JB Journals.ASM.org

Postgenomic Analyses Reveal Development of Infectious *Anaplasma phagocytophilum* during Transmission from Ticks to Mice

Juliana E. Mastronunzio, Sebastian Kurscheid, and Erol Fikrig

Section of Infectious Diseases, Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, USA

Obligate intracellular bacteria of the *Rickettsiales* order have evolved to colonize both arthropod and mammalian hosts, but few details are known about the bacterial adaptations that occur during transmission from blood-feeding arthropods to mammals. Here we apply proteomics and transcriptome sequencing to *Anaplasma phagocytophilum*, the agent of human granulocytic anaplasmosis, in *Ixodes scapularis* tick salivary glands, to detect proteins or genes expressed by the pathogen during transmission feeding by the tick. We detected expression of 139 genes, representing 11% of the open reading frames (ORFs) in the *A. phagocytophilum* genome. The predominant categories of proteins were ribosomal proteins, cell surface proteins, chaperones, and uncharacterized proteins. There was no evidence of DNA replication enzymes, suggesting that most of the *A. phagocytophilum* cells were no longer dividing. Instead, protein expression reflected conversion to the extracellular, infectious "dense-core" (DC) form. High expression of a DC-specific marker, APH_1235, further suggested this developmental transition in ticks. We showed that blocking APH_1235 with antibodies reduced *A. phagocytophilum* infection levels in mammalian cell culture. This work represents a starting point for clarifying essential proteins expressed by *A. phagocytophilum* during transmission from ticks to mammals and demonstrates that the abundantly expressed, DC-associated APH_1235 protein is important during *in vivo* infection by *A. phagocytophilum*.

Anaplasma phagocytophilum causes human granulocytic anaplasmosis, preferentially infecting neutrophils after transmission by deer tick (*Ixodes scapularis*) nymphs. Anaplasmosis is the second most common tickborne disease in North America, and its reported incidence has doubled over the last decade (6). *A. phagocytophilum* encompasses a heterogeneous group of strains, some of which infect rodents, deer, and horses, but the primary reservoir of human-infective strains is thought to be the white-footed mouse (18). *A. phagocytophilum* belongs to the *Rickettsiales* order, which includes emerging arthropod-transmitted pathogens such as *Ehrlichia* spp. and *Rickettsia* spp. and insect symbionts of the genus *Wolbachia*. These obligate intracellular bacteria have evolved to colonize diverse hosts, from nematodes and arthropods to various mammalian cells. Because many of these bacteria can cause severe disease, it is important to clarify the molecular strategies underlying their ability to thrive in both arthropod and mammalian hosts. In part because genetic approaches have not been optimized in these intracellular pathogens, few details are known about the bacterial adaptations that occur during transmission. To investigate these adaptations, postgenomic approaches such as shotgun proteomics and transcriptome sequencing (RNA-seq) are highly useful. These techniques are sensitive enough to detect bacterial gene expression even within a complex sample derived from infected hosts (4).

To investigate gene expression during transmission of *A. phagocytophilum*, we applied shotgun proteomics (based on peptide separation with liquid chromatography coupled to tandem mass spectrometry, or LC MS/MS) and Illumina transcriptome sequencing (RNA-seq) to infected *Ixodes scapularis* tick salivary glands (SG), collected after 48 h of transmission feeding on mice. We used the genome of *A. phagocytophilum* strain HZ as a reference genome to identify peptides and transcripts from our salivary gland isolate, collected in Connecticut and infectious to both mice and humans (see Materials and Methods). Reference genomes are very useful for unsequenced strains, as previously shown for other host-associated bacteria (26). By direct sampling of tick salivary glands, our analysis complements culture-based studies and identifies additional proteins that may be particularly important for survival during transmission to mammalian hosts. Knowledge of these proteins would also aid in designing vaccines or antibacterial therapies, as they include surface molecules expressed when *A. phagocytophilum* first enters the mammal.

The main published evidence regarding proteins required during the tick stage of the *A. phagocytophilum* life cycle comes from a microarray analysis comparing bacterial gene expression in embryonic tick cells to that in human cells. Forty-one genes, nearly all encoding hypothetical proteins, were upregulated in tick cells (28). While useful as model systems, tick embryonic cell lines likely do not recreate the complex changes occurring during blood feeding in differentiated tick salivary glands, so our report is the first to our knowledge that catalogs genes and proteins expressed by *A. phagocytophilum* in tick salivary glands. Recently, two shotgun proteomics studies reported *A. phagocytophilum* proteins expressed in culture in the human leukemia promyelocytic cell line HL-60. With a high depth of sampling, nearly all predicted open reading frames (ORFs) were detected by Lin and colleagues (23). The analysis performed by Troese et al. detected approximately 24% of the proteome and identified a highly expressed protein, APH_1235, that can serve as a marker specific for the infectious

Received 22 December 2011 **Accepted** 24 February 2012

Published ahead of print 2 March 2012

Address correspondence to Juliana E. Mastronunzio, juliana.mastronunzio@gmail.com.

Supplemental material for this article may be found at http://jb.asm.org/.

doi:10.1128/JB.06791-11

"dense-core" (DC) form of *A. phagocytophilum* (36). Our analysis also pointed to the upregulation of APH_1235 during transmission from ticks, as shown below.

Here we provide a summary of *A. phagocytophilum* genes expressed in the tick salivary gland, discuss the insights provided from this analysis, and investigate an abundantly expressed protein, APH_1235, throughout the pathogen's transmission cycle.

MATERIALS AND METHODS

Tick transmission feeding. *A. phagocytophilum*-infected *Ixodes scapularis* tick nymphs were generated by feeding uninfected larvae on infected C3H/HeJ mice. The mouse stock was originally infected by an *A. phagocytophilum* strain from adult ticks collected from North Branford, CT, and infection was maintained between mice by blood inoculation. Infectivity for humans was confirmed by *in vitro* infection of human neutrophils. After *Ixodes* larvae were fed on infected mice, they were allowed to molt to nymphs, and 10% of nymphs per batch (fed on one mouse) were tested for *A. phagocytophilum* infection by the use of quantitative PCR (qPCR) of the 16S rRNA gene. Between feedings, *Ixodes scapularis* ticks were maintained in 23°C incubators with 85% relative humidity. For transmission feeding, groups of 25 to 30 ticks were placed on mice and allowed to feed. Ticks were removed after 48 h, and salivary glands were dissected as described below and processed for either protein or RNA extraction. All animal protocols used were approved by the Yale University IACUC.

Salivary gland protein extraction and trypsin digestion. *A. phagocytophilum*-infected *Ixodes scapularis* nymphs were placed on mice and allowed to feed for 48 h. Salivary glands were dissected on glass slides in phosphate-buffered saline (PBS) with Complete protease inhibitor cocktail (Roche, NJ). Salivary glands from 200 ticks were dissected and stored in PBS with protease inhibitors at −80°C. Glands were resuspended in cell lysis buffer (Cell Signaling, Danvers, MA) with Complete protease inhibitor cocktail and lysed by three freeze-thaw cycles of 37°C for 5 min and −80°C for 5 min. The protein concentration in the extract was measured with the bicinchoninic acid (BCA) assay (ThermoFisher). The pooled salivary gland sample containing 160 μg of total protein was lyophilized using a SpeedVac concentrator, and the lypophilized pellet was dissolved in 8 M urea–0.4 M NH_4HCO_3 (pH 7.5 to 8.5). Dithiothreitol was added (at 7.5 mM), and samples were incubated for 20 min at 37°C, followed by addition of 15 mM iodoacetamide and 20 min of incubation at room temperature. The solution was diluted in water to 2 M urea, and sequencing-grade trypsin (Promega) was added according to the manufacturer's instructions. The trypsin digest was incubated overnight at 37°C and then frozen at −80°C prior to LC MS/MS analysis.

