Research article

Restricted changes in major surface protein-2 (*msp2*) transcription after prolonged in vitro passage of *Anaplasma phagocytophilum*

Diana G Scorpio[1,2], Karen Caspersen[1,2], Hiroyuki Ogata[3], Jinho Park[1,4] and J Stephen Dumler*[1]

Address: [1]Division of Medical Microbiology, Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore MD, 21205, USA, [2]Division of Comparative Medicine, The Johns Hopkins University School of Medicine, Baltimore MD, 21205, USA, [3]Information Génomique et Structurale, CNRS UPR 2589, Marseille, France and [4]Department of Veterinary Internal Medicine, College of Veterinary Medicine, Chonbuk National University, Jeonju Jeonbuk, Korea

Email: Diana G Scorpio - dscorpio@jhmi.edu; Karen Caspersen - casp_21117@yahoo.com; Hiroyuki Ogata - hiroyuki.Ogata@igs.cnrs-mrs.fr; Jinho Park - jpark@mail.chonbuk.ac.kr; J Stephen Dumler* - sdumler@jhmi.edu

* Corresponding author

Abstract

Background: *Anaplasma phagocytophilum* strains often vary in Msp2 expression, a situation assumed to be related to immune evasion. However, Msp2 is also an adhesin, and little is known about the role of endogenous *msp2* transcriptional changes in the absence of immune selection. Thus, Msp2 profiles and *msp2* transcripts of low passage *A. phagocytophilum* Webster strain, initially comprised of a single abundant *msp2* transcript, were re-examined after ≥ 20 in vitro passages without immune selection.

Results: Using an Msp2 monoclonal antibody, immunoblots revealed an unchanged dominant band and several weak bands that appeared with passage. Similarly, *msp2* transcript diversity changed, with a decrease in the initially abundant low passage transcript and appearance of a newly abundant and several minor *msp2* transcripts with high passage. BLASTN search of the *A. phagocytophilum* HZ strain genome revealed ≥ 52 *msp2* paralogs.

Conclusions: Msp2 expression and *msp2* transcription modulate even without immune selective pressures. However, the limited diversity of *msp2* transcripts in the absence of immune pressure suggests selection for Msp2 by specific functions beyond that of immune evasion, in spite of a large genomic reservoir for Msp2 diversity.

Background

Anaplasma phagocytophilum is the causative agent of human granulocytic anaplasmosis (formerly named HGE) and granulocytic anaplasmosis in animals. Human, horse, and canine infections are acute and self-limiting [1-3], whereas ruminant and rodent infections may persist [4,5]. As for *Anaplasma marginale*, *A. phagocytophilum* strains from different geographic regions are antigenically variable [6], and the emergence of antigenic variants in

vitro and in vivo suggests a role in immune evasion and persistent infection [7-10]. Antigenic variability results from differential expression of immunodominant surface proteins (Msp2, p44, 44 kDA), encoded by a multigene family [10-12]. Msp2 expression is polymorphic in vitro, in mice, and in ticks [13,14]. Based in part on these investigations where changes were characterized under immune selective pressure or in poorly defined polyclonal populations in vitro, the prevailing opinion is that

antigenic diversity is generated as a mechanism for immune evasion. However, among 18 unique *msp2* paralogs detected in uncloned low passage *A. phagocytophilum* Webster strain, only 1 of 7 transcripts was dominant in number [15]. With the discovery that Msp2 is an *A. phagocytophilum* adhesin for granulocytes, we hypothesized that Msp2s are also selected by non-immunological processes based on fitness for specific niches, such as growth in mammalian granulocytes. Before such hypotheses can be addressed, it is important to characterize and monitor endogenous changes in *msp2* transcription in the absence of selective pressures under defined limits of in vitro propagation. Thus, *msp2* transcripts of *A. phagocytophilum* Webster strain were investigated after 20 in vitro passages, compared to transcripts characterized after only 7 in vitro passages.

Results

Msp2 immunoblot

To ascertain whether Msp2 expression changes with in vitro passage, immunoblotting with a broad range Msp2 monoclonal antibody was performed. Msp2 was identified after 17 passages in vitro as an intense band and a weaker band at approximately 48 and 30 kDa, respectively. After 23 passages, additional weak bands at 46, 43, 40 and 37.5 kDa were detected, and the 30 kDa band became less intense [Figure 1].

