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′-GGGAGTGCCTCTGTTAGATTAG-3′) and AB1221 (5′-ATAGAACAGGGAGGGAAGAC-3′) or AB1207 and AB1058 (5′-GAACCATCCCCATTAAATACCTTC-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′-GAAGAAGAGATTGGAGCTTTTGATCTGTC-3′) or AB1221 and AB1267 (5′-GAAGAAGAGATTGGAGCTTTTGAGC-3′). The sequences determined here have been assigned GenBank accession nos. FJ467331–FJ467340.

The expression site encoding mRNA2/P44 is composed of the mRNA2/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-In (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; 

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 several 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 p44ESup1 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 divergence analysis and evolutionary study of antigenic variability in pathogens of veterinary interest.

**References**


Diversity of Anaplasma phagocytophilum, USA


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