Strong cation exchange peptide separation and liquid chromatography-tandem mass spectrometry (LC-MS/MS). Proteomic analysis was performed at the Keck Biotechnology and Resource Laboratory at the Yale School of Medicine. Tryptic digests were acidified by adding 2 μl of 1 M phosphoric acid, and cation exchange chromatography was performed on an Applied Biosystems Vision Workstation, using a PolySULFOETHYL A column (PolyLC Inc.) (2.1 by 200 mm) with a linear 118-min gradient. Each cation exchange fraction (10 total) was dried and resuspended in 5 μl of 70% formic acid and diluted with 0.1% trifluoroacetic acid (TFA) prior to loading onto the LTQ Orbitrap system. The LTQ Orbitrap is equipped with a Waters nanoAcquity ultraperformance LC (UPLC) system and uses a Waters Symmetry C18 trap column (180 μm by 20 mm) and a nanoAcquity UPLC column (1.7 μm pore size, 75 μm by 250 mm) (35°C) for peptide separation. Mass spectra were acquired in the Orbitrap using 1 microscan and a maximum inject time of 900 followed by four datum-dependent MS/MS acquisitions in the ion trap.

All MS/MS spectra were searched using the Mascot algorithm for uninterpreted MS/MS spectra after using the Mascot Distiller program to generate Mascot-compatible files (13). Peptide spectra were searched against a database consisting of the protein coding sequences from the *Ixodes scapularis* genome (www.vectorbase.org [accessed March 2011]) and the *Anaplasma phagocytophilum* strain HZ genome (Refseq NC_007797; updated December 2010). The Mascot significance score match is based on a MOWSE score and relies on multiple matches to two or more peptides from the same protein. Typical parameters used for searching are a peptide tolerance of ±20 ppm, MS/MS fragment tolerance of ±0.6 Da, and a peptide charge of +2 or +3. Normal and decoy database searches are run. Generally, proteins must have two or more unique peptides to be considered identified; however, we have also listed *A. phagocytophilum* proteins with one unique peptide match. Protein abundance was estimated with the exponentially modified protein abundance index (emPAI) score (17). The emPAI value and the percent sequence coverage were provided in the Mascot output.

Salivary gland RNA extraction and transcriptome sequencing (RNA-seq). *A. phagocytophilum*-infected *Ixodes scapularis* nymphs were fed on mice for 48 h. Salivary glands from 200 ticks were dissected on glass slides in RNase-free PBS and placed directly into TRIzol reagent (Invitrogen). Samples were either directly processed or stored at −80°C until RNA extraction. The total RNA was extracted following the manufacturer's protocol with minor modifications. Briefly, a 0.2 volume of chloroform was added per volume of TRIzol used for the initial suspension of sample. The tubes were shaken vigorously and incubated for 10 min at room temperature followed by 15 min of centrifugation at 12,000 × *g* at 4°C. The aqueous phase was transferred to fresh tubes, and a 0.5 volume per 1 initial volume of TRIzol and 5 μl of RNase-free glycogen (Ambion/Life Technologies) were added to precipitate RNA out of the solution. After 10 min of incubation at room temperature, the samples were centrifuged at 12,000 × *g* for 10 min at 4°C. The supernatant was removed, and the pellet was washed with 1 volume of 75% ethanol and finally centrifuged for 5 min at 7,500 × *g* at 4°C. After removal of ethanol, the pellet was air dried and the RNA resuspended in RNase-free water. The RNA extracted from the salivary glands was quantified by spectrophotometry and totaled 21 μg. Eukaryotic rRNA and mRNA were depleted from the total RNA by the use of a MicrobeEnrich kit (Ambion/Life Technologies). The remaining RNA was then treated with a MicrobeExpress kit (Ambion) to deplete bacterial rRNA. The resulting RNA was ethanol precipitated, resuspended in 20 μl of RNase-free water, and submitted to the Illumina sequencing facility at the Yale Center for Genome Analysis.

A cDNA library was constructed using an mRNA-seq Sample Prep kit with random hexamer primers (Illumina, San Diego, CA). The sequencing was carried out with an Illumina Genome Analyzer II system using standard protocols, generating 36-bp reads (27). Reads were assembled into contiguous sequences by the use of an ABySS *de novo* assembler, setting the k-mer size to 30 (33). The average coverage of assembled contigs was 175×. The assembled transcript data were then used as a query data set for separate BLASTn searches of *A. phagocytophilum* HZ transcripts (from the Comprehensive Microbial Resource of the J. C. Venter Institute) and *I. scapularis* transcripts (from www.vectorbase.org). To increase the sensitivity of the BLASTn search (Nucleotide-Nucleotide BLAST 2.2.14+), the parameter for word size was reduced to 7, and to limit the output of false-positive alignments, the e-value was set to $1e^{-10}$. The quality of the contig assemblies was confirmed by performing pairwise alignments between the *de novo* contigs and the known *A. phagocytophilum* HZ transcripts, resulting in identities > 99% for the highest-scoring pairs on the nucleotide level. To calculate the percentage of each *A. phagocytophilum* transcript covered, the lengths of all contigs matching one gene were summed and divided by the length of the gene.

qRT-PCR. For qRT-PCR analysis, salivary glands were pooled from 2 to 3 ticks and stored in RLT buffer at −80°C prior to RNA extraction with a Qiagen RNEasy kit (Qiagen, CA). Cultured HL-60 cells (sampled on days 1 to 4 postinfection) were also processed with this kit, while RNA from infected mouse blood (day 6 postinfection) was extracted using a mouse RiboPure blood RNA isolation kit (Ambion/Life Technologies). Reverse transcription was performed with an iScript cDNA synthesis kit (Bio-Rad) using 10 μl of RNA per reaction. Reactions were run on a Bio-Rad iCycler using typical SYBR green reaction conditions with 50 cycles of amplification and an annealing temperature of 58°C. A standard curve of dilutions containing 10^{-2} to 10^{-7} ng of the 16S rRNA amplicon

TABLE 1 Identification of *A. phagocytophilum* proteins and transcripts in tick salivary glands

	Proteome			RNA-seq		
Genome	No. of peptides	No. of proteins	No. of proteins with >1 peptide	No. of reads	No. of transcripts	No. of genes
A. phagocytophilum	147	71	23	187,284	255	82
I. scapularis	1,881	639	266	39,492,919	3,645	777

was used to calculate starting quantities and compare expression levels. The following primer sequences were used: for APH_1235, TCTTCTTC AGACGATAGCAGGA (forward) and GTTGTCGCTAGGTGCTTGAG (reverse); for 16S rRNA, GGTGAGTAATGCATAGGAATC (forward) and GCTCATCTAATAGCGATAAATC (reverse); for tick actin, GGCGA CGTAGCAG (forward) and GGTATCGTGCTCGACTC (reverse); and for human/mouse actin, AGCGGGAAATCGTGCGTG (forward) and CA GGGTACATGGTGGTGCC (reverse).

***A. phagocytophilum* culture.** *A. phagocytophilum* was initially isolated both from infected mouse blood and from infected tick salivary glands. This isolate originated from ticks collected in North Branford, CT, and has been subjected to repeated passage between ticks and C3H/HeJ mice at Yale University, New Haven, CT. *A. phagocytophilum* was propagated in the human promyelocytic cell line HL-60 (ATCC CCL-240) in Iscove's modified Dulbecco's medium with 10% fetal bovine serum at 37°C with 5% CO_2. For routine culture, highly infected (>80%) cultures were diluted between 1:10 and 1:20 with uninfected HL-60 cells every 3 to 4 days.

