Multiple p44 Genes Encoding Major Outer Membrane Proteins Are Expressed in the Human Granulocytic Ehrlichiosis Agent* (Received for publication, February 8, 1999, and in revised form, March 31, 1999)

Ning Zhi, Norio Ohashi, and Yasuko Rikihisa†

From the Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, Columbus, Ohio 43210-1093

Human granulocytic ehrlichiosis (HGE), a tick-borne zoonosis, was first reported in 1994 and increasingly recognized in the United States (1–5). Serologic and PCR analyses suggest that HGE also exists in Europe (4–6). HGE is characterized by chills, headache, myalgia, and hematological abnormalities, including leukopenia and thrombocytopenia. It frequently requires prolonged hospitalization, and when the treatment is delayed due to misdiagnosis or in immunocompromized patients, HGE can be fatal (7). HGE is caused by infection with an obligate intracellular bacterium, HGE agent. Comparison of 16 S rRNA gene sequences among these P44-homologous proteins, whether genes corresponding to the mRNAs from the genomic DNA of the HGE agent; (iii) showed that the genes being expressed were not clustered in the HGE agent genome; (iv) estimated that a minimum copy number of the p44-homologous genes in the genome is 18; (v) detected two different P44-homologous proteins expressed by the HGE agent; and (vi) demonstrated existence of antibodies specific to the two proteins in sera from patients with HGE. These findings showed that p44 multigenes have several active expression sites and the expression is regulated at transcriptional level, suggesting a potentially unique mechanism for generating the diversity in major antigenic outer membrane proteins of the HGE agent. Characterization of p44-homologous genes expressed by the HGE agent in a tissue culture would assist in understanding a role of the p44 multigene family in pathogenesis and immune response in HGE.

Human granulocytic ehrlichiosis (HGE), a tick-borne zoonosis, was first reported in 1994 and increasingly recognized in the United States (1–5). Serologic and PCR analyses suggest that HGE also exists in Europe (4–6). HGE is characterized by chills, headache, myalgia, and hematological abnormalities, including leukopenia and thrombocytopenia. It frequently requires prolonged hospitalization, and when the treatment is delayed due to misdiagnosis or in immunocompromized patients, HGE can be fatal (7). HGE is caused by infection with an obligate intracellular bacterium, HGE agent. Comparison of 16S rRNA gene sequences among these P44-homologous proteins suggests the existence of additional p44-homologous genes in the HGE agent genome. By Western blot analysis, a mouse antiserum against the recombinant P44 (rP44) protein recognized two to six P44-homologous proteins in each of the five isolates of the HGE agent and in USG3 isolate. Three monoclonal antibodies that react to the rP44 recognized one to four 44-kDa-range proteins in the whole organism as well as in the outer membrane fraction from these six isolates (14). Passive immunization with these antibodies induced a partial protection against infection with the live HGE agent in mice. These studies suggest that multiple antigenically cross-reactive proteins of 38–49 kDa are expressed in each isolate of the HGE and USG3 isolate in HL-60 cell culture and are potential protective antigens in HGE infection.

Murphy et al. (15) cloned three p44-homologous genes (msp-2a, msp-2b, and msp-2c) of USG3 isolate. The chemically determined amino acid sequences at N termini or internal segments of native 43- and 45-kDa proteins of the isolate approximately match with the segments of amino acid sequences predicted from one to two each of the three cloned genes. However, due to existence of highly conserved amino acid sequences among these P44-homologous proteins, whether these genes are actually expressed by the isolate was not conclusive in that study. Ijdo et al. (16) cloned a p44-homologous gene (hge-44) from the NCH-1 isolate of HGE agent and showed by RT-PCR and sequencing the product that this gene is expressed by the HGE agent in HL-60 cells.

So far, little is known about the relationship between the...
diversity of antigenically cross-reactive proteins of 38–49 kDa and multiple p44-homologous genes in the HGE agent. In this study, we characterized the structure, distribution, and expression of the p44-homologous genes of the HGE agent cultivated in HL-60 cells. Thereby, the study is expected to facilitate understanding of a role of the p44 multigene family in pathogenesis and immune responses in HGE infection and also to be helpful in designing a vaccine candidate by using these gene products.

**Experimental Procedures**

**Culture and Purification of the HGE Agent**—The HGE agent (HZ isolate) was cultivated in HL-60 cells (human promyelocytic leukemia cell line) and purified by the Sephacryl S-1000 chromatography method as described elsewhere (17).

