1 Proteomics informed by transcriptomics identifies novel secreted proteins in *Dermacentor*
2 *andersoni* saliva ★

3
4 Lwiindi Mudenda^{a,b}, Sebastián Aguilar Pierle^{a,b}, Joshua E. Turse^b, Glen A. Scoles^c, Samuel O.
5 Purvine^d, Carrie D. Nicora^e, Therese R. W. Clauss^e, Massaro W. Ueti^c, Wendy C. Brown^{b,*},
6 Kelly A. Brayton^{a,b,*}

Abstract

Dermacentor andersoni, known as the Rocky Mountain wood tick, is found in the western United States and transmits pathogens that cause diseases of veterinary and public health importance including Rocky Mountain spotted fever, tularemia, Colorado tick fever and bovine anaplasmosis. Tick saliva is known to modulate both innate and acquired immune responses, enabling ticks to feed for several days without detection. During feeding ticks subvert host defenses such as hemostasis and inflammation, which would otherwise result in coagulation, wound repair and rejection of the tick. Molecular characterization of the proteins and pharmacological molecules secreted in tick saliva offers an opportunity to develop tick vaccines as an alternative to the use of acaricides, as well as new anti-inflammatory drugs. We performed proteomics informed by transcriptomics to identify *D. andersoni* saliva proteins that are secreted during feeding. The transcript data generated a database of 21,797 consensus sequences, which we used to identify 677 proteins secreted in the saliva of *D. andersoni* ticks fed for 2 and 5 days, following proteomic investigations of whole saliva using mass spectrometry. Salivary gland transcript levels of unfed ticks were compared with 2 and 5 day fed ticks to identify genes upregulated early during tick feeding. We cross-referenced the proteomic data with the transcriptomic data to identify 157 proteins of interest for immunomodulation and blood feeding. Proteins of unknown function as well as known immunomodulators were identified.

Keywords: Tick, Gene expression, Transcriptomics, Proteomics, Salivary proteins, Saliva, Mass spectrometry

1. Introduction

Ticks are obligate hematophagous parasites that transmit a wide range of pathogens worldwide including protozoa, bacteria and viruses. They can feed on their hosts for days and even weeks without being detected by the host immune system (Deruaz et al., 2008; Kramer et al., 2008). Immune evasion is achieved through a wide range of proteins that are secreted in saliva (Ribeiro and Francischetti, 2003; Francischetti et al., 2009; Schwalie and Schultz, 2009). Tick salivary proteins are able to inhibit or modulate both innate and adaptive immune responses (Juncadella et al., 2007; Schuijt et al., 2008; Guo et al., 2009). Infectious organisms carried by ticks exploit these immunoevasion mechanisms to proliferate within the vertebrate host and cause disease (Ramamoorthi et al., 2005). *Dermacentor andersoni* ticks are important vectors of human and animal pathogens including *Rickettsia rickettsii* (Rocky Mountain Spotted Fever) (Niebylski et al., 1999), Colorado tick fever virus (Florio et al., 1950), *Francisella tularensis* (tularemia) (Reif et al., 2011) and *Anaplasma marginale* (bovine anaplasmosis) (Kocan et al., 1985). Control of tick infestation in livestock is typically by application of chemical acaricides, which has resulted in the selection of acaricide-resistant tick populations, contamination of meat and milk products, and is an important source of environmental pollution, that kill non-target organisms (Kunz and Kemp, 1994; Lambson et al., 2005; de la Fuente et al., 2007; Mulenga et al., 2007b).

Tick saliva contains hundreds of different proteins and other pharmacologically active molecules that confer antihemostatic, anti-inflammatory and immunomodulatory properties (Guo et al., 2009; Chmelar et al., 2011). Molecular characterization and identification of novel *D. andersoni* saliva proteins could provide targets for drug development, in addition to new and improved tick vaccine targets as an environmentally friendly alternative to the use of acaricides.