For adherence assays performed in the presence of anti-APH_1235 antibody, host cell-free *A. phagocytophilum* cells were isolated by lysing host cells with five passages through a 27-gauge syringe, followed by differential centrifugation to separate bacteria from host cell debris (3). IgG was purified from rabbit antiserum by the use of a protein A/G Spin kit (ThermoScientific, IL). Host-cell-free bacteria were incubated for 30 min at 37°C in medium containing anti-APH_1235 IgG or preimmune IgG control (50 μg/ml). HL-60 cells (equivalent to the original number of infected cells from which *A. phagocytophilum* bacteria were isolated) were added, and the mixture was incubated for 40 additional min at 37°C to allow binding of bacteria to host cells. After two washes in medium, HL-60 cells were subjected to a cytospin procedure on glass slides and stained with antibodies against *A. phagocytophilum* (as described below), and the numbers of bound *A. phagocytophilum* bacteria per host cell were counted.

For the propagation experiment in the presence of anti-APH_1235 IgG, highly (>80%) infected cells were diluted to 3% of the concentration of uninfected cells and cultured in medium containing rabbit anti-APH_1235 IgG or preimmune IgG (20 μg/ml), based on methods described previously (29). Fresh medium containing IgG was added after 48 h. Cells were collected after 4 days of propagation. An aliquot of cells was subjected to a cytospin procedure on slides, and the cells were subjected to Giemsa staining and counted to determine the percentage of HL-60 cells containing *A. phagocytophilum* morulae. The remaining culture was centrifuged to collect cells for DNA extraction using a Qiagen DNEasy kit (Qiagen, CA). *A. phagocytophilum* levels were quantified by qPCR using the 16S rRNA gene normalized to host actin.

Surface trypsin digestion and Western blotting. For surface trypsinolysis of *A. phagocytophilum* cells, host cell-free bacteria were isolated by lysing highly infected host cells (approximately 1×10^7) with five passages through a 27-gauge syringe, followed by differential centrifugation to separate bacteria from host cell debris (3). The pellet of purified *A. phagocytophilum* was resuspended in SPG buffer (0.25 mM sucrose, 10 mM sodium phosphate, 5 mM L-glutamic acid, pH 7.2), and 5 μl of sequencing-grade trypsin (Promega, WI) was added to half of the cell reaction mixture. Bacteria were incubated at 37°C for 30 min and then pelleted at 10,000 $\times$ *g* and resuspended in Laemmli protein loading buffer, boiled for 5 to 10 min, and processed on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels. After protein transfer to a nitrocellulose membrane, blots were probed with either anti-APH_1235 antiserum (1:1,000) or anti-p44 antiserum (1:5,000) followed by anti-rabbit horseradish peroxidase (HRP).

Antibody generation and immunofluorescent staining. The *APH_1235* gene locus from *A. phagocytophilum* strain HZ was amplified with BamHI and NotI linkers and cloned into pGEX-6P-2 (GE Healthcare). The construct was cloned into *Escherichia coli* BL21, and expression was induced with IPTG (isopropyl-β-D-thiogalactopyranoside). N-terminally glutathione transferase (GST)-tagged APH_1235 was purified using Sepharose beads, and the GST tag was cleaved with PreScission protease. Purified APH_1235 was injected with Freund's complete adjuvant into rabbits and mice, with two additional boosts, to generate anti-APH_1235 antiserum.

For immunostaining of infected HL-60 cells, cells were subjected to a cytospin procedure on slides and fixed with Hema 3 fixative (Fisher Scientific), which also permeabilizes cells. After blocking for 1 h in PBS–5% bovine serum albumin (BSA), cells were stained in PBS–5% BSA with rabbit antiserum against whole *A. phagocytophilum* cells (1:500) and mouse antiserum against APH_1235 (1:100), washed, and then stained in PBS–1% BSA with secondary antibodies (1:1,000) conjugated to (tetramethyl rhodamine isocyanate (TRITC) or Alexa 488. Host cell nuclei were stained with DAPI (4′,6′diamidino-2-phenylindole) (1:2,000). Images were viewed with a Zeiss Axiovision fluorescence microscope.

RESULTS

***A. phagocytophilum* gene and protein expression in tick salivary glands.** We present here the first profile of proteins expressed by *A. phagocytophilum* in infected tick salivary glands, sampled *in vivo* during transmission feeding on mice. Tick salivary glands were dissected after 48 h of feeding, which is roughly 12 to 18 h after the onset of transmission but well before ticks are fed to repletion and fall off the host (at 72 to 96 h) (19). By combining the results from the proteome and transcriptome analyses, 139 genes were identified, representing 11% (139/1,264) of annotated ORFs for this organism. Table S1 in the supplemental material shows a combined list of genes identified by matching peptides or transcripts to the *A. phagocytophilum* strain HZ genome and includes the percentage of the transcript or protein detected by each method. Table 1 shows the total number of proteins and transcripts identified in the *A. phagocytophilum* genome compared to the tick host (*I. scapularis*) genome. Each method yielded approximately 10% of identifications matching *A. phagocytophilum* (Table 1, columns 2 and 6). Expression of 14 *A. phagocytophilum* genes was detected by both approaches. When classified into functional categories (Clusters of Orthologous Groups [COGs]), the results of the breakdown of protein functions identified by the two approaches were quite similar, with the main differences being a higher proportion of chaperones and stress proteins detected in the proteome and more hypothetical protein expression detected in the transcriptome (Fig. 1; see also Table S1 in the supplemental material). The most abundant groups of proteins were those in COG J (translation and ribosomal proteins), COG M (cell envelope), COG O (protein turnover and chaperones), and proteins not found in COGs.

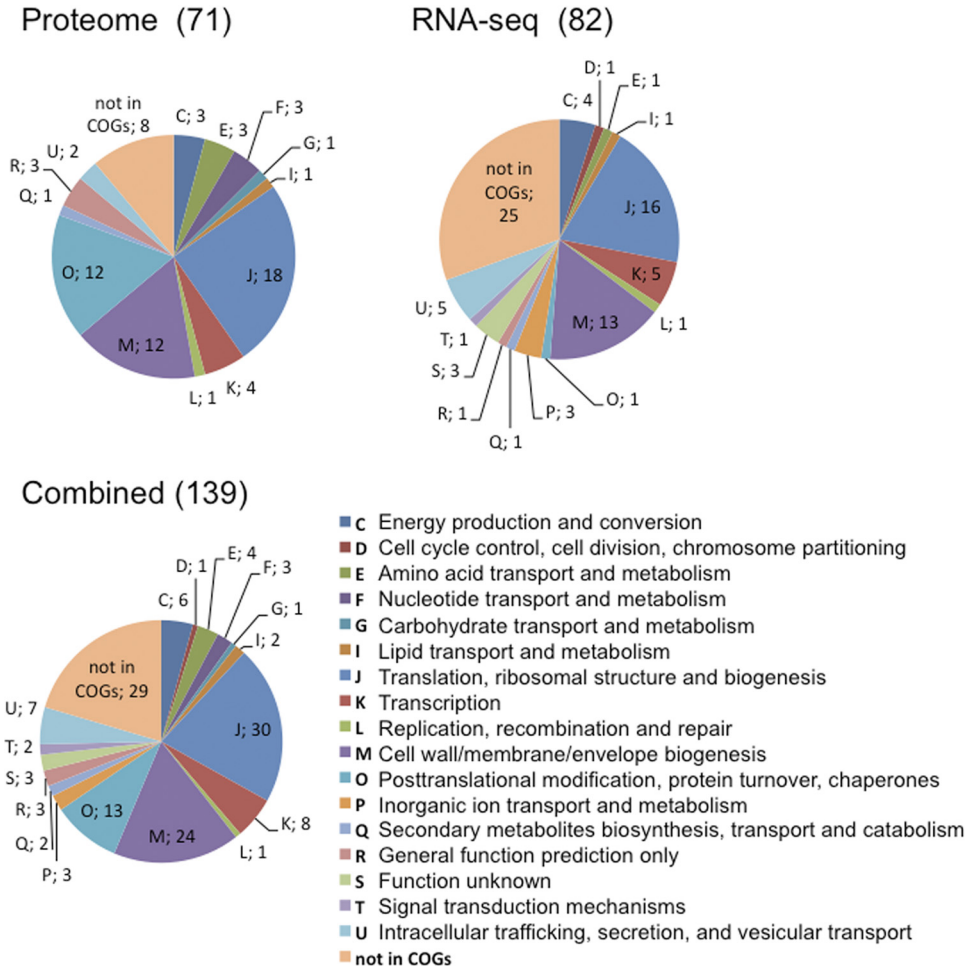

FIG 1 *A. phagocytophilum* proteins and genes detected in tick salivary glands, separated by identification method, and categorized into functional categories of Clusters of Ortholog Groups (COGs). COG classifications from the NCBI genome website for *A. phagocytophilum* strain HZ are shown, with the exception of COG M, which was assigned by us to include known surface proteins.