**RT-PCR and Cloning of the cDNA**—A pair of oligonucleotides used for RT-PCR (p3708 and p4257 as shown in Fig. 1 and Table I) was designed based on the conserved regions between DNA sequences of the p44 gene and a truncated p44-homologous gene downstream from the p44 gene. Total RNA was extracted from HL-60 cells infected with the HGE agent by using TRIzol reagent (Life Technologies, Inc.). The isolated RNA (3 μg) was heated at 70 °C for 10 min and reverse-transcribed in a 20-μl reaction mixture (0.5 mM deoxynucleoside triphosphate mixture (dNTP), 200 units of SuperScript II reverse transcriptase (Life Technologies, Inc.), 2 pmol of p4257 primer, and 3 mM MgCl₂) at 42 °C for 50 min. PCR was performed in a 100-μl reaction mixture containing 2 μl of the cDNA product, 10 pmol each of p3708 and p4257 primers, 0.2 mM dNTP mixture, 5 units of Taq DNA polymerase, and 1.5 mM MgCl₂ with 3 min of denaturation at 94 °C followed by 30 cycles consisting of 1 min of denaturation at 94 °C, 1 min of annealing at 52 °C, and 2 min of extension at 72 °C. To rule out contamination of DNA in the RNA preparation, RT-PCR without reverse transcriptase was carried out (negative control). The amplified RT-PCR products were cloned in a pCRII vector by using the TA Cloning Kit (Invitrogen Co., San Diego, CA). Twenty-five cDNA clones, which were randomly selected from the transformants, were sequenced by dideoxy chain termination method with an Applied Biosystems 373 DNA sequencer.

**Genomic Southern Blot Analysis**—DNA probes specific to each of the cDNAs for Southern blotting were designed based on a comparison of deduced amino acid sequences among these cDNAs. The central hypervariable regions (approximately 94 amino acid residues) in each cDNA and in the p44 gene were amplified by PCR with primer pairs as shown in Table I. The amplicons were cloned into a pCRII vector. The DNA insert excised from each recombinant plasmid was labeled with [α-32P]dATP by the random primer method with a kit (Amersham Pharmacia Biotech) and used as a probe. Hybridization was performed in rapid hybridization buffer (Amersham Pharmacia Biotech) as described elsewhere (13). The membrane was exposed to a Hyperfilm (Amersham Pharmacia Biotech).

**Genomic Cloning of p44-homologous Genes of the HGE Agent**—The XbaI DNA fragments of the HGE agent, which were detected by genomic Southern blot analysis, were inserted into a βlisscript II KS (+) vector, and the recombinant plasmids were introduced into E. coli DH5α. By using the colony hybridization method (18) with the specific DNA probes as those used for Southern blot analysis, three positive clones were isolated, and the DNA inserts were sequenced. The clones were designated pHGE3.0, pHGE3.4, and pHGE3.9, containing ehrlichial DNA fragments of 3.0, 3.4, and 3.9 kb, respectively.

**Northern Blot Analysis**—Total RNA (15 μg) from the HGE agent-infected cells was separated on 1.2% denaturing agarose gel containing 0.22 M formaldehyde and transferred to Hybond-N+ nylon membrane (Amersham Pharmacia Biotech). The [α-32P]CTP-labeled single-stranded RNA probes (Table I) that were specific to each of cDNAs derived from the recombinant plasmids used for Southern blotting were prepared by using the riboprobe in vitro transcription system (Promega Corp., Madison, WI). The hybridization was performed separately with 1 × 10⁶ cpm/ml of each [32P]-labeled RNA probe (a specific activity of 2 × 10⁶ cpm/μg) in Prehyb/Hyb solution (Ambion, Inc., Austin, TX) at 65 °C for 16 h. After being washed twice each with low stringency solution and high stringency solution (Ambion) at 65 °C for 15 min, the membranes were exposed to a Hyperfilm. The densities of hybridized bands were analyzed with a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

**Pulsed-field Gel Electrophoresis (PFGE)**—The purified HGE agent

---

**Table I**

| Target genes | Primers | Nucleotide sequences (5'-3') | Regions | PCR products | Usage |
|--------------|---------|-----------------------------|---------|--------------|-------|
| p44-2        | 2hvf    | TACGTGAGCTGATGCAGCAT        | 542–563 | 231          | Southern/Northern blot |
| p44-18       | 18hvf   | GGTGCTTGAACGATTTTGGAGA      | 542–563 | 240          | Southern blot |
| p44-15       | 15hvf   | GAGAGCTTACATCAGGCAGG        | 542–563 | 234          | Southern/Northern blot |
| p44-19       | 19hvf   | CATGCAGGTGACCATGTTAGGG      | 542–564 | 234          | Southern blot |
| p44          | 4hvf    | AGGAGGTTATGAGAGGCCAG        | 542–563 | 234          | Southern blot |
| 3'-end of p44-12 | pmf12  | ATGCGGCCGCATCGTCGACATCGG   | 218–239 | 247          | Southern/Northern blot |
| (p44-12N)    | pmf12   | CAGTTCAGTGACCAAGGCGG       | 478–498 | 280          | Southern blot |
| 5'-end of p44-12 |pcf12  | CTACAGCCGTTAGGAGGGACG       | 880–901 | 280          | Southern blot |
| (p44-12C)    | per12   | GAGAGAGCTGATGACCAAGG        | 1072–1093 | 210         | Southern blot |

a Numbers correspond to the nucleotides in the p44 gene that was cloned in Ref. 34 (GenBank™ accession number AF059181) and the primer positions are shown in Fig. 1.

b Primer pairs are used to amplify hypervariable regions of each of cDNAs.