RT-PCR amplification and sequencing of msp2 transcript amplicons

To detect changes in transcription after in vitro passage, RT-PCR using *msp2* universal primers was performed at 7 and 20 in vitro passages. A 550-bp RT-PCR product was amplified from high passage *A. phagocytophilum* Webster RNA template, and no DNA was amplified. Of 64 clones, 58 were successfully sequenced. When aligned with the previous 18 unique *A. phagocytophilum* Webster strain *msp2* paralogs [15], 4 additional *msp2* paralogs with unique hypervariable regions [Figure 2] were identified. Of all the sequences from these clones, 28 (48.3%) were transcribed from one newly identified *msp2*. Each of the 3 other new sequences were identified only once (1.7%) among the 58 transcript clones. The remaining 27 clones (46.6%) had sequences identical to those previously identified among the *msp2*s transcribed in passage 7, including 15 (28%) derived from the predominant paralog transcribed in low passage. When aligned with predicted open reading frames for fully sequenced *msp2*, amino acid translations identified conserved N- and C-terminal peptides flanking a hypervariable domain ranging from 70 to 79 residues.
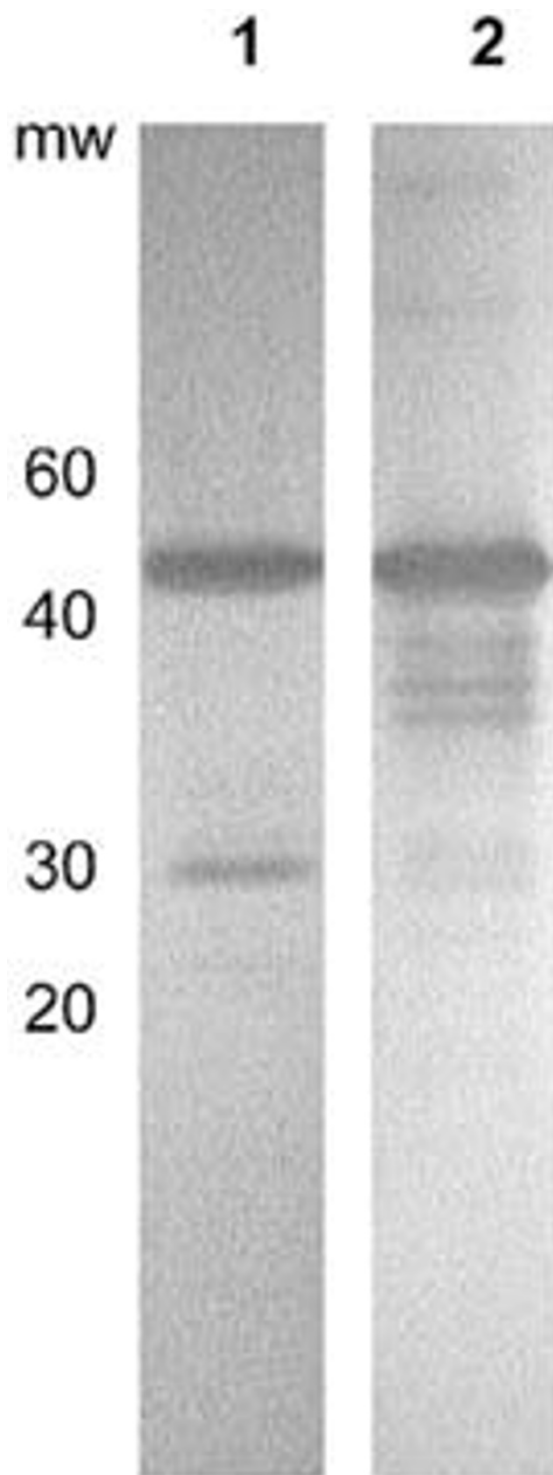

Figure 1
Protein immunoblot comparing Msp2 expression at 17 vs. 23 in vitro passages. Note the appearance of a dominant 48 kDa band in both lanes, the presence of a 30 kDa band after 17 passages that is very weakly detected at 23 passages, and the appearance of new bands at 37.5, 40, 43, and 46 kDa
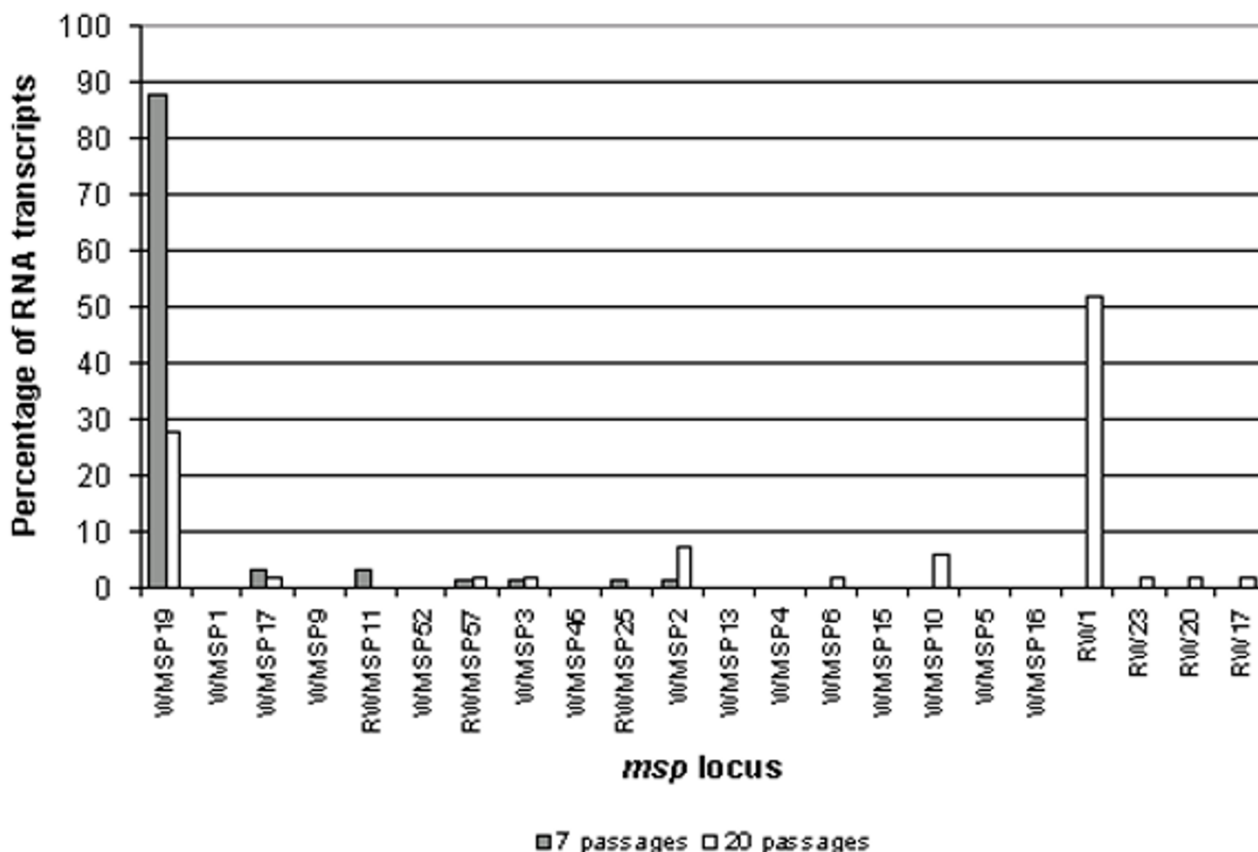
Figure 2
Histogram of 22 *msp2* loci and percentage of total RNA transcripts detected at low and high passage for *A. phagocytophilum*. Frequency of *msp2* transcript detection in low passage (7 passages in vitro, represented by shaded bars) [15] and high passage (20 passages in vitro, represented by open bars) *A. phagocytophilum* Webster strain for each of 22 distinct loci detected.