Replication, transcription, and translation. *A. phagocytophilum* is known to require at least 30 to 36 h of tick feeding for transmission to mice (19). At 48 h, when our samples were collected, there were no DNA replication enzymes detected by proteomics or RNA-seq. Although the cell division protein FtsZ was identified, it is likely that the major phase of pathogen chromosome replication is less active by this time point, based on the lack of detection of DNA replication machinery. Compared to the proteome observed in the HL-60 culture, in which DNA replication machinery proteins were abundant (23, 36), the tick-expressed proteome instead revealed only RNA polymerase subunits, a DNA-binding HU protein, and numerous ribosomal proteins, indicating that *A. phagocytophilum* transcription and translation are more active than replication at 48 h of tick transmission feeding. This correlates well with the developmental cycle of *A. phagocytophilum*, in which cells convert from an intracellular, replicative form (termed "RC" for a reticulate appearance with electron microscopy) to the nondividing, infectious dense-core (DC) form, akin to the biphasic cycles described in other intracellular bacteria such as *Ehrlichia* spp. and *Chlamydia* spp. (35). This developmental process has not been explicitly demonstrated for *A. phagocytophilum* in *Ixodes scapularis* ticks, but it is well established in culture, where DC forms mature between 28 to 36 h, which is consistent with the duration of tick feeding required to transmit the pathogen (36).

Carbon and energy metabolism. Energy metabolism was represented in our analyses by four enzymes of the tricarboxylic acid (TCA) cycle, along with a proline dehydrogenase and one subunit of ATP synthase. One enzyme of glycolysis was detected, but the *A. phagocytophilum* genome, like those of other *Rickettsiales* bacteria, does not encode enzymes of a complete glycolytic pathway, so *A. phagocytophilum* does not metabolize sugars in the tick host (8). The carbon sources for *A. phagocytophilum* growth have not been confirmed experimentally but are most likely host-derived TCA cycle intermediates and amino acids (8). Such compounds could potentially be imported via porin channels formed by the p44 family of surface proteins (16). Although one ATP synthase subunit was identified, no electron transport chain dehydrogenases were detected, so it is possible that the electron transport chain is not active at this stage. A proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehydrogenase, involved in metabolism of arginine, proline, and glutamate, was detected by both proteomics and RNA-seq, which further supports the picture of *A. phagocytophilum* metabolizing amino acids in the tick. Four proteins or

TABLE 2 Top 20 most abundant *A. phagocytophilum* proteins detected[a]

Locus tag	Annotation	% coverage	emPAI	COG	Category
APH_0240	Chaperonin GroEL	63.6	2.54	O	Posttranslational modification, protein turnover, chaperones
APH_0174	P44-55 outer membrane protein, silent	44.8	1.9	(M)	Cell wall/membrane/envelope biogenesis
APH_1215	P44-14 outer membrane protein	39.7	1.7	(M)	Cell wall/membrane/envelope biogenesis
APH_1168	P44-36 outer membrane protein	43.4	1.59	(M)	Cell wall/membrane/envelope biogenesis
APH_1229	P44-2b outer membrane protein	39.8	1.51	(M)	Cell wall/membrane/envelope biogenesis
APH_1235	Hypothetical protein APH_1235	56	1.48		Not in COGs
APH_1221	P44-18ES, P44 outer membrane protein expression locus with P44-18	39.4	1.46	(M)	Cell wall/membrane/envelope biogenesis
APH_1311	P44-24 outer membrane protein	23.4	1.4	(M)	Cell wall/membrane/envelope biogenesis
APH_0241	Cochaperonin GroES	38.3	1.36	O	Posttranslational modification, protein turnover, chaperones
APH_0784	DNA-binding protein HU	30.8	1.36	L	Replication, recombination, and repair
APH_1275	P44-16b outer membrane protein	30.4	1.25	(M)	Cell wall/membrane/envelope biogenesis
APH_0278	Translation elongation factor Tu	53.4	0.9	J	Translation, ribosomal structure, and biogenesis
APH_1025	50S ribosomal protein L7/L12	75.4	0.88	J	Translation, ribosomal structure, and biogenesis
APH_1266	50S ribosomal protein L19	58.6	0.83	J	Translation, ribosomal structure, and biogenesis
APH_1169	P44-19 outer membrane protein, silent	46.8	0.74	(M)	Cell wall/membrane/envelope biogenesis
APH_0292	50S ribosomal protein L5	34.5	0.59	J	Translation, ribosomal structure, and biogenesis
APH_0135	CSD family cold shock protein	14.6	0.56	K	Transcription
APH_1359	Major outer membrane protein OMP-1A	35.2	0.54	(M)	Cell wall/membrane/envelope biogenesis
APH_1051	30S ribosomal protein S21	25.3	0.52	J	Translation, ribosomal structure, and biogenesis

[a] Protein abundance data are based on the Exponentially Modified Protein Abundance Index (emPAI) value, provided in the Mascot output. Parentheses indicate COG categories that were not assigned on the NCBI genome website but are appropriate categories based on previously published work.

transcripts from amino acid metabolic pathways, three from nucleotide metabolism, and two from lipid metabolism were identified, but there was no evidence of vitamin or cofactor biosynthesis (see Table S1 in the supplemental material).

Surface proteins and protein secretion. Cell surface proteins formed the second most frequently detected category of proteins, after the ribosomal proteins. We assigned these confirmed surface-expressed proteins, including the p44 family proteins Asp55 and Asp62, and outer membrane proteins (OMPs) to COG category M (cell envelope) for this analysis, although these genes were not assigned to COGs by NCBI (Fig. 1; see also Table S1 in the supplemental material) (11). The p44 (msp2 homolog) family proteins comprise the majority of the abundantly detected proteins (Table 2). This multigene family, with over 100 paralogs in the *A. phagocytophilum* genome, has undergone dramatic lineage-specific expansion in this species. "Silent" genetic loci recombine at the main p44 expression locus, p44-18-ES, allowing expression of distinct forms of p44 in populations of *A. phagocytophilum* in ticks and mammalian hosts (38).

Many of the surface-expressed proteins, including Omp-1A and four p44 family proteins, contain predicted signal peptides for secretion by the general secretory (Sec) pathway (see asterisks in Table S1 in the supplemental material). Along with the Sec pathway, the *A. phagocytophilum* genome encodes the Twin Arginine Translocation (Tat) pathway and type IV secretion (T4S) system. We found evidence of each of these protein secretion pathways in tick salivary glands: SecG, YajC, and signal recognition particle (SRP) of the Sec secretion system, TatA from the Tat system, and VirB6, VirB9, and VirB2 homologs of the T4S system. Protein substrates likely exported by these systems were also detected. Thirteen proteins with predicted Sec signal peptides were observed, though no Tat substrates are known in *A. phagocytophilum*. Of Ats-1 and AnkA, the two putative effectors secreted by the T4S system in *A. phagocytophilum*, only AnkA was identified in ticks by RNA-seq (31).