Although a few tick saliva proteins have been previously identified, the majority of proteins in saliva and their full impact on the host-tick interaction remain unknown. Moreover, some of the proteins described have only been identified in salivary gland extracts but have not been confirmed to be secreted into saliva (Steen et al., 2006). The objective of this study was to identify novel, potentially immunosuppressive proteins that are upregulated early during feeding and secreted in the saliva of *D. andersoni* ticks. As the *D. andersoni* genome has not been sequenced, we established an expressed sequence tag (EST) library with over 21,000 consensus sequences from unfed, 2 day fed and 5 day fed ticks and used a method which relies on an all-frame translation of a transcriptomic dataset to generate a database of predicted proteins (Evans et al., 2012). Evans and colleagues (Evans et al., 2012) demonstrated that de novo transcriptome assembly generates a data set that can be used to identify virtually all detectable proteins using a method known as proteomics informed by transcriptomics. We used our EST library to identify 677 proteins expressed in *D. andersoni* tick saliva by mapping peptide sequences obtained from 2 day and 5 day fed tick saliva to the EST database following mass spectrometry. This comprehensive and novel catalog of *D. andersoni* saliva proteins serves as a foundation for research on the development of novel tick vaccines, which could also diminish transmission of tick-borne pathogens.

2. Materials and methods

2.1. Tick feeding

The *D. andersoni* ticks used in this study were from the Reynold's Creek stock maintained at the United States Department of Agriculture (USDA) Agricultural Research Service tick unit in Moscow, Idaho, USA. Adult female *D. andersoni* ticks were allowed to feed

on Holstein calves to study differential protein expression during tick feeding. All experiments involving animals were approved by the University of Idaho Institutional Animal Care and Use and Biosafety Committees (Protocol Numbers, IACUC: 2013-66, Biosafety: B-010-13) in accordance with institutional guidelines based on the U.S. National Institutes of Health (NIH) Guide for the Care and Use of Laboratory Animals.

2.2. Tick salivary gland EST database

Unfed ticks as well as ticks fed for 2 or 5 days were dissected under a microscope, using a pair of forceps and a sharp, sterile razor blade as previously described by Mulenga (Mulenga et al., 2007a). Salivary glands were teased out, washed in Hank's balanced salt solution and immersed in 5 - 10 volumes of *RNAlater* (Applied Biosystems, USA). Samples were left in *RNAlater* at 4°C overnight to allow thorough penetration of the tissue by *RNAlater* solution and stored at -80°C after removing the supernatant, as per the manufacturer's instructions.

Total RNA was extracted from 40 pairs of salivary glands for each cohort of ticks using TRI Reagent solution (Applied Biosystems). Then, mRNA was isolated from total RNA using the Ambion Poly (A) Purist kit in accordance with the manufacturer's instructions. A cDNA library was prepared in nine steps; fragmentation of RNA, double stranded cDNA synthesis, fragment end repair, AMPure bead preparation, adaptor ligation, small fragment removal, library quantitation, cDNA library quality assessment and preparation of working aliquots according to the Genome Sequencer FLX manual. The cDNA library was then sequenced using the 454 pyrosequencer (Roche, Germany) with GS FLX Titanium series reagents. Sequences were assembled from all cohorts to generate a database using the CLC Genomics workbench 6.5 de novo assembly tool (Gnerre et al., 2011). The algorithm was applied using the following

parameters: word size of 20 nucleotides (nt), bubble size of 50 nt and the minimum contig length was 200 bp. Annotation was aided by BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). The cut-off criteria for protein prediction were set as: E-value $\leq e^{-20}$, identity $\geq 45\%$, and $\geq 60\%$ coverage to the corresponding best hit (Zeng and Extavour, 2012). Sequences have been deposited in the GenBank Sequence Read Archive with accession number **SRP041899**.