The origin and characterization of new msp2 *transcripts*
It was possible that the 4 newly identified *msp2* transcripts in high passage cultures resulted from recombination of existing *msp2*s. However, by targeting hypervariable regions in DNA PCR, each new *msp2* paralog was detected in the genome of the nearly clonal low passage *A. phagocytophilum*, establishing their a priori existence. In addition, each new *msp2* was subjected to BLASTN against the 18 low passage partial *msp2*s [15] and low E value hits were examined by pairwise alignment for blocks of shared nucleic acids. Overall sequence identities were between 69 and 74% with the best matches. When 5', 3', and hypervariable regions were examined separately, sequence identities between 78 and 84% were obtained for 5' regions, between 86 and 96% for 3' regions, and between 49 and 69% for hypervariable regions, reflecting a lack of large

recombined *msp2* regions that would result in a simple mosaic appearance. Sequence dissimilarity was distributed throughout the entire hypervariable regions, mostly resulting from single or short stretches of insertions or deletions [Figure 3].

Translation of hypervariable region sequences revealed a similar level of divergence in multiple positions except for a high degree of conservation in only 2 positions. Compared with the amino acid sequence of *A. phagocytophilum* GenBank accession no. AF029323 [16], all hypervariable domains had a conserved Cys at residue 124, and a conserved consensus Lys-Asp-Trp-Pro-Thr-Ser/Gly motif at residues 148 through 153, as described by Zhi [14]. The latter residues were preceded by a predicted helical structure in 17 of 22 paralogs. The C-terminus of each
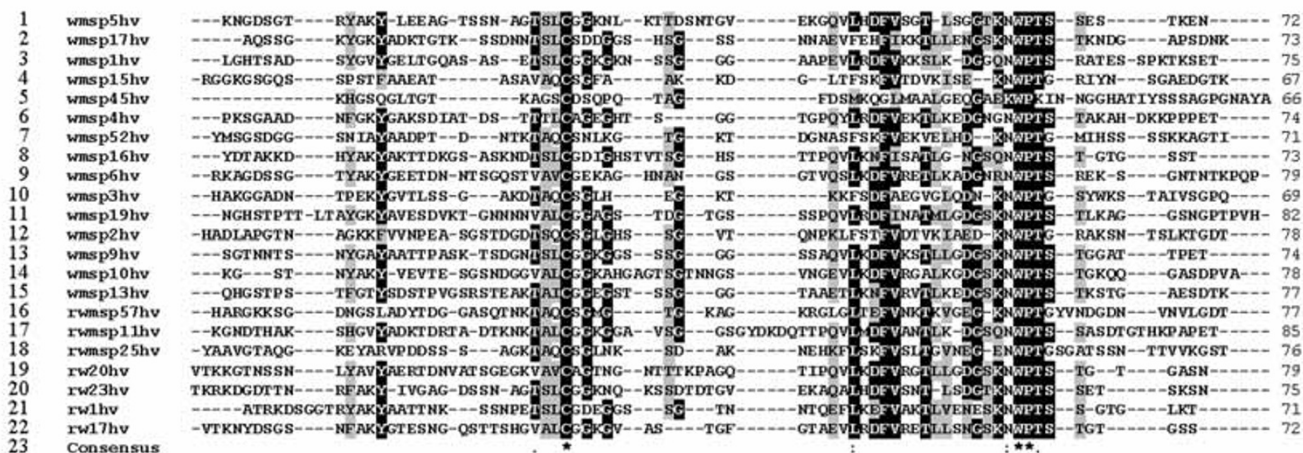
```
 1  wmsp5hv    ---KNGDSGT----RYAKY-LEEAG-TSSN-AGTSLCGGKNL--KTTDSNTGV------EKGQVLHDFVSGTLSGCTKNWPTS--SES------TKEN------ 72
 2  wmsp17hv   ----AQSSG----KYGKYADKTGTK--SSDNNTSLCGSDDGGS--HSG---SS------NNAEVFEHSIKKTLLENCSKNWPTS--TKNDG----APSDNK---- 73
 3  wmsp1hv    ---LGHTSAD----SYGVYGELTGQAS-AS--ETSLCGSKGKN--SSG---GG------AAPEVLRDFVKKSLK-DCGQNWPTS--RATES-SPKTKSET----- 75
 4  wmsp15hv   -RGGKGSGQS----SPSTFAAEAT------ASAVAQCSFA-----AK---KD--G--LTFSKSVTDVKISE--KNWPTG--RIYN---SGAEDGTK---- 67
 5  wmsp45hv   ------------KHGSQGLTGT--------KAGSCDSQPQ--TAG--------FDSMKQGLMAALGEOCAKWPKIN-NGGHATIYSSSAGPGNAYA 66
 6  wmsp4hv    ---PKSGAAD----NFCKYGAKSDIAT-DS--TPTILCASEGHT--S---GG------TGPQYLRDFVEKTLKEDGNGNWPTS--TAKAH-DKKPPPET---- 74
 7  wmsp52hv   --YMSGSDGG----SNIAYAADPT--D---NTKTAQCSNLKG----TG---KT------DGNASFSKSVEKVELHD--KNWPTG--MIHSS--SSKKAGTI---- 71
 8  wmsp16hv   ---YDTAKKD----HYAKYAKTTDKGS-ASKNDTSLCGDIGHSTVTSG--HS------TTPQVLKNFISATLG-NCSQNWPTS--T-GTG----SST------- 73