Hypothetical protein APH_1235 is upregulated during transmission from ticks and exit from mammalian cells. We detected expression of 27 hypothetical proteins in tick salivary glands by proteomics or RNA-seq (Table 3 and Table 4). Some of these proteins have previously been characterized in HL-60 cultures for their roles in *A. phagocytophilum* infection (14, 15, 34) or identified as immunoreactive proteins (24). Two hypothetical proteins, APH_0220 and APH_0792, were classified as among the most abundant in the HL-60 culture proteome (23). The majority of the *A. phagocytophilum* hypothetical proteins detected in ticks lack any conserved protein domains (Table 4). Most have orthologs in *Anaplasma marginale*, though several proteins do not show amino acid similarity to any sequences in the NCBI nr database (Table 4).

The most abundant hypothetical protein identified by proteomics was APH_1235 (Table 2). This was also the only hypothetical protein detected by both proteomics and RNA-seq. Transcriptome sequencing covered the entire 405-bp transcript of APH_1235, making the gene encoding APH_1235 one of only two

TABLE 3 Characterized "hypothetical" *A. phagocytophilum* proteins detected in tick salivary glands

Locus tag	Protein name	Characterized function	Reference
APH_0032		*A. phagocytophilum* vacuolar membrane	14
APH_0233	AptA	Toxin; activates Erk1/2 MAP kinase	34
APH_1387	HGE-2	*A. phagocytophilum* vacuolar membrane	15
APH_0387	HGE-14	Immunoreactive protein	24
APH_0452	HGE-14	Immunoreactive protein	24
APH_1235		Highly expressed by dense core (DC) cells	36

TABLE 4 Uncharacterized "hypothetical" *A. phagocytophilum* proteins detected in tick salivary glands

Locus tag	Conserved domain(s)	Top-scoring BLAST hit	E value
APH_0005	No	AM1285 (*Anaplasma marginale* strain St. Maries)	3.00E-78
APH_0062	DUF1013	AmarPR_00231 (*Anaplasma marginale* strain Puerto Rico)	9.00E-119
APH_0117	No	ACIS_01138 (*Anaplasma centrale* strain Israel)	3.00E-17
APH_0220	COG 5388; PAS_5	AM967 (*Anaplasma marginale* strain St. Maries)	2.00E-84
APH_0248	No	AM936 (*Anaplasma marginale* strain St. Maries)	2.00E-34
APH_0328	No	ACIS_00480 (*Anaplasma centrale* strain Israel)	1.00E-75
APH_0713	No	ACIS_00663 (*Anaplasma centrale* strain Israel)	6.00E-33
APH_0723	No	Rvan_2191 (*Rhodomicrobium vannielii* ATCC 17100)	0.033
APH_0785	No	None	n.a.[a]
APH_0792	No	ACIS_00868 (*Anaplasma centrale* strain Israel)	3.00E-12
APH_0819	No	Vacuolar protein sorting 35 (*Loa loa*)	0.76
APH_0843	No	APH_0838 (*Anaplasma phagocytophilum* HZ)	8.00E-30
APH_0915	No	None	n.a.
APH_0922	No	APH_0919 (*Anaplasma phagocytophilum* HZ)	0
APH_1069	No	AM216 (*Anaplasma marginale* strain St. Maries)	3.00E-28
APH_1142	No	Serine phosphatase (*Desulfohalobium retbaense* DSM 5692)	1.2
APH_1237	No	None	n.a.
APH_1280	Thioredoxin-like	ACIS_00004 (*Anaplasma centrale* strain Israel)	5.00E-83
APH_1378	No	AM1228 (*Anaplasma marginale* strain St. Maries)	6.00E-103
APH_1394	Ttg2D	AmarM_05210 (*Anaplasma marginale* strain Mississippi)	4.00E-51
APH_1412	No	PEBP family protein (*Thiomonas intermedia* K12)	0.13

[a] n.a., not available.

genes with 100% of the transcript observed (see Table S1 in the supplemental material). APH_1235 is fairly conserved among other members of the *Anaplasmataceae*, exhibiting 49% amino acid sequence identity to the ortholog in *A. marginale*. A PSI-BLAST search using the APH_1235 amino acid sequence against genomes of the *Chlamydiales* revealed low similarity to *ompH* of several *Chlamydiae* species. In addition to its high expression, we chose to investigate this protein further because it contains a high proportion (22%) of serines, which is characteristic of some bacterial adhesins (39), and because its gene is located near the main expression locus (*p44-18-ES*) of the p44 major surface proteins, the most prevalent surface antigens of *A. phagocytophilum*.

To investigate APH_1235 expression throughout the transmission cycle, we used qRT-PCR to measure mRNA levels in salivary glands at different stages of transmission feeding, as well as levels in infected mice and in cultured human and tick cells. After normalization to 16S rRNA expression, APH_1235 levels were 10- to 20-fold higher in tick salivary glands compared to mammalian cells, and expression in salivary glands increased drastically from 24 to 36 h of feeding (Fig. 2). APH_1235 expression was not detected during acquisition of *A. phagocytophilum* by uninfected tick nymphs fed on infected mice (not shown).

Levels of *A. phagocytophilum* expression of APH_1235 also differed during the infection cycle in mammalian cells. During routine culture in human promyelocytic HL-60 cells, *A. phagocytophilum* bacteria form intracellular clusters, called morulae, within a pathogen-containing vacuole. The dense-core (DC) forms of *A. phagocytophilum* bind and infect new cells and then switch to intracellular, replicating cells (RC form) and later differentiate back to DCs after 28 to 36 h (36). Using immunofluorescence, we observed an increase in APH_1235 production as *A. phagocytophilum* reached high numbers before exiting host cells. Figure 3 (top panels) shows increased APH_1235 fluorescence (green) expressed by *A. phagocytophilum* (red) in larger, mature morulae, as cells are transitioning to the DC form, and in extracellular bacteria after host cell lysis. The bottom panels depict parallel Giemsa-stained HL-60 cultures at equivalent stages of infection. Early in infection, at 16 h (Fig. 3B and G), the replicating RC cells within smaller morulae do not express APH_1235. A recent independent study by Troese and colleagues has verified that APH_1235 is expressed mainly by the DC form of *A. phagocytophilum* in HL-60 cell cultures, appearing during the transition from RC to DC form and in abundance as cells are released to initiate a second round of infection (36).

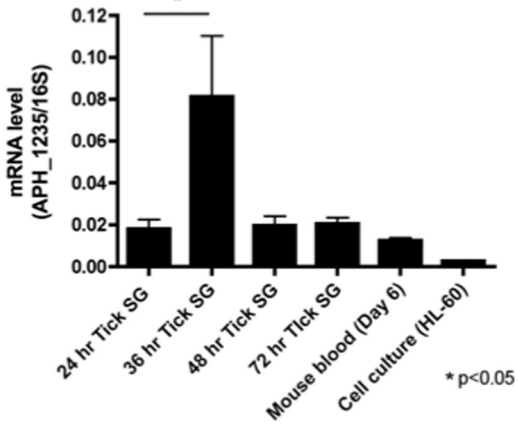

FIG 2 APH_1235 expression is induced during transmission feeding. Expression of APH_1235 in different hosts was measured by qRT-PCR and normalized to 16S rRNA levels. Tick salivary gland (SG) samples were collected at 24, 36, 48, and 72 h of feeding. Blood from tick-infected C3H/HeJ mice was collected on day 6 after removal of ticks. Cultured human promyelocytic cells (HL-60) were sampled at 24-h intervals after infection, from 24 to 96 h. Data for each host environment were averaged from the results obtained with 6 to 16 RNA samples, and a Student's *t* test was used to assess the statistical significance of differences between groups.

