

Diversity of *Anaplasma phagocytophilum* Strains, USA

Eric Morissette, Robert F. Massung,
Janet E. Foley, A. Rick Alleman, Patrick Foley,
and Anthony F. Barbet

We analyzed the structure of the expression site encoding the immunoprotective protein MSP2/P44 from multiple *Anaplasma phagocytophilum* strains in the United States. The sequence of *p44ESup1* had diverged in *Ap-variant 1* strains infecting ruminants. In contrast, no differences were detected between *A. phagocytophilum* strains infecting humans and domestic dogs.

Anaplasma phagocytophilum (order Rickettsiales) has a broad host range and infects humans as well as numerous other animal species (1). It has been known as a ruminant pathogen in Europe since at least 1932 but in recent years has emerged as a cause of disease in humans in the United States and Europe. The number of cases reported to the Centers for Disease Control and Prevention has increased from 537 in 2004 to 834 in 2007. Similarly, the number of dogs with clinical anaplasmosis has apparently increased. Strains clearly differ; not all appear to be capable of infecting humans or mice (2) or to cause persistent infections. These differences have stimulated the search for molecular markers of strain phenotypes and host tropisms. Although much strain variation has been identified, these markers have not been clearly linked to host tropisms except for 16S rRNA and the US *Ap-variant 1* (*Ap-V1*). *Ap-V1* differs from a human strain (*Ap-ha*) by a 2-bp substitution in the 16S rRNA sequence (3) and appears to be restricted to ruminant species (2,4). Because a 2-bp difference in 16S rRNA is minimally informative, we selected the *msp2/p44* expression site to investigate other potential differences between *Ap-V1* and other strains. In the related organism, *A. marginale*, the genomic repertoire of *msp2* pseudogenes has been associated with the ability, or lack of ability, of strains to superinfect and cause epidemic spread of the organism (5).

Author affiliations: University of Florida, Gainesville, Florida, USA (E. Morissette, A.R. Alleman, A.F. Barbet); Centers for Disease Control and Prevention, Atlanta, Georgia, USA (R.F. Massung); University of California, Davis, California, USA (J.E. Foley); and California State University, Sacramento, California, USA (P. Foley)

DOI: 10.3201/eid1506.081610

The Study

Different isolates of the *Ap-V1* strain were obtained from Rhode Island and Minnesota. Genomic DNA was extracted from infected goat blood, infected *Ixodes scapularis* ticks, and cell culture-grown isolates (ISE6) as described by Massung et al. (6). Other genomic DNA was isolated as described previously (7) from whole infected blood or HL-60 cultures. Dog blood was obtained from naturally infected animals identified by private practitioners in New York and Minnesota. Genomic DNA from the horse MRK strain of *A. phagocytophilum* (8) was isolated from infected equine neutrophils.

PCR amplification, sequencing, and analysis of the *msp2/p44* expression site were performed by methods described previously (7). We used oligonucleotide primers AB1207 (5'-GGGAGTGCTCTGGTTAGATTTAGG-3') and AB1221 (5'-ATAGAACAAGAGCAGGGAGAAGAC-3') or AB1207 and AB1058 (5'-GAACCATCCCCTTA AAATACTTTC-3') to amplify the *msp2/p44* gene, the upstream gene *p44ESup1*, and the intergenic region between them. To amplify just the *msp2/p44* gene in the expression site and to determine hypervariable region sequences, we used AB1221 and AB1266 (5'-GAAGAAGAGATTGGACTTTTGATCTGTC-3') or AB1221 and AB1267 (5'-GAGGAAGAGATTGGA CTTTTGAGCTGTC-3'). The sequences determined here have been assigned GenBank accession nos. FJ467331–FJ467340.

The expression site encoding MSP2/P44 is composed of the MSP2/P44 coding region itself, including a central hypervariable region, an intergenic region containing binding sites for a regulatory transcription factor ApxR, and an upstream gene known either as *p44ESup1* or *omp-1n* (9,10). In our study, the *p44ESup1* gene appeared to be experiencing purifying or stabilizing selection. Evolutionary analysis using MEGA4 (11; Nei-Gojobori method with the Jukes-Cantor correction) showed a low ratio of nonsynonymous-to-synonymous substitutions (mean dN 0.053, dS 0.296; dN/dS ratio 0.179). The *Ap-V1* strains from Rhode Island and Minnesota shared many substitutions with a sheep strain from Norway that were not present in the other *A. phagocytophilum* strains (Figure 1). The *p44ESup1* gene in strains isolated from 5 persons from Wisconsin, Minnesota, and New York most closely resembled *p44ESup1* in strains from clinical infections identified in 3 dogs from New York and Minnesota.