 9  wmsp6hv    ---RKAGDSSG----TYAKYGEETDN-NTSGQSTVAVCGEKAG--HNAN---GS------GTVQSLKDFVRETLKADGNRNWPTS--REK-S---GNTNTKPQP- 79
10  wmsp3hv    --HAKGGADN----TPEKYGVTLSS-G---AKDTAQCSLH-----EG---KT------KKFSDFAEGVGLQDN-KNWPTG--SYWKS--TAIVSGPQ---- 69
11  wmsp19hv   ---NGHSTPTT-LTAYGKYAVESDVKT-GNNNNVALCGFAGS--TDG-TGS------SSPQVLRDFINATMLGDCSKNWPTS--TLKAG----GSNGPTPVH- 82
12  wmsp2hv    --HADLAPGTN----AGKKFVVNPEA-SGSTDGDTSQCSCLGHS--SSG--VT------QNPKLFSTFVDTVKIAED-KNWPTG--RAKSN--TSLKTGDT---- 78
13  wmsp9hv    ---SGTNNTS----NYGAYAATTPASK-TSDGNTSLCGCGGS--SSG---GG------SSAQVLKDFVKSTLLGDCSKNWPTS--TGGAT---TPET------ 74
14  wmsp10hv   ---KG---ST----NYAKY-VEVTE-SGSNDGGVALCGKAHGAGTSGTNNGS------VNGEVLKDFVRGALKGDGSKNWPTS--TGKQQ---GASDPVA--- 78
15  wmsp13hv   ---QHGSTPS----TFGTYSDSTPVGSRSTEAKTAICGCEGST--SSG---GG------TAAETLKNFVRVTLKEDCSKNWPTS--TKSTG---AESDTK---- 77
16  rwmsp57hv  --HARGKKSG----DNGSLADYTDG-GASQTNKTAQCSKG----TG--KAG------KRGLGLTEFVNKTKVGEDCGKNWPTGYVNDGDN---VNVLGDT--- 77
17  rwmsp11hv  --KGNDTHAK----SHGVYADKTDRTA-DTKNKTAICGKGGA--VSG---GSGYDKDQTTPQVLDFVANTLK-DCSQNWPTS--SASDTGTHKPAPET----- 85
18  rwmsp25hv  --YAAVGTAQG----KEYARVPDDSS-S---AGKTAQCSCLNK---SD--AK------NEHKFLSKFVSLTVKILVENSKNWPTS--GSGATSSN--TTVVKGST---- 76
19  rw20hv     VTKKGTNSSN----LYAVYAERTDNVATSGEGKVAVCAGTNG--NTTTKPAGQ------TIPQVLKDFVRGILLGDCSKNWPTS--TG--T----GASN------ 79
20  rw23hv     TKRKDGDTTN----RFAKY-IVGAG-DSSN-AGTSLCGGKNQ--KSSDTDTGV------EKAQALHDFVSNTLSDCTKNWPTS--SET------SKSN------ 75
21  rw1hv      -----ATRKDSGGTRYAKYAATTNK--SSNPETSLCGDEGGS---SG---TN------NTQEFLKEFVLVENESKNWPTS--S-GTG----LKT------ 71
22  rw17hv     -VTKNYDSGS----NFAKYGTESNG-QSTTSHGVALCGKGV--AS---TGF------GTAEVLRDFVRETLLSHGSKNWPTS--TGT------GSS------- 72
23  Consensus                               .       *                            :  **.
```

Figure 3
ClustalX alignment and consensus amino acid sequences of Msp2 hypervariable domains. Domains were translated from genomic and RT-PCR amplicons of *A. phagocytophilum* Webster strain *msp2* after 7 and 20 passages in vitro. The alignment demonstrates the heterogeneity of the hypervariable domains and the lack of significant similarities that would suggest combinatorial gene conversion as a mechanism for ongoing Msp2 antigenic variation. Paralog names are listed on the left, the total number of amino acid residues is shown on the right; the consensus sequence is shown on the bottom – "." indicates a semi-conserved position, ":" indicates a highly conserved position, and "*" indicates a conserved position.

translated amplicon had 2 tandem helical hydrophobic domains, and the N-terminus of each shared 2 consecutive helical domains and 2 tandem beta strand-like elements.