Probes p44-12N and p44-12C are specific to 5' and 3'-end conserved regions of p44 homologous genes.
Expressed p44 Genes in the HGE Agent

Fig. 2. Alignment of amino acid sequences deduced from cDNA clones and the corresponding p44-homologous genes of the HGE agent. Aligned positions of identical amino acids with P44 of the HGE agent are shown with dots. Gaps indicated by dashed lines were introduced for an optimal alignment of all proteins. A boxed area in the middle indicates the amino acid sequences deduced from nucleotide sequences of cDNAs. Hypervariable regions are shown in boldface. A bar indicates the N-terminal amino acid sequence of the native P44 protein, and an arrowhead shows the cleavage site of the putative signal peptide. The amino acid sequences underlined in the hypervariable regions of P44-2 and boxed area P44-18 indicate the sequences that were used to prepare synthetic oligopeptides, Pep2 and Pep18, respectively. The arrows point out the positions of the putative signal peptide.

Table II

Properties of p44-homologous genes of the HGE agent

| Genes | Coding region | Amino acid no./AL | Copy number | No. of isolated cDNA clones* | Universal start codon | SD/Putative promoter (-10) |
|-------|---------------|------------------|-------------|-----------------------------|-----------------------|---------------------------|
| p44-2 | bp            | 1275             | 2           | 2                           | +                      | AGGA/G/TATTAT             |
| p44-18| 759           | 253/26,400       | 5           | 2                           | +                      | NG/NG                     |
| p44-12| 1173          | 391/41,179       | 3           | 2                           | +                      | AGGA/G/TATTAT             |
| p44-14| 834           | 278/29,287       | 3           | 2                           | +                      | NG/NG                     |
| p44-19| ND            | ND               | 3           | 2                           | +                      | AGGA/G/NG                 |
| p44c  | 1173          | 391/41,293       | 1           | +                           |                        | AGGA/G/TATTAT             |
|       | 225           | 759/9,026        | ND          | ND                          | +                      | AGGA/G/NG                 |

* A total of 25 cDNA clones were analyzed.

Western Immunoblotting and Indirect Fluorescent Antibody (IFA) Labeling—To examine the existence of p44-homologous gene products in the HGE agent organisms, Western immunoblot analysis and IFA were performed by using antisera against synthetic oligopeptides specific to each P44-homologous proteins. Briefly, by using DNASTAR program (DNASTAR Inc., Madison, WI), we analyzed the antigenic index, surface probability, and hydrophilicity and selected the unique amino acid sequence from the hypervariable region of P44-2 and P44-18 proteins. Two oligopeptides with sequences of GHSSGVTQNPKLFST and KNQKSSDTDTGVEKA were synthesized (Alpha Diagnostic, San Antonio, TX) and named Pep2 and Pep18, respectively (Fig. 2). Antisera against these synthetic oligopeptides were generated by immunization of a rabbit and a mouse with keyhole limpet hemacyanin (KLH) (Pierce)-conjugated synthetic oligopeptides Pep2 and Pep18, respectively. Western immunoblot analysis and IFA test with the above sera were performed as described elsewhere (14, 15). The mouse anti-recombinant P44-18 fragment was used as a positive control. For double IFA staining, a lissamine rhodamine-labeled goat anti-rabbit IgG and fluorescein isothiocyanate-labeled goat anti-mouse IgG (Jackson Immunonotechnology, West Grove, PA) were used as secondary antibodies.

Dot Immunoblot Assay—The assay was carried out as described elsewhere (13). The KLH-conjugated oligopeptides Pep2 and Pep18...
RESULTS

mRNAs Transcribed from p44-homologous Genes—By RT-PCR, approximately 550-bp cDNAs were amplified from the total RNA of the HGE agent in HL-60 cells. The amplicon was observed as a single band in the gel by ethidium bromide, but no amplicon was detected without a reverse transcriptase (data not shown), indicating the absence of contamination of genomic DNA in the RNA preparation. After cloning the amplicons, only five different nucleotide sequences were found in the 25 randomly selected cDNA clones, and the deduced amino acid sequences are shown in Fig. 2 (boxed area). The five sequences of the cDNAs were represented as P44-2, P44-18, P44-12, P44-15, and P44-19. The numbers of cDNA clones with nucleotide sequences identical to P44-2, P44-18, P44-12, P44-15, and P44-19 were 12, 5, 3, 3, and 2, respectively (Table II). One or two substitutions were found at the 3'-end of three nucleotide sequences in a set of 12 cDNA clones represented by P44-2. We assume that these substitutions were generated by nucleotide misincorporation with reverse transcriptase or Taq polymerase. No cDNA sequence identical to the p44 gene was found in these 25 cDNA clones. This indicates that at least five different mRNAs from p44-homologous genes are transcribed by the HGE agent cultivated in HL-60 cell.