2.3. Comparative transcriptional analysis and K-means clustering

Reads from each sample were mapped against the EST database using CLC Genomics Workbench 6.5. (CLC Bio). Gene expression levels were determined by counting the number of reads that mapped back to each contig. Transcripts were quantified in reads per kilobase of exon model per million mapped reads (RPKM) as described by Mortazavi (Mortazavi et al., 2008) and assigned *P* values using Kal's test. Although there is no genome on which to map the sequences, the error rate of the assembly algorithm used in this study is estimated to be low, thus we consider that the success rate is high enough for appropriate abundance measurements using RPKM where $RPKM = \text{total exon reads/mapped reads (millions)} * \text{exon length (kb)}$ (Misner et al., 2013).

Transcription values were compared and analyzed following adjustment of their distribution by quantile normalization (Pierlé et al., 2013). Fold change was calculated by comparison of normalized transcript expression (RPKM) values at the different feeding time points. We used K-means clustering to assign the mean transcriptional value of each gene to the cluster whose center is nearest. Lloyd's algorithm was used for these experiments (Lloyd, 1982).

Euclidean distance was used as a distance metric; five partitions were used to generate the clusters. For each gene, the mean gene expression value over all input samples was subtracted. Normalized expression values were used for clustering.

2.4. Quantitative PCR (qPCR)

In order to validate the RNAseq data we performed quantitative PCR for a set of genes. Total RNA was extracted from 40 pairs of pooled salivary glands from unfed, 2 day and 5 day fed ticks, respectively, using TRI Reagent solution (Applied Biosystems), then treated with the TURBO DNA-free kit (Dnase treatment) and subsequently used for first strand cDNA synthesis with a Superscript VILO cDNA Synthesis Kit (Invitrogen, USA). The primers used for qPCRs are listed in Table 1. The qPCR conditions were 50 °C for 2 min, 95 °C for 10 min, 44 cycles of 95 °C for 15 s, 60 °C for 1 min and 72 °C for 7 min, and a final hold at 4 °C. All reactions were done in triplicate using 0.4 ul of 10 µM primers per well and 0.5 ng of cDNA per well. The qPCR data were analyzed using the Relative Expression Software Tool (REST) and BootstRatio to assess the statistical significance of the gene expression ratios across different techniques (Pfaffl et al., 2002; Clèries et al., 2012). The data were normalized using DaSG8252, a ribosomal protein S1 RNA binding domain-containing protein, as a reference gene.

2.5. Tick saliva sample preparation

Saliva was collected from 2 day fed and 5 day fed ticks after stimulating salivation by injecting the ticks with 5 - 10 µl of 10 mg/ml dopamine in 1.2% NaCl solution, using a 50 µl syringe and a 12.7 x 0.33 mm needle. Saliva was collected using a 10 µl capillary tube fitted over

the ticks' mouth parts, and kept on ice. All of the saliva collected was pooled into a microfuge tube and stored at -80 °C. Saliva was collected from 200 5 day fed ticks and 400 2 day fed ticks.

The protein concentration of the saliva was determined using a Pierce bicinchoninic acid (BCA) protein assay (Pierce, Rockford, IL, USA). Urea (all chemicals were purchased through Sigma Aldrich, St. Louis, MO, USA unless otherwise stated) was then added to the sample so that the final urea concentration in the sample was 8 M DTT was then added to a final concentration of 10 mM. After incubating the sample in a thermo mixer at 37 °C for 1 h with constant shaking, iodoacetamide was added to reach a concentration of 40 mM. The sample was incubated again as described above. Following incubation, the sample was diluted with 50 mM ammonium bicarbonate, pH 8.0, to bring the concentration of urea in the sample to 1.5 M, and adjusted to 1 mM CaCl₂. Sample digestion was performed by adding trypsin at a ratio of 1:50 (trypsin to protein) and incubating at 37 °C for 3 h. The sample was desalted by passing the sample through a 1 ml C18 SPE column (Supelco, Bellefonte, PA, USA). Briefly, the columns were conditioned with methanol and rinsed with acidified water (0.1% trifluoroacetic acid (TFA)). The sample solution was put through the column at a rate no faster than 1 ml/min. The column was washed with 95:5 H₂O:Acetonitrile (ACN), 0.1% TFA and allowed to dry. Samples were eluted in 1 ml of 80:20 ACN:H₂O, 0.1% TFA. The sample was concentrated to a volume of approximately 50 - 100 µl in a speed vac (Thermo Savant, Holbrook, NY, USA). Protein concentration was determined by performing the BCA protein assay and the sample was stored at -80°C until needed for two-dimensional (2D) LC-MS/MS analysis.