Estimation of msp2 *content in* A. phagocytophilum *strain HZ incomplete genome sequences*

A TBLASTN search against the *A. phagocytophilum* strain HZ genomic DNA sequences detected 219 homologous segments in 19 different contigs with E-values <0.001. The average and median sizes of contigs were 62 and 20 kb, respectively. From the TBLASTN output, 86 translated amino acid sequences ≥ 100 residues and uninterrupted by in-frame stop codons were extracted. Because *msp2* has conserved 5' and 3' ends, some segments might have originated from one gene; thus, one from each sequence pair within 1 kb on the same strand of a single contig was removed leaving 57 sequences. One from each sequence pair that had identical amino acid sequences was also removed, resulting in 54 remaining segments. The average pair-wise sequence identity among these was 59%, similar to that of *A. phagocytophilum* Webster strain paralogs. Finally, BLASTP searches carried out against the SWISS-PROT/TrEMBL non-redundant sequence database [16] confirmed that 52 of 54 sequences were best-matched with known Msp2s. One was best-matched to *Ehrlichia chaffeensis* major outer membrane protein OMP-1A

(TrEMBL accession no. O52102) and the other to *Anaplasma ovis* major surface protein 4 (Q8KWV5).

Discussion

Msp2 is the immunodominant protein expressed in both *A. marginale* and *A. phagocytophilum*. The encoding *msp2*s are believed to allow the generation of antigenic variation and persistence in reservoir hosts [8-12,15,17,18]. In addition, *A. phagocytophilum* Msp2s are heterogeneous among geographically distinct isolates [6] and this diversity may be important for infection in reservoirs such as rodents, ruminants, and cervids [4,5,19]. The mechanisms of persistence of *Anaplasma marginale* and many other microorganisms involve gene recombination or conversion and expression of new antigens that allow immune evasion [20-24]. However, another hypothesis is whether a non-immunologic functional role for individual *msp2* paralogs influences selection when expressed in specific niches such as the mammalian or tick host [13,14]. Experiments conducted under immune selective pressure, such as in vivo mammalian infection, or with cultures of undefined bacterial clonality are unable to provide supportive data for such alternate roles. Previous characterization showed that very low passage strains of *A. phagocytophilum* possess a single abundant *msp2* transcript [15], which coupled with new data confirming a single expression site for *A. phagocytophilum* [25], indicates a relative uniform clonal population. If immune selection

is the driving force for Msp2 diversity, propagation in vitro in the absence of immune pressure should allow the generation of multiple Msp2 variants. Thus, starting with a low passage, nearly clonal strain of *A. phagocytophilum*, changes in endogenous *msp2* transcription were analyzed when immune selection was absent.

The identification of 4 additional unique *msp2* paralog transcripts in *A. phagocytophilum* Webster strain after prolonged in vitro passage without immune selection suggests a role for either combinatorial recombination or gene conversion, as described for *A. marginale* [21]. However, all 4 new *msp2* genes were identified in the genome of clonal low passage *A. phagocytophilum* excluding the possibility that each was generated by a new recombination event. Although proof of a role for *A. phagocytophilum msp2*s in immune evasion is lacking, changing Msp2 expression does alter antigenicity, and could contribute to evasion [13,14]. However, only minimal antigenic changes were observed, and these could equally reflect structural changes anticipated to occur with selection of alternate adhesins.

It is strongly suspected that transcription of *msp2* occurs from perhaps a single site in the chromosome, and that *msp2* transcriptional changes imply at minimum, emergence of new populations after gene conversion [7,8,25]. The incomplete *A. phagocytophilum* strain HZ genome revealed more than 50 *msp2* paralogs and/or pseudogenes of over 100 amino acids. This suggests that a vast capacity for Msp2 diversity is already "hard-wired" and that only simple gene conversion into a transcriptional site is required. Regardless, these findings illustrate that even without immune pressure, spontaneous genetic changes can accrue, but do so in a very limited fashion. It may be that human infection selects for certain Msp2s, explaining the abundance of a single or small number of *msp2* transcripts in human HL-60 cell cultures, whereas other Msp2s may more efficiently bind tick cells or neutrophils in reservoir animals. In partial support of this concept, Carlyon et al. recently described an alternative murine neutrophil ligand that differs in scaffold protein from platelet selectin glycoprotein-1 (PSGL-1) on human neutrophils [26].

What evidence is there that Msp2 has a role other than in immune evasion? Although *msp2* transcript diversity increases with mammalian cell passage, the effects are not neutral. Given the implied requirement for frequent emergence of variation to resist mammalian immunity [21], one might predict that a larger diversity and a more uniform frequency of *msp2* transcripts would appear. Instead, prolonged passage is associated with the appearance of only 2 major *msp2* transcripts over 20 or more passages – an interval usually greater than 4 months of continuous in

vitro propagation. Moreover, despite the relative lack of selective pressures, the absence of any transcripts that would indicate significant gene recombination suggests that *msp2*s of individual strains are moderately stable, at least in vitro. Taken together, these data further suggest that some Msp2s have specific functions that are selected for by environmental conditions and niches. Indeed, both Zhi et al. and IJdo et al. have also demonstrated a restricted profile of *msp2* transcripts under several conditions including tick infections [13,14].