When we performed a concatenated analysis of the *p44ESup1* and intergenic region *p44ESup1* to *msp2(p44)* using maximum-likelihood methods, we found strong support for 3 clades: a clade of strains from eastern North America; a clade of strains from western North America; and a clade comprising a sheep strain from Norway, a dog strain from Sweden, and the 4 *Ap-V1* strains from North America

(Figure 2). Except for *Ap-V1*, the strains from eastern North America appeared to be closely related among themselves;

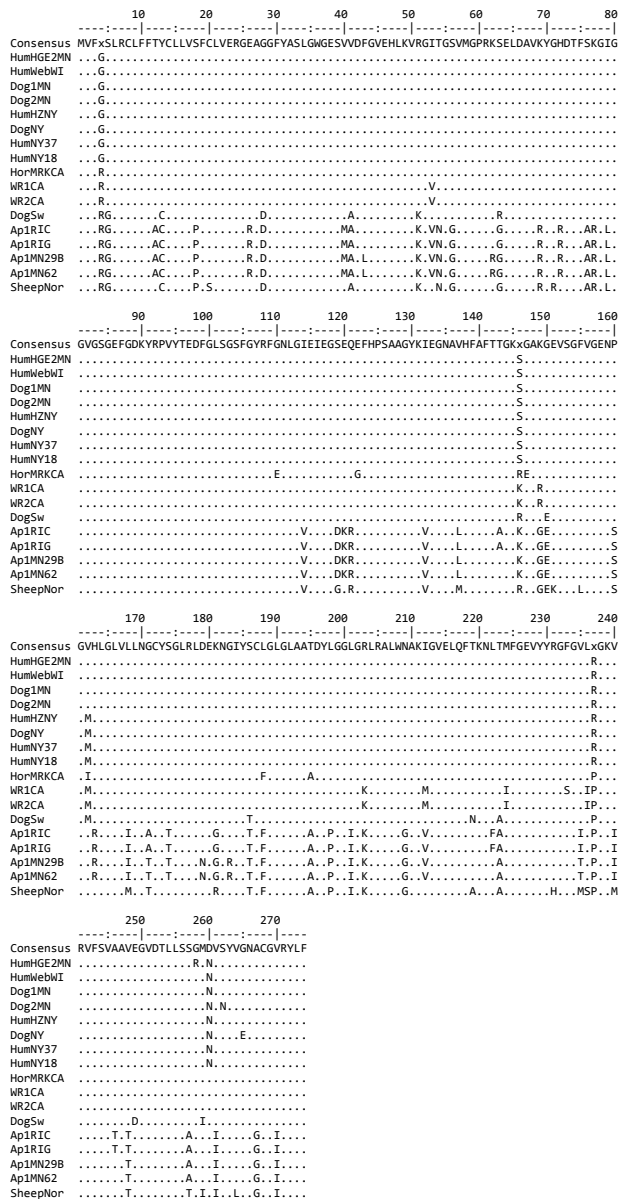


Figure 1. Diversity in the amino acid sequences encoded by *p44ESup1/omp-1n* in US and European strains of *Anaplasma phagocytophilum*. All strains are from the United States (the state is indicated in the strain designation) except for the strain from the sheep from Norway (SheepNor) and the dog from Sweden (DogSw). Human-origin strains are HZNY, NY18, NY37, WebWI, and HGE2MN; dog strains are Dog1MN, Dog2MN, and DogNY; wood rat (*Neotoma fuscipes*) strains are WR1CA and WR2CA; the horse strain is HorMRKCA; *Ap-V1* strains are Ap1RIC (culture derived), Ap1RIG (isolated from an infected goat), Ap1MN29B, and Ap1MN62 (both Ap1MN strains were derived from naturally infected *Ixodes scapularis* ticks). Sequences were from either this study or GenBank: accession nos. DQ519565 (SheepNor), DQ519566 (DogSw), CP000235 (HZ), AY164490 (NY18), AY137510 (NY37), AY164491 (Webster), and AY164492 (HGE2).

the dog and human strains of *A. phagocytophilum* were indistinguishable from each other. Of note, a strain isolated from a dog in Sweden with clinical disease is on a separate branch from all US strains, including those from dogs in the United States.

The central hypervariable regions of *msp2/p44* and the flanking conserved sequences from 34 *Ap-V1* sequences were also aligned. The alignments showed the typical structure, including flanking LAKT residues and conserved framework residues such as C and WP described previously (9,12). Also, multiple hypervariable region variants were identified in each population of *A. phagocytophilum* (organisms characterized at a single time point from a single host). Some of the same variants were identified in different Rhode Island populations. No shared expression site variants were found between the Rhode Island and Minnesota *Ap-V1* strain sequences.