A comparison of the deduced amino acid sequences among the five different cDNAs and p44 gene previously cloned (13) revealed that a central region of approximately 94 amino acid residues corresponding to the 175th to 269th amino acid sequence of a protein (P44) encoded by the p44 gene was hypervariable, and the flanking regions of approximately 30 residues each were highly conserved (Fig. 2). Within the hypervariable region, the highest amino acid sequence similarity was 32.8%, between the P44-2 and P44-18 proteins, and the lowest similarity was 19.9%, between P44-15 and P44 proteins. In comparison with flanking regions, the hypervariable region had higher hydrophilicity and antigenic index.

Southern Blotting and Genomic Cloning of the p44-homologous Genes Expressed by the HGE Agent—By using a DNA probe specific to each of cDNAs of P44-18, P44-12, P44-19, and the p44 gene, a single DNA band was detected in all restriction digestions tested (Fig. 3). Because restriction enzymes used do not cut within a p44 gene or any cDNA clones, this suggests that each gene corresponding to the four probes is a single copy in the HGE agent genome. However, the probes specific to the cDNAs of P44-2 and P44-15 generated two or three bands (Fig. 3), showing that two or three gene copies with sequences identical to each other in the four restriction digestions. This is consistent with the result of our sequence analysis data that the p44 gene and a p44-homologous gene (p44-12) without a universal start codon (AUG) were overlapped by 21 bp in an XbaI fragment in PGE1221 plasmid (Fig. 4). Overall, at least nine copies of p44-homologous genes (a total of the copies of genes identified in Table II) either expressed and/or highly homologous were detected in the genome by using cDNA-specific probes.

Three DNA fragments of 3.0, 3.4, and 3.9 kb were cloned from the XhoI-digested genomic DNA of the HGE agent with the probes specific to the cDNAs of P44-2, P44-12, and P44-15 (Fig. 4). Sequencing of the 3.9- and 3.0-kb fragments revealed two complete ORFs of 1275 and 1173 bp encoding 425- and 391-amino acid proteins with molecular weights of 44,969 and 41,179, respectively. As expected, these ORFs (p44-2 and p44-12) contained the ORFs identical to the cDNAs of P44-2 and P44-12, respectively (Figs. 2 and 4). In the 3.9-kb DNA, an additional small ORF (p44-c), which had a DNA sequence corresponding to 82 amino acid residues (9062 Da) at the C terminus of the p44-2 protein, was found at 111 nucleotides downstream from the p44-2 gene. The 3.4-kb DNA fragment contained an ORF of 834 bp encoding a 278-amino acid protein with a molecular weight of 29,387. This ORF (p44-15) included a sequence identical to P44-15 cDNA. The p44-15 did not have a universal start codon and lacked DNA sequence corresponding to 82 amino acid residues at the N terminus of the P44 protein. Consensus sequences of e challenged rich region about 10 base pairs upstream of the transcription start site; the −10 sequence) and Shine-Dalgarno sequence were found in the regions upstream from the start codon of p44-2, p44-12, and p44-c. The properties of the p44-homologous genes identified in this study were summarized in Table II.

The N-terminal amino acid sequence (HDDVSALETG) of the native 44-kDa protein previously determined (13) was also found in P44-2 and P44-12 proteins as shown in Fig. 2. A comparison of deduced amino acid sequences of these cloned genes suggests that the identical amino acid sequence consisting of 15 residues at the N terminus of P44-2, P44-2, or P44-12 protein is a signal peptide. The alignment also showed that the N- and C-terminal portions are highly conserved among three P44 homologs (P44-2, P44-12, and P44), except for the existence of additional 34 amino acid residues at C terminus of P44-2 (Fig. 2). At N and C termini of both P44-14 and P44-18, proteins that are encoded by the ORFs without universal start codons had a short stretch of an amino acid sequence consisting of 8–28 residues without any homology to other P44 homologs, including P44, P44-2, and P44-12 (Fig. 2).

Northern Blot Analysis—To confirm whether the relative proportion of the cDNA clones represented different levels of transcription of the p44-homologous genes and to determine the transcript sizes, the Northern blotting was performed with
three RNA probes. The probe specific to P44-2 and P44-18 cDNAs strongly hybridized with the respective 1.5-kb mRNA transcript (Fig. 5). Using same amounts of total RNA and each probe, the density of band hybridized with P44-2-specific probe was at least 2-fold greater than that with P44-18-specific probe. No hybridized band was observed when p44-specific probe was used. These findings support the cDNA cloning results that, among the p44-homologous genes, p44-2 was more abundantly expressed in the HGE agent than p44-18, and the p44 gene was silent.