2.6. 2D-LC-MS/MS analysis

Tick saliva was separated into 15 fractions by 2D-LC. 2D-LC was performed by applying the acidified saliva peptide mixture to a strong cation exchange (SCX) column (first dimension separation) followed by a binary gradient to high salt to elute the peptides. Fractions of peptides were loaded onto a reverse phase HPLC column which removed salt while performing a second separation of peptides based on their hydrophobicity. The 2D-LC system was custom built using two Agilent 1200 nanoflow pumps and one 1200 capillary pump (Agilent Technologies, Santa Clara, CA, USA), Valco valves (Valco Instruments Co., Houston, TX, USA), and a PAL auto sampler (Leap Technologies, Carrboro, NC, USA). Custom software allowed for full automation of parallel event coordination and near 100% MS duty cycle by the use of two trapping columns and two analytical columns. The columns were made by slurry packing media into fused silica (Polymicro Technologies Inc., Phoenix, AZ, USA) using a 0.5 cm sol-gel frit for media retention (Maiolica et al., 2005).

The first dimension SCX column consisted of 5 μm Polysulfoethyl A (PolyLC Inc., Columbia, MD, USA), 15 cm x 360 μm outer diameter (o.d.) x 150 μm inner diameter (i.d.). The trapping columns were 5 μm Jupiter C₁₈ (Phenomenex, Torrance, CA, USA), 4 cm x 360 μm o.d. x 150 μm i.d. Second dimension reversed-phase columns comprised 3 μm Jupiter C₁₈ (Phenomenex), 35 cm x 360 μm o.d. x 75 μm i.d. Mobile phases consisted of 0.1 mM NaH₂PO₄ (A) and 0.3 M NaH₂PO₄ (B) for the first dimension, and 0.1% formic acid in water (A) and 0.1% formic acid acetonitrile (B) for the second dimension (Table 2).

MS analysis was performed using a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, San Jose, CA, USA) outfitted with a custom-built electrospray ionization (ESI) interface. Electrospray emitters were custom made using 150 μm o.d. x 20 μm i.d. chemically

etched fused silica as described by Kelly (Kelly et al., 2006). A heated capillary temperature of 325 °C and a spray voltage of 2.4 kV were used. Data were acquired for 100 min starting 50 min after sample injection and 10 min into the gradient. Orbitrap spectra (automatic gain control (AGC) 1×10^6) were collected from 400-2000 m/z at a resolution of 60 k followed by data-dependent ion trap collision-induced dissociation (CID) MS/MS (collision energy 35%, AGC 3×10^4) of the 10 most abundant ions. A dynamic exclusion time of 60 s was used to discriminate against previously analyzed ions.

2.7. Protein identification

RNAseq data was used to inform protein identification. Proteins were identified by mapping peptide masses and/or sequences to our *D. andersoni* EST database following all-frame translation into amino acid sequences and using SEQUEST (Thermo Finnigan, USA). SEQUEST filtering criteria included a mass tolerance of $\pm 1-3$, ΔCn 0.1, Xcorr 1.5 for charge states of +1 and +2 and an Xcorr of 2.5 for a charge state of +3. Peptide assignments to MS/MS spectra were validated by the generating function (MS-GF) cut off $\leq 1E-9$ and peptide prophet score of ≥ 0.5 . MS-GF values are a peptide spectral probability that assigns a statistical significance to each individual spectral identification (Kim et al., 2008). An MS-GF cut off $\leq 1E-9$ gives a false discovery rate (FDR) of $\leq 1\%$ (Kim et al., 2008). Peptide prophet computes a probability of being correct for each peptide assignment to a spectrum and was used to validate SEQUEST search results (Keller et al., 2002; Ma et al., 2012). Identified peptides were searched against a reversed sequence decoy database to determine FDR values at given cut-offs.