An important function of Msp2 may be to provide bacterial antigenic diversity wherein immune recognition of the abundant bacterial population is targeted. Antigenically diverse bacterial clones temporarily survive and contribute to reservoir persistence. However, the high level of expression, the conserved flanking regions with relative conservation of structure, and highly conserved residues within the hypervariable domain suggest the potential for other roles, such as cell adhesion [27-29]. Moreover, the possibility that changes in Msp2 expression are associated with changes in *A. phagocytophilum* pathogenicity has not been investigated. Low passage *A. phagocytophilum* Webster strain causes disease in horses similar to natural human infection, but prolonged in vitro propagation, which leads to changes in *msp2* transcription and Msp2 expression, attenuates clinical signs in the horse model [30]. The change in virulence may ultimately be unrelated to changes in Msp2, but as host immune response is the predominant correlate of histopathologic lesions and disease [31,32], this hypothesis and the functions of specific Msp2s in various niches warrant further investigation.

Conclusions
Msp2 expression and *msp2* transcription in *A. phagocytophilum* display limited diversity under nonimmune-selective conditions despite long intervals of passage, suggesting that Msp2 has other functions beyond generation of antigenic variation for immune evasion. With the recognition of its role in adhesion to host cells, Msp2 diversity could also be an important adaptation to alternate niches such as small mammal or cervid reservoir neutrophils or tick cells.

Methods
A. phagocytophilum *in vitro cultivation and immunoblots*
A. phagocytophilum (Webster strain) was propagated in HL-60 cells [6]. Infected cells were harvested for nucleic acid preparation [15] and bacterial density gradient purification [6]. Protein immunoblotting [6] used Msp2 monoclonal antibody 20B4 that reacts with all *A. phagocytophilum* strains tested. Webster strain in vitro passage 20 *msp2* RNAs were compared to those obtained at in vitro passage 7 [15]. To demonstrate changes over prolonged passage, a sample of convenience that compared

Table 1: *msp2* hypervariable region primers to determine newly identified msp2 paralogs in low passage *A. phagocytophilum* genomic DNA.

msp2 transcript locus	Forward primer sequence (name)	Reverse primer sequence (name)
msp2 (pan-*msp2*)	5'-TGATGTTGTTACTGGACAGA-3' (MSP465f)	5'-CACCTAACCTTCATAAGAA-3' (MSP980r)
RW1	5'-GCGACGAGGAAGGACAGTGGT-3'	5'-AGTCTTCAACCCAGTCCCGCT-3'
RW17	5'-GTGACGAAGAATTACGACAGT-3'	5'-GCTAGAAGATCCGGTTCCTGT-3'
RW20	5'-GTGACGAAGAAGGGGACAAAC-3'	5'-GTTAGAGGCTCCGGTTCCTGT-3'
RW23	5'-ACAAAGCGGAAGGATGGTGAT-3'	5'-ATTCGATTAGACGTTTCACT-3'

Msp2 expression at passage 17 with passage 23 was examined by protein immunoblotting. Cell cultures were harvested when >90% of cells were infected, and identical protein quantities were loaded onto SDS-PAGE gels for comparison.

PCR and RT-PCR

DNA was prepared from infected HL-60 cells using the QIAamp DNA Mini Kit (Qiagen, Valencia, CA). RNA was prepared from infected cells using the RNeasy Mini Kit and a QIAshredder (Qiagen, Valencia, CA). Genomic DNA contamination was eliminated by the use of RNase-free DNase (Qiagen, Valencia, CA) during extraction and again after preparation of the final RNA. DNA and RNA concentrations of each preparation were determined.

Primers for amplification of *msp2* were created from the alignment of 10 *A. phagocytophilum msp2* sequences in GenBank, as previously described [15]. Primers [Table 1] were used in both PCR and RT-PCR reactions at a final concentration of 1 μM. Amplification from *A. phagocytophilum* nucleic acids was determined by PCR using DNA from the Webster strain and RT-PCR was performed on RNA also from the Webster strain. DNA PCR was performed as previously described [15], except for the use of primers designed to specifically amplify the hypervariable regions in the newly identified *msp2* transcripts [Table 1]. The specificity of all primers was tested by using *E. coli* and *Ehrlichia chaffeensis* DNA templates.

For RT-PCR, a one step PCR protocol using RT-Taq (SUPERSCRIPT One-Step RT-PCR, Life Technologies, Gaithersburg, MD) and 30 ng of RNA from infected cells was employed as per manufacturer's recommendations [15]. DNA contamination in the RNA preparations was assessed by the same PCR method substituting Taq for RT-Taq polymerase. Agarose gel electrophoresis and ethidium bromide staining were again conducted to assess a 550-bp cDNA fragment.

Cloning and sequencing of PCR products

cDNA RT-PCR products were cloned into the pCR4-TOPO vector (Invitrogen, Carlsbad, CA) by transformation into TOP 10 *E. coli* plated onto LB agar (50 μg/ml kanamycin). Plasmids from each of 64 colonies (Wizard *Plus* Minipreps DNA purification System, Promega, Madison, WI) were assessed for insert size after *Eco*RI digestion.