Expression Sites of p44-homologous Genes in the HGE Agent Genome—In a PFGE gel, EagI and BamHI digestions produced 14 and 23 bands ranging from 9 to 145 kb and from 2 to 194 kb, respectively (data not shown). In the Southern blotting (Fig. 6), the probes hybridized to larger fragments of 100, 97, 49, and 12 kb in EagI digestion and 50, 20 and 18 kb in BamHI digestion. These results show that these five different expressed genes did not form a cluster in the genome of the HGE agent.

Copy Number of p44-homologous Genes—In the genomic Southern blot analysis (Fig. 7), the hybridization patterns of major bands with p44-12N and p44-12C probes (no homology each other) in two restriction digestions were almost identical to each other, except that two restriction fragments of 2.2 and 1.5 kb in XbaI and three restriction fragments of 2.0, 1.7, and 1.2 kb in PstI hybridized only with either p44-12N or p44-12C probes (Fig. 7, asterisks). This suggests that two and three copies of p44-homologous genes lack the 3'- or 5'-end conserved regions, respectively. The restriction sites of these enzymes were not found in the DNA sequences of nine p44-homologous genes so far cloned (five in the present study, three in Ref. 15, and one in Ref. 16). However, one or two p44-homologous genes that have not been cloned may be digested by the restriction enzymes. Because any sequence divergences between the probes and target p44-homologous genes reduce the hybridization signal, ratios of densities of the bands to the 3.0-kb band approximately represent a minimum copy number of p44-homologous genes in each of these hybridization bands. Therefore, we concluded that there are at least 18 of p44-homologous genes in the HGE agent genome. Because by genomic Southern blotting with cDNA-specific probes, only nine copies of p44-homologous genes were detected (Fig. 3 and Table II), at least nine more copies with sequences in their hypervariable regions divergent from any of those expressed genes appear to exist in the HGE agent genome.

Analysis of HGE Agent Organisms Expressing P44-homologous Proteins—The rabbit anti-Pep2 and the mouse anti-Pep18 sera reacted with a single band of 44 and 43 kDa, respectively, in the HGE agent lysate (Fig. 8). In agreement with our RT-PCR and Northern blot results, the size of the protein recognized by the anti-Pep18 serum was not 26.4 kDa, as was
cDNA clones, we suggest that multiple p44-18 antigens (Fig. 10). Serum 7 reacted strongly to both Pep2 and Pep18. Serum 11 also reacted strongly to Pep18 but weakly to Pep2. This result indicates that p44-2 and p44-18 genes are actually expressed by the HGE agent in infected patients; thus, antibodies specific to these P44 homologs were developed.

**Phylogenetic Analysis**—A phylogenetic analysis revealed nine P44 homologs had identities of 58.3–99.5% (Table III) and made a cluster separated from two major surface protein 2 (MSP2) (2-11.2 and 2-DF (20, 21)) and a MSP4 (22) of *Anaplasma marginale* (Fig. 11). The result is in agreement with phylogenetic analysis based on 16 S rRNA sequence comparison. This also indicates that P44 homologs of HGE and USG3 isolates were not segregated either by geography (e.g. New York versus Massachusetts) or host origin (e.g. human versus tick).

**DISCUSSION**

This study demonstrated that multiple p44 genes are expressed by the HGE agent in HL-60 cells and probably in patients, which may be the reason why multiple bands of 38–49 kDa were found in our previous Western blot analysis studies with patients’ sera (12) or with monoclonal antibodies (14). Our data set the framework for a better understanding of the mechanism(s) underpinning the expression of p44 multigenes in the HGE agent. Four of expressed p44-homologous genes (p44-2, p44-12, p44-15, and p44-18) and one silent gene (p44) from the genomic DNA were cloned. The proteins encoded by these p44-homologous genes were found to consist of a single central hypervariable region of approximately 94 amino acid residues and N- and C-terminal regions highly conserved among the homologs. It appears that two kinds of mechanisms may involve in the expressions of p44-homologous genes: (i) the normal expression of the genes, such as p44-2 and p44-12, with complete ORFs from the respective expression sites, and (ii) the unique expression with a specific event, probably transcriptional modification such as splicing, of the genes, such as p44-18 and p44-15, which lack a universal start codon. The latter mechanism may come to existence to overcome the deficiency in creating antigenic diversity by a common mechanism, such as recombination.