3. Results

3.1. Development of an EST database

mRNA was extracted from salivary glands from unfed, 2 day fed and 5 day fed ticks and sequenced using the Roche 454 sequencing technology. The number of sequence reads obtained for each of the feeding time points (day 0, day 2 and day 5) were 156,620, 157,647 and 318,000, respectively. The data from all three time points were pooled and assembled, resulting in 21,797 unique sequences. Twenty-eight sequences were not annotated as a gene or transcript, resulting in 21,769 sequences that were further analyzed. The distribution of 21,506 sequences at the three time points sampled is illustrated in Fig. 1. The remaining 263 sequences had very low representation, such that following normalization of gene expression they had a value of zero at all three time points.

The total number of transcripts that were assigned a putative function was 2,810 with an additional 706 being annotated as conserved hypothetical proteins (Supplementary Table S1). Genes upregulated early during feeding included those coding for metabolic proteins, genes involved in endocytosis or nutrient uptake, reproduction and cytoskeletal network proteins, transporter proteins, splicing and transcription factors, signaling molecules, and proteins required for successful feeding and immunomodulation including genes coding for immunosuppressant protein p36, serine/threonine protein kinases, cement protein, suppressor of cytokine signaling, lipocalins, cystatin and serpin 8.

3.2. Differential transcription

We were interested in identifying transcripts that were upregulated early during feeding and still present at day 5, but absent or low at day zero, as well as those transcripts upregulated

early during feeding (at 2 days) that were low or absent at 5 days but whose proteins are present at both 2 and 5 days, assuming this set of transcripts would include immunomodulators. Table 3 summarizes the differential transcription of genes at the different feeding time points. Genes were classified as differentially expressed if their fold change was >2 with $P < 0.05$. The number of genes upregulated at 2 days compared with day zero was 7,132. Among these, the number of genes that were still upregulated at 5 days was 2,357.

3.3. Cluster analysis

We partitioned our comparative transcriptional analyses into five clusters (Fig. 2) in order to visualize the expression patterns of the different genes. Clusters 2 and 3 represent expression patterns that are not of interest to us in the current study: they contain genes that are high in unfed ticks and downregulated early during feeding, suggesting that they are not required for host immune suppression and successful feeding. Clusters 1 and 5, on the other hand, represent the ideal gene expression patterns for immunomodulation of host responses and successful feeding of the tick. These genes are upregulated at day 2 and their transcripts are still present at day 5. The transcription patterns observed in clusters 1 and 5 can also be seen in cluster 4, however genes in cluster 4 are regulated with a smaller magnitude. Supplementary Table S1 shows the affiliation of different genes with the clusters shown in Fig. 2.

All of the genes in cluster 1 are unknown. There is a conserved hypothetical protein (DaSG10689) in cluster 2 while the rest of the genes are unknown. Two genes in cluster 3 have been assigned putative function, DaSG5281 (ribosomal protein L17) and DaSG10972 (cement protein RIM36); the rest are unknown. Only six genes belong to cluster 5 and their functions are

not known. The rest of the 2,810 genes with assigned putative function, 706 genes coding for conserved hypothetical proteins and other genes with unknown function belong to cluster 4.

