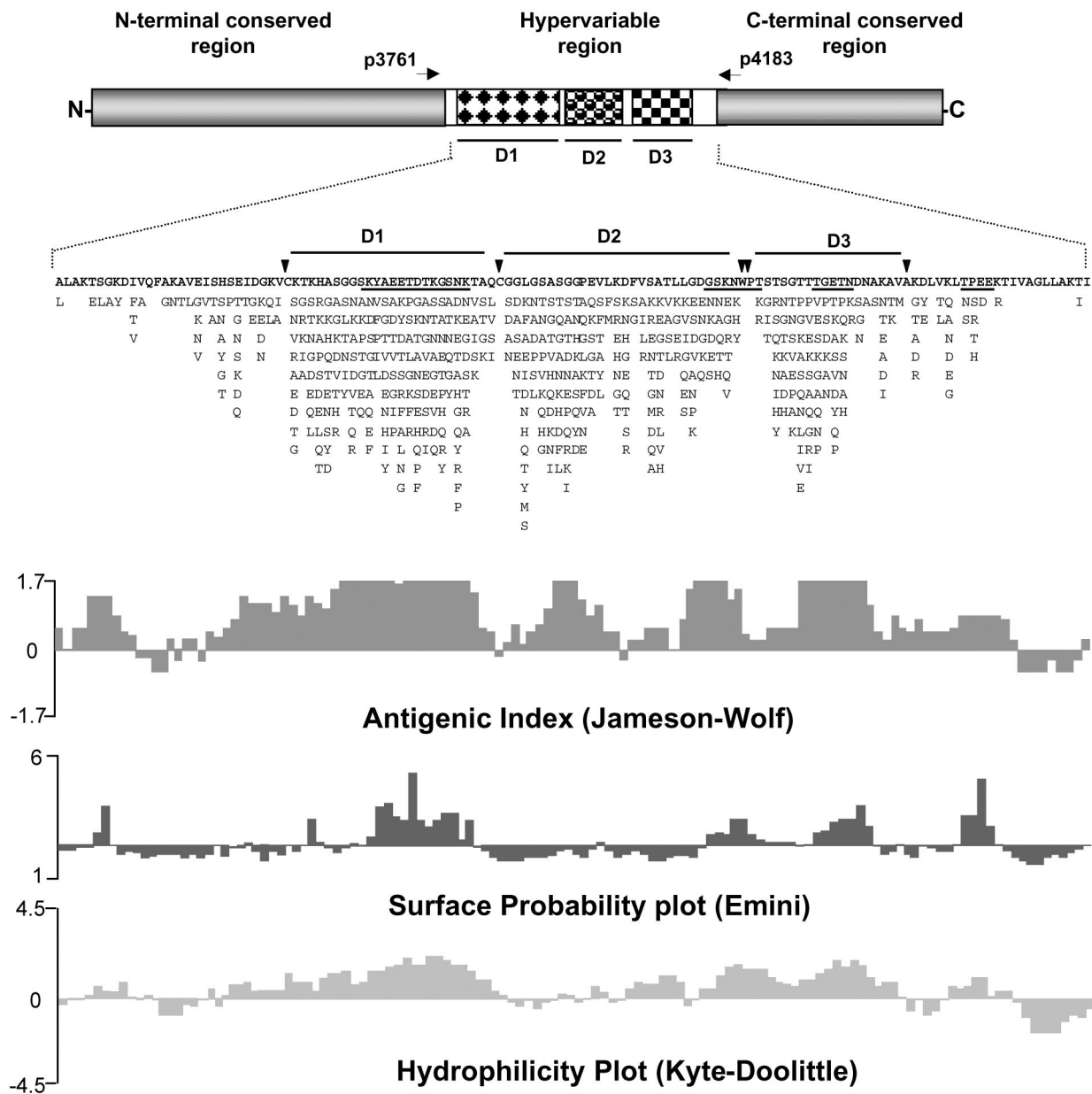
FIG. 7. Characteristics of the central hypervariable region of the *p44* multigene family. The deduced amino acids at each position in 43 different aligned sequences are shown. The amino acid found most frequently at each position is shown above the alignment. Other amino acid residues found at each position are listed in decreasing order of frequency from top to bottom. Vertical arrowheads indicate highly conserved amino acids (C, W, P, and A). Three major hypervariable domains inside the hypervariable region are marked D1, D2, and D3. Four underlined amino acid sequences are regions of high surface exposure probability. The forward and reverse primers used for RT-PCR in the present study are shown as p3761 and p4183, respectively, with horizontal arrows. Representative antigenic index, surface probability, and hydrophilicity plot were determined based on the consensus amino acid sequence with the Protean program in DNAStar.

conserved cysteines, and thus a disulfide bond may create a loop exposing this region of high surface exposure probability and hydrophilicity to the bacterial surface. The four regions of high surface exposure probability, hydrophilicity, and high antigenic index may be exploited in designing P44-based vaccines in the future.

Anaplasma marginale, a tick-borne pathogen that persistently infects red blood cells of cattle, is phylogenetically closely related to *A. phagocytophila*. The major surface protein

2s (MSP2s) of *A. marginale* (23) are orthologous to the P44 proteins of *A. phagocytophila*. MSP2s consist of a central hypervariable region and N- and C-terminal conserved regions. They are encoded by a multigene family (3–5, 8). However, our alignment analysis of sequences of all MSP2s available at the GenBank data base showed that the hypervariable region of MSPs lacks the feature described for P44s in the present study.