p44 multigenes are homologous to two *A. marginale* msp2 genes that had been cloned (~45% amino acid similarity (13, 15, 16)). Recently, the study on *A. marginale* revealed that

| Protein | Amino acid sequence identity (upper right)/evolutionary distance (lower left) |
|---------|--------------------------------------------------------------------------|
| P44 (HGE agent HZstrain) | 78.2 82.4 59.0 60.7 75.3 78.9 68.7 72.3 38.4 38.5 28.4 |
| P44-2 (HGE agent HZstrain) | 0.20294 80.1 58.3 58.7 80.2 81.9 77.3 70.9 43.0 43.4 29.1 |
| P44-12 (HGE agent HZstrain) | 0.19796 0.22277 59.7 59.5 78.8 99.5 65.3 72.3 39.9 40.2 27.7 |
| P44-15 (HGE agent HZstrain) | 0.50564 0.55686 0.46298 56.7 61.9 59.7 53.6 79.9 33.8 33.1 18.3 |
| P44-18 (HGE agent HZstrain) | 0.51409 0.54133 0.42141 0.41841 61.9 59.5 53.6 59.5 29.4 29.8 17.1 |
| Hge-44 (HGE agent NCH-1 strain) | 0.26919 0.23124 0.23950 0.48470 0.43694 81.4 73.9 75.5 44.0 43.2 30.1 |
| Msp-2a (GE agent USG3 strain) | 0.21239 0.20045 0.00513 0.46298 0.47188 0.20823 76.1 73.6 43.5 43.7 29.4 |
| Msp-2b (GE agent USG3 strain) | 0.26463 0.20803 0.27820 0.43688 0.50621 0.30798 0.24265 49.1 45.1 44.5 23.0 |
| Msp-2c (GE agent USG3 strain) | 0.26633 0.30231 0.24447 0.12056 0.42164 0.25831 0.24810 0.43146 35.7 35.2 21.3 |
| MSP2-11.2 (A. marginale) | 0.84141 0.75118 0.72860 1.06102 1.41654 0.86472 0.73031 0.73031 0.73031 0.73031 0.73031 |
| MSP2-DF (A. marginale) | 0.85587 0.78480 0.74784 1.11780 1.48052 0.86222 0.74511 0.73637 0.73637 0.73637 0.73637 |
| MSP4 (A. marginale) | 1.09759 1.01450 0.98031 1.24793 1.39230 1.06037 1.03435 0.87160 1.12230 1.31835 1.31835 |

Legend:
The numbers in the table indicate the amino acid sequence identity (upper right) and evolutionary distance (lower left). The alignments were performed using ClustalW version 2.1. The amino acid sequence identity was calculated using the formula: (a + b) / (a + b + c), where a is the number of identical amino acid residues, b is the number of amino acid residues that are semantically similar, and c is the number of different amino acid residues. The evolutionary distance was calculated using the Jukes-Cantor model.

**TABLE III**

Identities and evolutionary distances among amino acid sequences predicted from entire genes of HGE agent P44 and its homologs

| Protein | P44 | P44-2 | P44-12 | P44-15 | P44-18 | Hge-44 | Msp-2a | Msp-2b | Msp-2c | MSP2-11.2 | MSP2-DF | MSP4 |
|---------|-----|-------|--------|--------|--------|--------|--------|--------|--------|-----------|----------|------|
| HGE agent HZstrain | 78.2 | 82.4 | 59.0 | 60.7 | 75.3 | 78.9 | 68.7 | 72.3 | 38.4 | 38.5 | 28.4 |
| MSP2-11.2 (A. marginale) | 0.84141 | 0.75118 | 0.72860 | 1.06102 | 1.41654 | 0.86472 | 0.73031 | 0.73031 | 0.73031 | 0.73031 | 0.73031 |
| MSP2-DF (A. marginale) | 0.85587 | 0.78480 | 0.74784 | 1.11780 | 1.48052 | 0.86222 | 0.74511 | 0.73637 | 0.73637 | 0.73637 | 0.73637 |
| MSP4 (A. marginale) | 1.09759 | 1.01450 | 0.98031 | 1.24793 | 1.39230 | 1.06037 | 1.03435 | 0.87160 | 1.12230 | 1.31835 | 1.31835 |
multiple (at least 4) msp2 genes were expressed in each peak of rickettsia that occurred at 6–8-week intervals in two cattle (20). The distribution of the genes in the genome (broad distribution throughout the genome) and the protein structure (e.g., a central hypervariable region flanked with conserved regions) are similar between msp2 genes and p44 multigenes. Because the copy number of msp2 multigenes in A. marginale was estimated as 10 (21), the relative ratio (4 in 10) of expressed...
genes against the total copy number of msp2 multigens at a
given stage of infection is higher than that (5 in 18) of p44
multigens of the HGE agent. In other words, the HGE agent
may have more potential genetic capacity to generate the
diversity of the P44-homologous proteins. Because for A. margi-
nale, the msp2 genes corresponding to the transcripts were not
identified, the expression mechanism is unknown.