3.4. qPCR results

In order to validate the gene expression levels observed following deep sequencing, we performed qPCRs on nine genes listed in Table 1. The genes were selected to include various expression patterns across the three feeding time points. DaSG8252, a ribosomal protein S1 RNA binding domain-containing protein, was used as a reference gene to normalize the data. Based on our RNAseq data, DaSG8252 is constitutively expressed at a stable rate across the different feeding time points. In support of the use of ribosomal genes as housekeeping genes, (Koči et al., 2013) found that ribosomal proteins S4 and L13A showed a highly stable expression pattern over a 7 day feeding period in female *Ixodes scapularis* ticks. Ribosomal genes appear to be reliable reference genes in ticks as previously reported (Scharlaken et al., 2008; Nijhof et al., 2009; Mamidala et al., 2011). Following qPCR, changes in transcription were analyzed using REST. We compared transcriptional activity at day 0 versus day 2, day 0 versus day 5 and lastly, day 2 versus day 5. At each time point, six out of the nine genes tested had a similar expression pattern to our RNA sequence data and out of the six consistent genes, four had statistically significant ratios ($P < 0.05$). This result was then confirmed by BootstrapsRatio analyses, which showed that 70% of the genes had statistically significant transcriptional trends across the different techniques (Fig. 3). Sequencing depth is known to have an effect on the correlation of gene expression levels. Good congruency may be achieved with genes that have high numbers of reads mapped while genes that are less strongly transcribed may exhibit lower reproducibility (Mortazavi et al., 2008). Consequently, some transcripts may be assessed with good precision,

while others may be difficult to measure reliably using the different techniques (Łabaj et al., 2011). The set of genes that were selected for qPCR had a range of RPKM values and at some of the time points the RPKM values were normalized to zero. Interestingly, incongruent results were most often obtained when comparing time points where the RPKM value was normalized to zero in only one of the time points. Thus, considering that the set of genes selected for qPCR had low relative transcription levels at some time points, this result suggests that for less strongly transcribed genes approximately 70% will be reliably measured by the different techniques.

3.5. Identified saliva proteins

Saliva was collected from 2 day fed and 5 day fed ticks for LC-MS/MS and subsequent identification of proteins. The 30 fractions collected from LC generated a total of 19,460 MS/MS spectra that passed the MSGF filter. These spectra mapped to 2,104 unique peptide sequences. Of these peptides, 1,847 mapped to a total of 677 proteins translated from our EST database described in Section 3.1 (Supplementary Table S2). To our knowledge, this is the most comprehensive search to date for the individual proteins produced in *D. andersoni* tick saliva. The differential expression of the identified proteins showed that 140 proteins were detected only at 2 days, 165 at 5 days, and 372 proteins were detected at both time points. Cross-referencing with the transcriptomic data identified 148 proteins that are upregulated early and present at both 2 and 5 days of feeding and/or have sequence identity to previously reported immunomodulators (Table 4) for a total of 157 genes/proteins of interest. Of the 677 identified proteins, 136 (20%) have homologs in the public databases, based on the BLAST cut-off: $<E-20$, 45% sequence identity and 60% coverage length (Supplementary Table S3).

4. Discussion

A few studies have previously reported the use of transcriptomics to inform proteomics in other tick species. A recent study examined proteins from homogenized *Dermacentor reticularis* larvae and identified 287 proteins following 1D gel electrophoresis and LC-MS/MS (Villar et al., 2014). A second study compared proteins in the saliva of partially and fully engorged *Rhipicephalus (Boophilus) microplus* (Tirloni et al., 2014). This study employed 1D gel electrophoresis and LC-MS/MS, and identified 187 tick proteins and 68 host/bovine proteins. A third study used 1D gel electrophoresis and LC-MS/MS to analyze saliva from male and female *Ornithodoros moubata*, and identified a total of 193 proteins with only 10 of those proteins being found in both sexes (Diaz-Martin et al., 2014). A proteomic analysis of saliva from pilocarpine- or dopamine-induced salivation in fed female *Rhipicephalus sanguineus* ticks identified 19 tick proteins and 56 host proteins (Oliveira et al., 2013). While each of these studies had a different focus, it is interesting to note that they report categories of proteins with functional annotations and biological processes that are consistently identified in all of the studies and are also represented in our data (Supplementary Tables S2, S3). These broad categories of proteins include genes/proteins that encode metabolic processes, signaling, regulation, transporter activity, structural molecules, electron carrier activity, enzyme regulator activity, catalytic activity, reproduction, binding, etc., and lastly unknown proteins. Tirloni et al. (2014) also identified some of the immunomodulators discussed in this study including serpins, cystatin, thyropins, metalloproteases, serine carboxipeptidase and alpha-2-macroglobulin. They reported differences in the saliva of partially and fully engorged ticks fed for 17 and 20 days. Similarly, we found that some proteins were only expressed at 2 days of feeding and others at 5 days. All of these studies report a total number of identified proteins far below our total of 677. The *D.*