When comparing the *Ap-V1* expression site variants to genomic copies of the sequenced US human HZ strain, we found sequence identities >90% between 20/34 *Ap-V1* variants, including 100% identities of 5/34 *Ap-V1* variants. This level is comparable to that seen in most other US *A. phagocytophilum* strains. When compared with variants (non-HZ) identified directly from human infections, 10/34 *Ap-V1* variants were >90% identical. In contrast, none of the *Ap-V1* variants matched, with at least 70% identity, any previously identified MSP2/P44 expression site variants from strains from sheep in Norway. In general, little similarity was found between the *msp2/p44* hypervariable regions of US and European strain variants.

Conclusions

Despite finding clear differences in the MSP2/P44 hypervariable region repertoire between US and European strains, we did not discover distinct repertoires in any US strains, including in *Ap-V1*. These findings agree with previous data that showed few differences by pulsed-field gel electrophoresis of 7 US strains (13) or by comparative microarray hybridization of 3 US strains (14). Our analysis focused on those hypervariable regions found frequently in the expression site. Because the genome repertoire contains ≈100 functional pseudogenes in each strain, complete genome sequencing may show differences in this repertoire not detected here.

The *p44ESup1* gene, upstream from *msp2/p44* on the same polycistronic mRNA transcript, gave the most phylogenetically useful information. This gene clearly distinguished *Ap-V1* from other US strains. Moreover, the resemblance of the *p44ESup1* gene in *Ap-V1* and in a strain from a sheep in Norway suggests that it may be a marker for a ruminant tropism of *A. phagocytophilum*. Also, phylogenetic trees based on the *p44ESup1* gene grouped *A. phagocytophilum* strains that cause clinical infections in US dogs

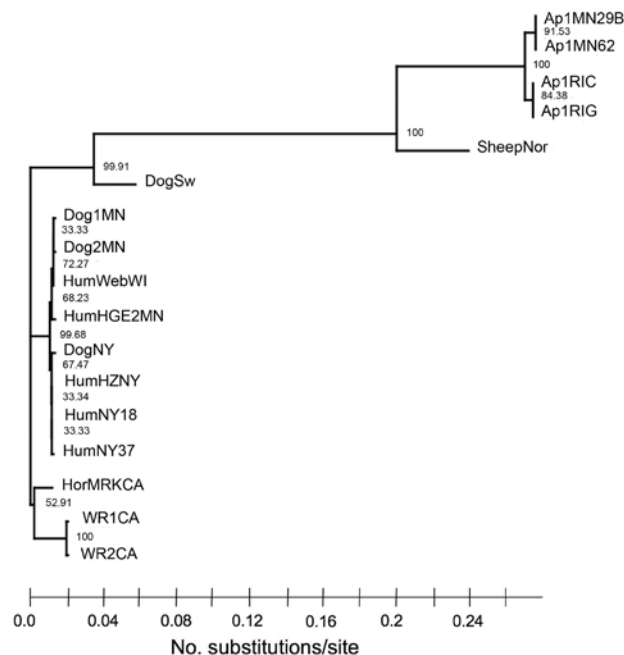


Figure 2. Maximum-likelihood phylogram of different variants of *Anaplasma phagocytophilum* based on the *p44ESup1/omp-1n* and intergenic region gene sequences created by using TREEFINDER (www.treefinder.de) with default values. The number of substitutions per site over 1,092 total sites is shown under the tree, and bootstrap support for each split (percentage of times recovered) is shown next to each branch of the tree.

or humans on the same branch. In fact, the genes from the 2 sources are indistinguishable, which may suggest a recent and common evolutionary origin of the US dog and human strains. Because these US data were obtained from a relatively small sampling of *A. phagocytophilum* infections (although from at least 2 states for the human, dog, and *Ap-V1* strains), these findings should be verified in a larger dataset.

The sequence divergence between strains in *p44E-Sup1* is similar to that in the downstream intergenic region. This intergenic region includes 2 divergent (54% and 58% identity in *Ap-V1*) binding sites for the transcription factor ApxR, which has been postulated to upregulate *msp2/p44* transcription in mammalian cells (15). Either the ApxR transcription factor has low specificity for sequence compared with secondary structure or it does not have the same biological mode of action in *Ap-V1* as in some other strains.

In summary, the *Ap-V1* expression site encoding *msp2/p44* was most similar to a strain from sheep in Norway. Strains causing clinical disease in humans and domestic dogs in the United States were indistinguishable.

Acknowledgments

We gratefully acknowledge the excellent technical assistance of Anna M. Lundgren and the contributions of samples from naturally infected dogs from Matt Eberts and Robin Jaeger.

This project was supported by National Institute of General Medical Sciences grant no. RO1 GM081714.