Sequence analysis of recombinant clones

Plasmid inserts were cycle-sequenced using the Big Dye Terminator Cycle Sequencing Ready Reaction (Applied Biosystems, Foster City, CA) and was performed [15] with 175 ng of plasmid DNA and M13 forward and reverse primers. Products were purified using Centrisep spin columns (Princeton Separations, Adelphia, NJ) before sequence determination (ABI Model 377, Foster City, CA), assembly (Assemgel, PCGENE, Intelligenetics, Mountain View, CA), alignment (CLUSTALW v1.4) [33], and visualization of phylodendograms (TreeView v1.6.1) [34]. Some nucleotide comparisons used BLASTN analysis [35]. Translated protein secondary structure was predicted with NNPREDICT http://www.cmpharm.ucsf.edu/cgi-bin/nnpredict.pl[36].

Analysis of msp2 in A. phagocytophilum strain HZ incomplete genome sequences

Using the amino acid sequence of *A. phagocytophilum* Msp2 (GenBank accession no. AF037599) as query, the *A. phagocytophilum* strain HZ unfinished genomic sequences were searched using TBLASTN on the Institute for Genomic Research server (TIGR; http://www.tigr.org). The shotgun phase of sequencing is complete, and as of February 6, 2003, the data included 54 contigs totaling 1,507,831 bp. BLASTP searches against the SWISS-PROT/TrEMBL non-redundant sequence database [16] were used to confirm best matches for TBLASTN-identified sequence segments.

Nucleotide sequence accession numbers

The nucleotide sequences of the 4 newly identified Webster strain *msp2* hypervariable regions are deposited in GenBank under accession numbers: AY253529, AY253530, AY253531, AY253532.

Authors contributions

DGS performed molecular biological techniques and sequence analysis, alignment, interpretation, and manuscript preparation. KC performed some amplification and cloning reactions and assisted in initial design. HO performed bioinformatics manipulations to assess the genomic content of the *A. phagocytophilum* HZ strain. JHP assisted in design of the study and assisted with molecular techniques. JSD was responsible for overall design of the study, analysis and interpretation of data, and assisted with manuscript preparation.