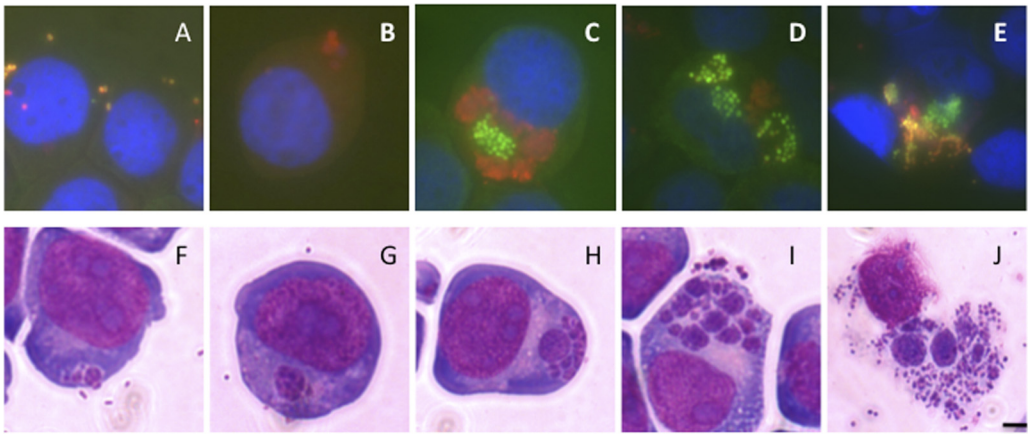

FIG 3 Production of APH_1235 (green) increases as *A. phagocytophilum* bacteria transition to the dense-core (DC) phase and exit host cells. Infected HL-60 cells were collected during the course of *A. phagocytophilum* infection, and representative images are shown. APH_1235 levels increase as morulae mature and DCs develop between 28 to 36 h (C and D), and APH_1235 is expressed on many extracellular *A. phagocytophilum* cells after exiting the host cell (E and A) but not in the early stages of intracellular division at 16 h (B). Top panels (A to E): infected HL-60 cells stained with rabbit anti-*A. phagocytophilum* (red [TRITC]) and mouse anti-APH_1235 (green, Alexa 488); nuclei stained with DAPI (blue). Bottom panels (F to J): bright-field images of Giemsa-stained infected HL-60 cells at equivalent stages of infection. Bacteria stain purple; host nuclei stain pink. Bar, 5 μm.

Surface localization and antibody-mediated inhibition of APH_1235. To test whether APH_1235 is exposed to the bacterial cell surface, we isolated host cell-free *A. phagocytophilum* cells and digested the cells with trypsin for 30 min, following a protocol used for surface proteins of *Chlamydia trachomatis* elementary bodies (37). After separating total protein by SDS-PAGE, we probed Western blots with antibodies to recombinant APH_1235 and to p44. As seen in Fig. 4, the APH_1235 band was greatly reduced by trypsin digestion, while p44, an integral membrane protein with many transmembrane helices, was only slightly digested. The APH_1235 protein band migrates at 18 kDa and also shows a fainter band closer to its predicted size of 14.5 kDa. The same 18-kDa band was detected in infected tick salivary glands (not shown).

Since targeted mutagenesis is not currently feasible in *A. phagocytophilum*, we tested whether adding an antibody against APH_1235 might interfere with infection, as has been demonstrated with antibodies to p44 (msp2) in mammalian cells (29). We tested whether anti-APH_1235 IgG could inhibit bacterial binding to HL-60 cells but found no inhibition of binding (not shown). To test whether the anti-APH_1235 antibody could reduce infection, we cultured *A. phagocytophilum* in HL-60 cells in the presence of anti-APH_1235 IgG or preimmune IgG (20 μg/ml) for 4 days, allowing multiple rounds of infection. When the bacteria were propagated with anti-APH_1235 IgG, *A. phagocytophilum* levels were reduced by 25% to 50% compared to the control, as measured by qPCR of the 16S rRNA gene (Fig. 5). This inhibition of infection was most pronounced in the *A. phagocytophilum* stock that had been isolated directly from tick salivary glands (termed the "SG" isolate) and subjected to passage for several months in mammalian cells. When we repeated the experiment with the more rapidly infecting NCH-1 strain, which has been subjected to passage for years in HL-60 cells, the total *A. phagocytophilum* levels were up to 3 times higher on day 4, but there was not a clear reduction when the bacteria were incubated with anti-APH_1235 IgG (not shown). However, for the SG isolate, the presence of the antibody in the medium, where it presum-

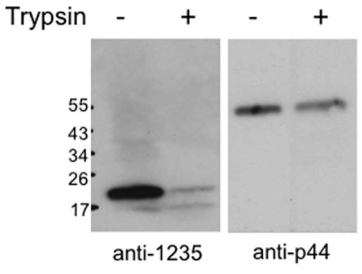

FIG 4 Surface trypsin digestion of host cell-free *A. phagocytophilum* indicates surface localization of APH_1235. Western blots show bacterial protein extracts with or without trypsin digestion, probed with anti-APH_1235 or anti-p44 antisera. After 30 min of surface trypsinolysis, the amount of APH_1235 was significantly reduced, while the p44 protein, an integral membrane protein, was only partially digested. Data are representative of the results of three independent experiments.

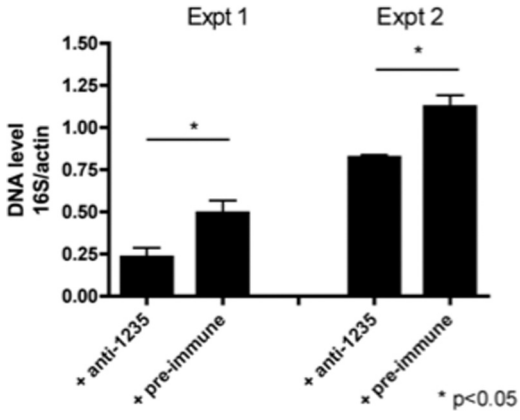

FIG 5 Propagation of *A. phagocytophilum* in the presence of anti-APH_1235 IgG reduces infection in HL-60 cells. Cultures were sampled at day 4 postinfection. Data represent the results of two independent experiments, each performed with three replicate samples. Statistical significance was assessed with a Student's *t* test.

ably contacts extracellular *A. phagocytophilum* cells, reduced total *A. phagocytophilum* infection over time, suggesting a role in entry or exit of host cells.

In summary, the APH_1235 protein is abundantly expressed during transmission from ticks to mice, can be digested on the bacterial cell surface by trypsin, and is most abundant in mammalian cell culture when *A. phagocytophilum* bacteria prepare to exit host cells. Adding anti-APH_1235 antibody to culture medium leads to a decrease in *A. phagocytophilum* infection levels, implying a role in escaping host cells or colonizing new cells.

DISCUSSION

The picture that emerges from these postgenomic analyses is that the extracellular, infectious DC form of *A. phagocytophilum* is the main cell type passed from ticks to mammalian hosts. We sampled ticks at 48 h of feeding, in the middle of the expected transmission time. Our results indicate that, at this time point, *A. phagocytophilum* cells have already undergone significant replication and no longer undergo extensive growth and division, as suggested by the lack of DNA replication enzymes and the small number of proteins involved in energy generation or biosynthesis. Rather, transcription and translation appear to be higher, and many proteins are directed to the cell surface. The p44 surface proteins were among the most abundant proteins detected (Table 2) and are known to be the major antigens on the *A. phagocytophilum* surface. The p44 proteins have conserved N- and C-terminal domains, with a hypervariable central domain. Note that in some cases, transcripts mapped to a region conserved between two or more p44 genes (see Table S1 in the supplemental material). Expression of distinct forms of p44, after recombination of paralogous loci at the *p44-18-ES* locus, enables the pathogen to evade adaptive immune defenses. The p44 proteins and other as-yet-unidentified adhesins facilitate interactions with host ligands, such as P-selectin glycoprotein 1 on neutrophils (29). In addition to roles in binding and entry of host cells, outer envelope proteins could also be important for maintaining cell structural integrity during transit from the tick to mammalian host cells, especially since *A. phagocytophilum* does not synthesize a typical peptidoglycan cell wall.