Dr Morissette is completing a veterinary clinical pathology residency at the University of Florida. His research interests include structure and sequence diversity analysis and evolutionary study of antigenic variability in pathogens of veterinary interest.

References

- Dumler JS, Barbet AF, Bekker CPJ, Dasch GA, Palmer GH, Ray SC, et al. 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*. 2001;51:2145–65.
- Massung RF, Priestley RA, Miller NJ, Mather TN, Levin ML. Inability of a variant strain of *Anaplasma phagocytophilum* to infect mice. *J Infect Dis*. 2003;188:1757–63. DOI: 10.1086/379725
- Massung RF, Mauel MJ, Owens JH, Allan N, Courtney JW, Stafford KC III, et al. Genetic variants of *Ehrlichia phagocytophila*, Rhode Island, and Connecticut. *Emerg Infect Dis*. 2002;8:467–72.
- Massung RF, Mather TN, Levin ML. Reservoir competency of goats for the Ap-variant 1 strain of *Anaplasma phagocytophilum*. *Infect Immun*. 2006;74:1373–5. DOI: 10.1128/IAI.74.2.1373-1375.2006
- Futse JE, Brayton KA, Dark MJ, Knowles DP Jr, Palmer GH. Superinfection as a driver of genomic diversification in antigenically variant pathogens. *Proc Natl Acad Sci U S A*. 2008;105:2123–7. DOI: 10.1073/pnas.0710333105
- Massung RF, Levin ML, Munderloh UG, Silverman DJ, Lynch MJ, Gaywee JK, et al. Isolation and propagation of the Ap-variant 1 strain of *Anaplasma phagocytophilum* in a tick cell line. *J Clin Microbiol*. 2007;45:2138–43. DOI: 10.1128/JCM.00478-07
- Barbet AF, Lundgren AM, Alleman AR, Stuen S, Björnsdóttir A, Brown RN, et al. Structure of the expression site reveals global diversity in MSP2(P44) variants in *Anaplasma phagocytophilum*. *Infect Immun*. 2006;74:6429–37. DOI: 10.1128/IAI.00809-06
- Gribble DH. Equine ehrlichiosis. *J Am Vet Med Assoc*. 1969;155:462–9.
- Barbet AF, Meeus PFM, Bélanger M, Bowie MV, Ji J, Lundgren AM, et al. Expression of multiple outer membrane protein sequence variants from a single genomic locus of *Anaplasma phagocytophilum*. *Infect Immun*. 2003;71:1706–18. DOI: 10.1128/IAI.71.4.1706-1718.2003
- Lin Q, Rikihisa Y, Ohashi N, Zhi N. Mechanisms of variable p44 expression by *Anaplasma phagocytophilum*. *Infect Immun*. 2003;71:5650–61. DOI: 10.1128/IAI.71.10.5650-5661.2003
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007;24:1596–9. DOI: 10.1093/molbev/msm092
- Lin Q, Zhi N, Ohashi N, Horowitz HW, Aguero-Rosenfeld ME, Raf-fali G, et al. Analysis of sequences and loci of *p44* homologs expressed by *Anaplasma phagocytophila* in acutely infected patients. *J Clin Microbiol*. 2002;40:2981–8. DOI: 10.1128/JCM.40.8.2981-2988.2002

13. Dumler JS, Asanovich KM, Bakken JS. Analysis of genetic identity of North American *Anaplasma phagocytophilum* strains by pulsed-field gel electrophoresis. *J Clin Microbiol.* 2003;41:3392–4. DOI: 10.1128/JCM.41.7.3392-3394.2003
14. Hotopp JC, Lin M, Madupu R, Crabtree J, Angiuoli SV, Eisen JA, et al. Comparative genomics of emerging human ehrlichiosis agents. *PLoS Genet.* 2006;2:e21 10.1371/journal.pgen.0020021. DOI: 10.1371/journal.pgen.0020021
15. Wang X, Cheng Z, Zhang C, Kikuchi T, Rikihisa Y. *Anaplasma phagocytophilum p44* mRNA expression is differentially regulated in mammalian and tick host cells: involvement of the DNA binding protein ApxR. *J Bacteriol.* 2007;189:8651–9. DOI: 10.1128/JB.00881-07

Address for correspondence: Anthony F. Barbet, Department of Infectious Diseases and Pathology, College of Veterinary Medicine, Box 110880, University of Florida, Gainesville, FL 32611-0880, USA; email: barbet@ufl.edu

EMERGING INFECTIOUS DISEASES

A Peer-Reviewed Journal Tracking and Analyzing Disease Trends

Vol.6, No.5, Sep–Oct 2000



Search
past issues

EID
online
www.cdc.gov/eid

Genomics

Toxin-Producing *E. coli*

Antigenic Variation