Among rickettsia closely related to HGE agent, persistent
infection and recurrence after recovery from the clinical disease
are known for Ehrlichia canis, Ehrlichia platys, E. phagocyto-
phila, Cowdria ruminantium, and A. marginale (20, 23). The
HGE agent was detected in the serum of one untreated patient
by PCR at 30 day after onset of illness (24), suggesting that the
HGE agent may also cause persistent infection in humans. In
addition, persistence of the HGE agent in reservoir rodent host
would be an important adaptation that allows greater access of
uninfected tick populations to an infectious blood meal. French
et al. (20) proposed that the expression of distinct sets of msp2
genes at each rickettsia peak in A. marginale infection of
cattle may allow immunoevasion of anaplasma to persist in
immunocompetent hosts. Multiple expression of p44 multi-
genies may also be related to immunoevasion in human and
rodent hosts and potential persistence.

The expression mechanisms of the several multigene fami-
lies in the human pathogens Neisseria gonorrhoeae (omp (25) and
opa (26)), Borrelia hermsii (vmp (27)), Borrelia burgdorferi
(vls (28)), African trypanosommes (vsg (29)), and Plasmodium
falciparum (var (30)), have been well studied. These multigene
families are involved in pathogenesis, e.g. antigenic variation
(or phase variation) and cytoadhesin. The switching of gene
expression can be divided into two major mechanisms: one
depends on DNA rearrangement, and another occurs at tran-
scriptional level. In pil, vmp, vls, and vsg, switching of expres-
sion between members of the corresponding gene families oc-
curs through programmed DNA rearrangements (gene conversion), moving a transcriptionally silent gene into an active
expression site. In oppa, expression is regulated by a revers-
ible frameshift mutation of DNA, i.e. a slipped-strand mispair-
ing of the number of pentanucleotide coding repeat units in the
single peptides (26), whereas in the var gene of P. falciparum,
each parasite expresses only a single and distinct var gene product. Each var gene is an independent transcription unit in
which promoter activity determines the expression status. P44
multigene expression appears to be different from that of pil,
vmp, vls, and vsg in that multiple expressed p44 genes are
located in several large DNA fragments in the HGE agent
genome, i.e. the transcription of the genes apparently does not
occur from a unique expression site. The expression p44 mul-
tigene, therefore, may not involve DNA rearrangement. Ex-
pressions of p44 multigenes, as well as msp2 genes (21), are also
different from that of var genes, because a single organism
appear to express more than two gene products.

In ehrlichiae, recently, additional multigene families encod-
ing major outer membrane proteins have been discovered. We
identified the omp-1 gene family of Ehrlichia chaffeensis (31)
and the p30 gene family of E. canis (19). More than a dozen
copies of the omp-1 and p30 gene families are tandemly ar-
nanged in the genome, in contrast to p44 multigenes and msp2
genes. Furthermore the protein structures encoded by the
omp-1 and p30 gene families is distinct from those of p44 and
msp2, consisting of three short, hypervariable segments inter-
posed with relatively conserved segments (19, 31). This sug-
gests that the expression mechanism of omp-1 and p30 gene
families is probably different from those of p44 and msp2.
Immunization with a recombinant P28 protein (one of the
omp-1 multigene products) of E. chaffeensis and native MSP2
of A. marginale has been demonstrated to induce almost com-
plete and partial protections against the infection in mice and
cattle, respectively (31, 32). We previously observed that pas-
sive immunization with monoclonal antibodies specific to P44-
related proteins of the HGE agent induced partial protec-
tion against challenge with the HGE agent in mice (14). Active
immunization with the recombinant P44 protein also partially
protected mice from HGE infection.2 This indicates that the
p44 multigene family encodes a potential protective antigen.
However, because only a fraction of genes is differentially ex-
pressed as shown in this study, it seems to be essential to
determine which genes are expressed in the host for identifi-
cation of the most protective antigen.

By a GenBankTM data base homology search, the P44-homol-
ogous proteins of the HGE agent were found to possess a
similarity (~64%) with the N-terminal region of Hsf protein
(surface fibrils) involved in binding of Hemophilus influenzae
type b to human epithelial cells (33). The similar sequence
identified was located in the center of all P44 homologs, in-
cluding the hypervariable domain (span of approximately
200 amino acid residues). This finding suggests that P44s may
have a similar function, such as cytoadhesion to human
granulocytes.

The basic information obtained in this study would facilitate
the understanding of the role of the p44 multigene family in
causing disease in humans, in the persistence of HGE agent
in white-footed mice, and in transmission of this agent from tick
to human. Further analysis of expression mechanism of these
genomes and the genes encoding more protective antigens would
assist in designing an effective vaccine candidate against the
human ehrlichiosis.