Concurrently, detection of components from the Sec, Tat, and T4S secretion systems suggests that protein secretion, particularly secretion of proteins that make up the cell envelope, is crucial during *A. phagocytophilum* transmission and that protein secretion pathways could be excellent targets for inhibiting transmission. Numerous Sec substrates and one T4S substrate were found, but the only protein predicted to contain a typical Tat secretion motif, APH_0401 (as predicted by Tatfind), was not observed in this study (32). APH_0401, an iron-sulfur subunit of ubiquinol-cytochrome *c* oxidase, is a likely candidate for Tat secretion, as redox proteins containing iron-sulfur clusters are often secreted via this pathway (22). The Tat pathway is required for host cell colonization in *Legionella pneumophila*, among other pathogenic microbes (7). Despite the lack of predicted substrates, it is worth noting that TatA, which forms the membrane-spanning channel of the Tat machinery in *E. coli* (12), was not detected in either of the HL-60 culture proteomics studies (23, 36).

Several other unique *A. phagocytophilum* proteins with defined roles in mammalian cell infection were detected in tick salivary glands, including AnkA, AptA, and *A. phagocytophilum* vacuolar membrane (AVM) proteins APH_0032 and APH_1387, but it is unclear whether they perform the same functions in the tick. The role of AnkA in the tick host has not been described, but this T4S effector may alter transcription of tick genes, as it does in mammalian cells (10). Future studies should uncover the mechanisms of host cell entry, modification, and exit shared by the two hosts.

The high expression of the DC-specific APH_1235 protein during tick feeding strongly suggests that the DC cell type prevails during transmission, which is logical in that the pathogen is exiting one host and preparing to enter its next host during that process. Transcriptional profiling comparing the equivalent intracellular ("RB") versus infectious ("EB") forms in *Ehrlichia ruminatum* also found that an APH_1235 ortholog, CDS_00640, was upregulated in the infectious EB form (30). Further evidence of the development of the DC form is provided by the detection of multiple chaperones and stress response proteins (see Table S1 in the supplemental material). The DC form, which must survive extracellularly before infecting new host cells, is known to be more resistant to osmotic stress and sonication stress than the intracellular RC form (35). Detection of a DNA-binding response regulator CtrA homolog, APH_0199, is also intriguing, as the CtrA homolog in the closely related pathogen *Ehrlichia chaffensis* is proposed to coordinate expression of late-stage genes in developing DC forms, especially those involved in resistance to environmental stress (5). In addition, the DNA-binding HU (histone-like) protein, detected by both proteomics and RNA-seq, may be involved in condensing the DNA of the DC form, potentially altering gene transcription, as shown in *E. coli* (20).

Examining proteins expressed in ticks by members of the *Anaplasmataceae* should lead to a better understanding of how bacteria interact with arthropod vectors. We compared the 139 proteins identified here to those reported to be upregulated by *A. phagocytophilum* in ISE6 tick cells and saw a low degree of overlap (28). Of the 41 ORFs whose transcription level was at least 2-fold higher in ISE6 tick cells than in HL-60 human cells, only two of these proteins, Msp4 (APH_1240) and the hypothetical protein APH_0723, were detected in our analysis. Differences in the environmental conditions in culture, or the growth stages sampled, may account for this discrepancy, as has been observed in other high-throughput analyses of pathogens sampled *in vivo* in hosts (2). One notable difference between the gene expression reported in the microarray study and the expression found in this *in vivo* analysis was the detection of the VirB2 paralogs APH_1144 and APH_1145 by RNA-seq in ticks; those two paralogs were not detected at all in ISE6 tick cells.

In the cattle pathogen *Anaplasma marginale*, the most closely related species with a sequenced genome, a proteomics study revealed 15 bacterial proteins that were upregulated in tick cells compared to mammalian cells. Ten of these had clear orthologs in the *A. phagocytophilum* genome, and one, APH_0149, a tripartite ATP-independent periplasmic (TRAP) transporter, was detected in this study. *A. marginale* colonizes *Dermacentor* ticks and infects bovine erythrocytes, but it is likely that the proteins required for interacting with tick hosts are found across *Anaplasma* spp. and other arthropod-transmitted *Rickettsiales* bacteria. Future work should elucidate whether such proteins are conserved at the sequence level or mainly at the functional level. Even those that are present only in *A. phagocytophilum* could be important for survival in the *Ixodes* tick and may represent functions performed by proteins with highly divergent amino acid sequences.

Due to detection limits, we are aware that the lack of detection

does not imply nonexpression of *A. phagocytophilum* proteins. Conversely, detection of a gene or protein does not prove activity. While this work does not purport to represent a complete inventory of all *A. phagocytophilum* proteins expressed in tick salivary glands, it highlights proteins that can be detected during tick transmission feeding and begins to define proteins required for survival during passage to mammalian hosts. In addition, these studies provide a valuable picture of how the pathogen "looks" in a particular host. Surprisingly, the two high-throughput approaches used yielded little overlap (10%), a phenomenon that has been reported for other parallel proteome and transcriptome studies (1). We believe that this discrepancy is due in large part to the limitations of the detection methods used—including the difficulty of detecting bacterially derived peptides and transcripts in a sample dominated by tick host material (see Table 1) and the differences in data processing and database search parameters between the two approaches—but it may also imply distinct expression patterns for protein and mRNA at this key transitional stage. Because proteins have much longer half-lives than mRNA in cells (25), the protein identifications could reflect proteins that are induced at an earlier stage and that are more important during development in the tick, while the transcripts identified may indicate gene expression needed later, in the mammalian host. This idea is supported by the high level of chaperone proteins, perhaps transcriptionally induced during early feeding, lacking mRNA transcripts at 48 h because stable proteins are already present (see Table S1 in the supplemental material). While such speculations are intriguing, additional sampling is required to determine to what extent the transcriptional and proteomic programs differ. Internal quality controls in our study include the similarity in the results of breakdowns of COG categories between the two methods, as well as the relatively high detection of p44 proteins and APH_1235 at both the mRNA and protein levels. Further enrichment of bacterially derived protein or mRNA could be achieved by physical separation of *A. phagocytophilum* cells from tick salivary glands, cDNA library synthesis using primers designed to preferentially amplify *A. phagocytophilum* genes, or selective capture of transcribed sequences (9).

The benefit of postgenomic strategies to investigate pathogen gene expression in hosts is clear, especially for revealing proteins whose expression is specific to the host environment rather than to the cultured cells (4, 21). By allowing us to sample an *in vivo* system during the transition between hosts, these techniques can both complement and inform culture-based studies. The high sensitivity of shotgun proteomics and RNA-seq approaches, combined with data from culture model systems, should clarify the changes in *A. phagocytophilum* gene expression across host environments and the functions of the many uncharacterized proteins in this species. This analysis provides strong evidence that the most relevant developmental stage to focus on for transmission studies is the DC form, which should provide a rich resource of protein targets for use in designing strategies to combat these and other vector-borne pathogens.

ACKNOWLEDGMENTS

We kindly thank Lindsay Rollend of the Durland Fish laboratory at Yale University for providing *A. phagocytophilum*-infected *I. scapularis* tick nymphs and thank Jason Carlyon at Virginia Commonwealth University for providing anti-p44 antiserum. We are grateful to Debby Beck, Aaron Bozzi, Kathleen Deponte, and Jingyi Pan for high-quality technical assistance.