Southern blotting with cDNA probes specific to two new *p44* transcripts expressed in patients NY31, NY36, and NY37 sug-

gests that these two genes exist in the genomes of all five *A. phagocytophila* strains used in this study and that the genes are present in two copies in some isolates. Whether these duplicate copies are related to the unique expression site described for *msp2* (3, 4) needs further evaluation. Unfortunately, we have not been able to determine the expression site in *A. phagocytophila* in patients because of the difficulty in obtaining sufficient amounts of purified genomic *A. phagocytophila* DNA and RNA from patients to analyze the site.

REFERENCES

1. **Aguero-Rosenfeld, M. E., F. Kalantarpour, M. Baluch, H. W. Horowitz, D. F. McKenna, J. T. Raffalli, T-C Hsieh, J. Wu, J. S. Dumler, and G. P. Wormser.** 2000. Serology of culture-confirmed cases of human granulocytic ehrlichiosis. J. Clin. Microbiol. **38:**635–638.
2. **Bakken, J. S., J. Krueth, C. Wilson-Nordskog, R. L. Tilden, K. Asanovich, and J. S. Dumler.** 1996. Clinical and laboratory characteristics of human granulocytic ehrlichiosis. JAMA **275:**199–205.
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., J. Yi, A. Lundgren, B. R. McEwen, E. F. Blouin, and K. M. Kocan.** 2001. Antigenic variation of *Anaplasma marginale*: major surface protein 2 diversity during cyclic transmission between ticks and cattle. Infect. Immun. **69:**3057–3066.
5. **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.
6. **Chen, S. M., J. S. Dumler, J. S. Bakken, and D. H. Walker.** 1994. Identification of a granulocytotropic *Ehrlichia* species as the etiologic agent of human disease. J. Clin. Microbiol. **32:**589–595.
7. **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.
8. **French, D. M., T. F. McElwain, T. C. McGuire, and G. H. Palmer.** 1998. Expression of *Anaplasma marginale* major surface protein 2 variants during persistent cyclic rickettsemia. Infect. Immun. **66:**1200–1207.
9. **Hardalo, C. J., V. Quagliarello, and J. S. Dumler.** 1995. Human granulocytic ehrlichiosis in Connecticut: report of a fatal case. Clin. Infect. Dis. **21:**910–914.
10. **Ijdo, J. W., W. Sun, Y. Zhang, L. A. Magnarelli, and E. Fikrig.** 1998. Cloning of the gene encoding the 44-kilodalton antigen of the agent of human granulocytic ehrlichiosis and characterization of the humoral response. Infect. Immun. **66:**3264–3269.
11. **Kalantarpour, F., I. Chowdhury, G. P. Wormser, and M. E. Aguero-Rosenfeld.** 2000. Survival of the human granulocytic ehrlichiosis agent under refrigeration conditions. J. Clin. Microbiol. **38:**2398–2399.
12. **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.
13. **Kim, H. Y., and Y. Rikihisa.** 2000. Expression of interleukin-1β, 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.
14. **Lodes, M. J., R. Mohamath, L. D. Reynolds, P. McNeill, C. P. Kolbert, E. S. Bruinsma, D. R. Benson, E. Hofmeister, S. G. Reed, R. L. Houghton, and D. H. Persing.** 2001. Serodiagnosis of human granulocytic ehrlichiosis by with novel combinations of immunoreactive recombinant proteins. J. Clin. Microbiol. **39:**2466–2476.
15. **Ohashi, N., N. Zhi, Y. Zhang, and Y. Rikihisa.** 1998. Immunodominant major outer membrane proteins of *Ehrlichia chaffeensis* are encoded by a polymorphic multigene family. Infect. Immun. **66:**132–139.
16. **Petrovec, M., S. L. Furlan, T. A. Zupanc, F. Strle, P. Brouqui, V. Roux, and J. S. Dumler.** 1997. Human disease in Europe caused by a granulocytic *Ehrlichia* species. J. Clin. Microbiol. **35:**1556–1559.
17. **Rikihisa, Y., S. A. Ewing, J. C. Fox, A. G. Siregar, F. H. Pasaribu, and M. B. Malole.** 1992. Analyses of *Ehrlichia canis* and a canine granulocytic *Ehrlichia* infection. J. Clin. Microbiol. **30:**143–148.
18. **Rikihisa, Y., S. A. Ewing, and J. C. Fox.** 1994. Western immunoblot analysis of *Ehrlichia chaffeensis*, *E. canis*, or *E. ewingii* infections in dogs and humans. J. Clin. Microbiol. **32:**2107–2112.
19. **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.
20. **van Dobbenburgh, A., A. P. van Dam, and E. Fikrig.** 1999. Human granulocytic ehrlichiosis in Western Europe. N. Engl. J. Med. **340:**1214–1216.
21. **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.
22. **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.
23. **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.
24. **Zhi, N., N. Ohashi, T. Tajima, J. Mott, R. W. Stich, D. Grover, S. R. Telford, I. I. I., 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.