REFERENCES
1. Bakken, J. S., Krueh, J., Wilson-Nordskog, C., Tilden, R. L., Asanovich, K.,
and Dumper, J. S. (1996) J. Am. Med. Assoc. 275, 199–205
2. Chen, S.-M., Dumper, J. S., Bakken, J. S., and Walker, D. H. (1994) J. Clin.
Microbiol. 32, 589–595
3. Goodman, J. L., C. Nelson, C., Vitale, B., Madigan, J. E., Dumler, J. S., Kurtti,
T. J., and Munderlich, U. G. (1996) N. Engl. J. Med. 334, 209–215
4. Brouqui, P., Dumler, J. S., Lienhard, R., Brossard, M., and Raoult, D. (1995)
Antimicrob. Agents Chemother. 39, 783–789
5. Pusterla, N., Huder, J. B., Woltersberger, C., Litschi, B., Parris, A., and Lutz,
H. (1997) J. Clin. Microbiol. 35, 2307–2309
6. Pusterla, N., Huder, J., Feige, K., and Lutz, H. (1998) J. Clin. Microbiol. 36,
2035–2037
7. Hardalo, C. J., Quagliarello, V., and Dumper, J. S. (1995) Clin. Infect. Dis. 21,
910–914
8. Rikihisa, Y., Zhi, N., Wormser, G. P., Wen, B. H., Horowitz, H. W., and
Hechmen, K. E. (1997) J. Infect. Dis. 175, 210–213
9. Telford, S. R., III, Dawson, J. E., Katavolos, P., Warner, C. K., Kelbert, C. P.,
and Perang, D. H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 6209–6214
10. Asanovitch, K. M., Bakken, J. S., Madigan, J. E., Auern-Rosenfeld, M., Worm-
ser, G. P., and Dumper, J. S. (1997) J. Infect. Dis. 176, 1029–1034
11. Ido, J. W., Zhang, Y., Hodzic, E., Magallarelli, L. A., Wilson, M. L., Telford,
S. R., III, Barthold, S. W., and Fikrig, E. (1997) J. Infect. Dis. 176, 687–692
12. Zhi, N., Rikihisa, Y., Kim, H. M., Wormser, G. P., and Horowitz, H. W. (1997)
J. Clin. Microbiol. 35, 2606–2611
13. Zhi, N., Ohashi, N., Rikihisa, Y., Wormser, G. P., Horowitz, H. W., and
Hechmen, K. E. (1998) J. Clin. Microbiol. 36, 1666–1673
14. Kim, H. Y., and Rikihisa, Y. (1998) J. Clin. Microbiol. 36, 3278–3284
15. Murphy, C. I., Storey, J. R., Brecchia, J., Doros-Bichert, L. A., Gingrich-Baker,
C., Munroe, K., Bakken, J. S., Coughlin, R. T., and Beltz, G. A. (1998) Infect.
Immun. 66, 3711–3718
16. Ido, J. W., Sun, W., Zhang, Y., Magallarelli, L. A., and Fikrig, E. (1998) Infect.
Immun. 66, 3264–3269
17. Rikihisa, Y., Ewing, S. A., Fox, J. C., Siregar, A. G., Yasari, F. H., and
Malone, M. B. (1999) J. Clin. Microbiol. 37, 143–148
18. Samborok, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A
Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring
Harbor, NY
19. Ohashi, N., Uver, A., Zhi, N., and Rikihisa, Y. (1998) J. Clin. Microbiol. 36,
2071–2080
20. French, D. M., McElwin, T. F., Mcguire, T. C., and Palmer, G. H. (1998)
Infect. Immun. 66, 1200–1207
21. Palmer, G. H., Eid, G., Barbet, A. F., McGuire, T. C., and McElwin, T. F.
(1994) Infect. Immun. 62, 3808–3816
22. Oberle, S. M., and Barbet, A. F. (1993) Gene 136, 291–294

---

2 N. Zhi, N. Ohashi, and Y. Rikihisa, unpublished data.
23. Rikihisa, Y. (1991) Clin. Microbiol. Rev. 4, 286–308
24. Dumler, J. S., and Bakken, J. S. (1996) J. Infect. Dis. 173, 1027–1030
25. Haas, R., and Meyer, T. F. (1986) Cell 44, 107–115
26. Stern, A., Brown, M., Nickel, P., and Meyer, T. F. (1986) Cell 47, 61–71
27. Barbour, A. G. (1990) Annu. Rev. Microbiol. 44, 155–171
28. Zhang, J. R., Hardham, J. M., Barbour, A. G., and Norris, S. J. (1997) Cell 89, 275–285
29. Vanhamme, L., and Pays, E. (1995) Microbiol. Rev. 59, 223–240
30. Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998) EMBO J. 17, 5418–5426
31. Ohashi, N., Zhi, N., Zhang, Y., and Rikihisa, Y. (1998) Infect. Immun. 66, 132–139
32. Palmer, G. H., Oberle, S. M., Barbet, A. F., Goff, W. L., Davis, W. C., and McGuire, T. C. (1988) Infect. Immun. 56, 1526–1531
33. St. Geme, J. W., III, Cutter, D., and Barenkamp, S. J. (1996) J. Bacteriol. 178, 6281–6287

Expressed p44 Genes in the HGE Agent