Acknowledgements

References

1. Greig B, Asanovich KM, Armstrong PJ, Dumler JS: **Geographic, clinical, serologic, and molecular evidence of granulocytic ehrlichiosis, a likely zoonotic disease, in Minnesota and Wisconsin dogs.** *J Clin Microbiol* 1996, **34:**44-48.
2. Madigan JE, Gribble D: **Equine ehrlichiosis in northern California: 49 cases (1968–1981).** *J Am Vet Med Assoc* 1987, **190:**445-448.
3. Bakken JS, Krueth J, Wilson-Nordskog C, Tilden RL, Asanovich K, Dumler JS: **Clinical and laboratory characteristics of human granulocytic ehrlichiosis.** *JAMA* 1996, **275:**199-205.
4. Stafford KC, Massung RF, Magnarelli LA, IJdo JW, Anderson JF: **Infection with agents of human granulocytic ehrlichiosis, Lyme disease, and babesiosis in wild white-footed mice (*Peromyscus leucopus*) in Connecticut.** *J Clin Microbiol* 1999, **37:**2887-2892.
5. Stuen S, Engvall EO, Artursson K: **Persistence of *Ehrlichia phagocytophila* infection in lambs in relation to clinical parameters and antibody responses.** *Vet Rec* 1998, **143:**553-555.
6. Asanovich KM, Bakken JS, Madigan JE, Aguero-Rosenfeld M, Wormser GP, Dumler JS: **Antigenic diversity of granulocytic *Ehrlichia* species isolates from humans in Wisconsin, New York, and a California horse.** *J Infect Dis* 1997, **176:**1029-1034.
7. Barbet AF, Lundgren A, Yi J, Rurangirwa FR, Palmer GH: **Antigenic variation of the ehrlichia *Anaplasma marginale* by expression of MSP2 sequence mosaics.** *Infect Immun* 2000, **68:**6133-6138.
8. Brayton KA, Knowles DP, McGuire TC, Palmer GH: **Efficient use of a small genome to generate antigenic diversity in tick-borne ehrlichial pathogens.** *Proc Natl Acad Sci USA* 2001, **98:**4130-4135.
9. French DM, Brown WC, Palmer GH: **Emergence of *Anaplasma marginale* antigenic variants during persistent rickettsemia.** *Infect Immun* 1999, **67:**5834-5840.
10. Murphy CL, Storey JR, Recchia J, Doros-Richert LA, Gingrich-Baker C, Munroe K, Bakken JS, Coughlin RT, Beltz GA: **Major antigenic proteins of the agent of human granulocytic ehrlichiosis are encoded by members of a multigene family.** *Infect Immun* 1998, **66:**3711-3718.
11. IJdo JW, Zhang Y, Magnarelli LA, Fikrig E: **Cloning of the gene encoding the 44-kilodalton antigen of the agent of human granulocytic ehrlichiosis and characterization of the humoral response.** *Infect Immun* 1998, **66:**3264-3269.
12. Zhi N, Ohasi N, Rikihisa Y: **Multiple *p44* genes encoding major membrane proteins are expressed in human granulocytic ehrlichiosis agent.** *J Biol Chem* 1999, **274:**17828-17836.
13. IJdo JW, Wu C, Telford SR III, Fikrig E: **Differential expression of the *p44* gene family in the agent of human granulocytic ehrlichiosis.** *Infect Immun* 2002, **70:**5295-5298.
14. Zhi N, Ohashi N, Tajima T, Mott J, Stich RW, Grover D, Telford SR III, Lin Q, Rikihisa Y: **Transcript heterogeneity of the *p44* multigene family in a human granulocytic ehrlichiosis agent transmitted by ticks.** *Infect Immun* 2002, **70:**1175-1184.
15. Caspersen K, Park J, Patil S, Dumler JS: **Genetic variability and stability of *Anaplasma phagocytophila* msp2 (*p44*).** *Infect Immun* 2002, **70:**1230-1234.
16. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M: **The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003.** *Nucleic Acids Res* 2003, **31:**365-370.
17. Palmer GH, Rurangirwa FR, McElwain TF: **Strain composition of the ehrlichia *Anaplasma marginale* within persistently infected cattle, a mammalian reservoir for tick transmission.** *J Clin Microbiol* 2001, **39:**631-635.
18. Rurangirwa FR, Stiller D, Palmer GH: **Strain diversity in major surface protein 2 during tick transmission of *Anaplasma marginale*.** *Infect Immun* 2000, **68:**3023-3027.
19. Walls JJ, Asanovich KM, Bakken JS, Dumler JS: **Serologic evidence of a natural infection of white-tailed deer with the agent of human granulocytic ehrlichiosis in Wisconsin and Maryland.** *Clin Diagn Lab Immunol* 1998, **5:**762-765.
20. Barbour AG, Restrepo BI: **Antigenic variation in vector-borne pathogens.** *Emerg Infect Dis* 2000, **6:**449-457.
21. Brayton KA, Palmer GH, Lundgren A, Yi J, Barbet AF: **Antigenic variation of *Anaplasma marginale* msp2 occurs by combinatorial gene conversion.** *Mol Microbiol* 2002, **43:**1151-1159.
22. Reddy GR, Sulsona CR, Barbet AF, Mahan SM, Burridge MJ, Alleman AR: **Molecular characterization of a 28 kDa surface antigen gene family of the tribe *Ehrlichieae*.** *Biophys Res Comm* 1998, **247:**636-643.
23. Reddy GR, Streck CP: **Variation in the 28 kDa surface antigen protein multigene locus of isolates of the emerging disease agent *Ehrlichia chaffeensis* suggests that it plays a role in immune evasion.** *Mol Cell Biol Res Comm* 1999, **1:**167-175.
24. Taylor-Robinson A, Phillips RS: **Protective CD4+ T-cell lines raised against *Plasmodium chabaudi* show characteristics of either Th1 or Th2 cells.** *Parasit Immunol* 1993, **15:**301-310.
25. Barbet AF, Meeus PF, Belanger M, Bowie MV, Yi J, Lundgren AM, Alleman AR, Wong SJ, Chu FK, Munderloh UG, Jauron SD: **Expression of multiple outer membrane protein sequence variants from a single genomic locus of *Anaplasma phagocytophilum*.** *Infect Immun* 2003, **71:**1706-1718.
26. Carlyon JA, Akkoyunlu M, Xia L, Yago T, Wang T, Cummings RD, McEver RP, Fikrig E: **Murine neutrophils require α1,3-fucosylation but not PSGL-1 for productive infection with *Anaplasma phagocytophilum*.** *Blood* 2003, **102:**3387-3395.
27. McGarey DJ, Allred DR: **Characterization of hemagglutinating components on the *Anaplasma marginale* initial body surface and identification of possible adhesins.** *Infect Immun* 1994, **62:**4587-4593.
28. McGarey DJ, Barbet AF, Palmer GH, McGuire TC, Allred DR: **Putative adhesins of *Anaplasma marginale*: major surface polypeptides 1a and 1b.** *Infect Immun* 1994, **62:**4594-4601.
29. Park J, Choi KS, Dumler JS: **Major surface protein 2 of *Anaplasma phagocytophilum* facilitates adherence to granulocytes.** *Infect Immun* 2003, **71:**4018-4025.
30. Pusterla N, Madigan JE, Asanovich KM, Chae J-S, DeRock E, Berger-Pusterla J, Lutz H, Dumler JS: **Experimental transmission of human granulocytic ehrlichia agent derived from high and low passage cell culture in horses.** *J Clin Microbiol* 2000, **38:**1276-1278.
31. Foley JE, Leutenegger CM, Dumler JS, Pedersen NC, Madigan JE: **Evidence for modulated immune response to *Anaplasma phagocytophila* sensu lato in cats with FIV-induced immunosuppression.** *Comp Immunol Micobiol Infect Dis* 2003, **26:**103-113.

32. Martin ME, Caspersen K, Dumler JS: **Immunopathology and ehrlichial propagation are regulated by interferon-γ and interleukin-10 in a murine model of human granulocytic ehrlichiosis.** *Am J Pathol* 2001, **158:**1881-1888.

33. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22:**4673-4680.

34. Page RDM: **TREEVIEW: An application to display phylogenetic trees on personal computers.** *Comput Appl Biosci* 1996, **12:**357-358.

35. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215:**403-410.

36. Kneller DG, Cohen FE, Langridge R: **Improvements in protein secondary structure prediction by an enhanced neural network.** *J Mol Biol* 1990, **214:**171-182.