This work was supported by National Institutes of Health grant R01AI041440 to E.F.

REFERENCES

1. **Basler M, et al.** 2006. The iron-regulated transcriptome and proteome of Neisseria meningitidis serogroup C. Proteomics **6**:6194–6206.
2. **Camejo A, et al.** 2009. In vivo transcriptional profiling of Listeria monocytogenes and mutagenesis identify new virulence factors involved in infection. PLoS Pathog. **5**:e1000449.
3. **Carlyon JA.** 2005. Laboratory maintenance of Anaplasma phagocytophilum. Curr. Protoc. Microbiol. 3A.2.1–3A.2.30.
4. **Cash P.** 2011. Investigating pathogen biology at the level of the proteome. Proteomics **11**:3190–3202.
5. **Cheng Z, Miura K, Popov VL, Kumagai Y, Rikihisa Y.** 2011. Insights into the CtrA regulon in development of stress resistance in obligatory intracellular pathogen Ehrlichia chaffeensis. Mol. Microbiol. **82**:1217–1234.
6. **Dahlgren FS, Mandel EJ, Krebs JW, Massung RF, McQuiston JH.** 2011. Increasing incidence of Ehrlichia chaffeensis and Anaplasma phagocytophilum in the United States, 2000–2007. Am. J. Trop. Med. Hyg. **85**:124–131.
7. **De Buck E, Lammertyn E, Anne J.** 2008. The importance of the twin-arginine translocation pathway for bacterial virulence. Trends Microbiol. **16**:442–453.
8. **Dunning Hotopp JC, et al.** 2006. Comparative genomics of emerging human ehrlichiosis agents. PLoS Genet. **2**:e21.
9. **Emboulé L, et al.** 2009. Innovative approach for transcriptomic analysis of obligate intracellular pathogen: selective capture of transcribed sequences of Ehrlichia ruminantium. BMC Mol. Biol. **10**:111.
10. **Garcia-Garcia JC, Rennoll-Bankert KE, Pelly S, Milstone AM, Dumler JS.** 2009. Silencing of host cell CYBB gene expression by the nuclear effector AnkA of the intracellular pathogen Anaplasma phagocytophilum. Infect. Immun. **77**:2385–2391.
11. **Ge Y, Rikihisa Y.** 2007. Identification of novel surface proteins of Anaplasma phagocytophilum by affinity purification and proteomics. J. Bacteriol. **189**:7819–7828.
12. **Gohlke U, et al.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc. Natl. Acad. Sci. U. S. A. **102**:10482–10486.
13. **Hirosawa M, Hoshida M, Ishikawa M, Toya T.** 1993. MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. Comput. Appl. Biosci. **9**:161–167.
14. **Huang B, et al.** 2010. Anaplasma phagocytophilum APH_0032 is expressed late during infection and localizes to the pathogen-occupied vacuolar membrane. Microb. Pathog. **49**:273–284.
15. **Huang B, et al.** 2010. Anaplasma phagocytophilum APH_1387 is expressed throughout bacterial intracellular development and localizes to the pathogen-occupied vacuolar membrane. Infect. Immun. **78**:1864–1873.
16. **Huang H, Wang X, Kikuchi T, Kumagai Y, Rikihisa Y.** 2007. Porin activity of Anaplasma phagocytophilum outer membrane fraction and purified P44. J. Bacteriol. **189**:1998–2006.
17. **Ishihama Y, et al.** 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol. Cell. Proteomics **4**:1265–1272.
18. **Jin H, Wei F, Liu Q, Qian J.** 4 January 2012, posting date. Epidemiology and control of human granulocytic anaplasmosis: a systematic review. Vector Borne Zoonotic Dis. [Epub ahead of print.] doi:10.1089/vbz.2011.0753.
19. **Katavolos P, Armstrong PM, Dawson JE, Telford SR III.** 1998. Duration of tick attachment required for transmission of granulocytic ehrlichiosis. J. Infect. Dis. **177**:1422–1425.
20. **Koli P, Sudan S, Fitzgerald D, Adhya S, Kar S.** 2011. Conversion of commensal Escherichia coli K-12 to an invasive form via expression of a mutant histone-like protein. mBio **2**:e00182–11.
21. **Kuriakose JA, Miyashiro S, Luo T, Zhu B, McBride JW.** 2011. Ehrlichia chaffeensis transcriptome in mammalian and arthropod hosts reveals differential gene expression and post transcriptional regulation. PLoS One **6**:e24136.

22. **Lee PA, Tullman-Ercek D, Georgiou G.** 2006. The bacterial twin-arginine translocation pathway. Annu. Rev. Microbiol. **60**:373–395.
23. **Lin M, Kikuchi T, Brewer HM, Norbeck AD, Rikihisa Y.** 2011. Global proteomic analysis of two tick-borne emerging zoonotic agents: Anaplasma phagocytophilum and Ehrlichia chaffeensis. Front. Microbiol. **2**:24.
24. **Lodes MJ, et al.** 2001. Serodiagnosis of human granulocytic ehrlichiosis by using novel combinations of immunoreactive recombinant proteins. J. Clin. Microbiol. **39**:2466–2476.
25. **Maier T, et al.** 2011. Quantification of mRNA and protein and integration with protein turnover in a bacterium. Mol. Syst. Biol. **7**:511.
26. **Mastronunzio JE, Benson DR.** 2010. Wild nodules can be broken: proteomics of Frankia in field-collected root nodules. Symbiosis **50**:13–26.
27. **Nagalakshmi U, Waern K, Snyder M.** 2010. RNA-Seq: a method for comprehensive transcriptome analysis. Curr. Protoc. Mol. Biol. **89**: 4.11.1–4.11.13.
28. **Nelson CM, et al.** 2008. Whole genome transcription profiling of Anaplasma phagocytophilum in human and tick host cells by tiling array analysis. BMC Genomics **9**:364.
29. **Park J, Choi KS, Dumler JS.** 2003. Major surface protein 2 of Anaplasma phagocytophilum facilitates adherence to granulocytes. Infect. Immun. **71**:4018–4025.
30. **Pruneau L, et al.** 2012. Global gene expression profiling of Ehrlichia ruminantium at different stages of development. FEMS Immunol. Med. Microbiol. **64**:66–73.
31. **Rikihisa Y, Lin M, Niu H.** 2010. Type IV secretion in the obligatory intracellular bacterium Anaplasma phagocytophilum. Cell Microbiol. **12**: 1213–1221.
32. **Rose RW, Bruser T, Kissinger JC, Pohlschroder M.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol. Microbiol. **45**:943–950.
33. **Simpson JT, et al.** 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. **19**:1117–1123.
34. **Sukumaran B, et al.** 2011. Anaplasma phagocytophilum AptA modulates Erk1/2 signalling. Cell. Microbiol. **13**:47–61.
35. **Troese MJ, Carlyon JA.** 2009. Anaplasma phagocytophilum dense-cored organisms mediate cellular adherence through recognition of human P-selectin glycoprotein ligand 1. Infect. Immun. **77**:4018–4027.
36. **Troese MJ, et al.** 2011. Proteomic analysis of Anaplasma phagocytophilum during infection of human myeloid cells identifies a protein that is pronouncedly upregulated on the infectious dense-cored cell. Infect. Immun. **79**:4696–4707.
37. **Wang Y, et al.** 2006. Identification of surface-exposed components of MOMP of Chlamydia trachomatis serovar F. Protein Sci. **15**:122–134.
38. **Zhi N, et al.** 2002. Transcript heterogeneity of the p44 multigene family in a human granulocytic ehrlichiosis agent transmitted by ticks. Infect. Immun. **70**:1175–1184.
39. **Zhou M, Wu H.** 2009. Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. Microbiology **155**:317–327.