andersoni proteome may, therefore, be the most comprehensive Ixodidae tick saliva proteome to date. The differences in the number of identified proteins may be attributed to differences in fractionation methods (Villar et al., 2014) as well as the instrumentation used for mass spectrometry.

The host's innate immune response occurs within 4 h of tissue injury (Murphy et al., 2008) thus ticks must overcome the host defense mechanisms, which include hemostasis, inflammation and cell mediated immunity, to acquire a blood meal (Sonenshine, 1991). Previous studies have shown that an array of pharmacologically active tick saliva proteins is secreted within the first 24 h of attachment (Narisimhan et al., 2007). The main objective of this study was to identify proteins that are secreted in *D. andersoni* saliva during feeding, particularly those proteins that modulate the host immune response. Given that the *D. andersoni* genome has not been sequenced, we developed an EST library to aid in identifying the saliva proteins following mass spectrometry. Comparative RNAseq analyses were employed to identify genes that were upregulated early during tick feeding, assuming they would include genes that were important for immunomodulation. As has been previously reported in other systems, we found many genes to be differentially regulated over time, with the majority (more than 80% of the transcripts (18,253)) being of unknown function (Ribeiro et al., 2006; Mercado-Curiel et al., 2011; McNally et al., 2012). We cross-referenced the transcriptome with protein data to identify proteins for testing as immunomodulators. Almost all of the genes (650) corresponding to identified proteins were found in cluster 4, with 12 genes in cluster 1, and four in cluster 5 (Fig. 2). The number of genes corresponding to protein data in clusters 2 and 3 were one and nine, respectively. The 157 proteins of interest were those that were upregulated early during feeding (at day 2) and those were found in clusters 1, 4 and 5.

In order to identify proteins that are secreted in saliva and potentially required for successful feeding, including those involved in immunomodulation, we chose two feeding time points within the slow feeding phase (days 2 and 5) for collection of saliva from *D. andersoni* ticks. Day 2 was the earliest time point at which we could collect a sufficient amount of tick saliva for mass spectrometry due to technical challenges with this system (Chen et al., 2012). Day 5 was chosen because the ticks would have established their feeding site by that time and thus they would be secreting proteins necessary for maintenance of the feeding site. After day 5, ticks feed rapidly to repletion as can be inferred from the rate of increase in the live weight of female *D. andersoni* during feeding (Bergman et al., 2000).

An additional criterion for identifying the proteins of interest was based on bioinformatics to include proteins with sequence identity to known immunomodulators and to exclude proteins with sequence identity to non-immunosuppressors or other irrelevant proteins. Among the functionally annotated proteins that may be involved in suppression of the host immune response to ticks were catalase, serpin 8 precursor, alpha-2-macroglobulin, tumor rejection antigen (gp96), serine carboxypeptidase and thioredoxin peroxidase, which were present in the 2 day and 5 day fed saliva, and were upregulated at 2 days. The remaining proteins, tumor rejection antigen (gp96), metalloprotease, immunosuppressant protein p36, Iospins 8 and 9, thyroperin precursor, serpin-1 precursor and intracellular cystatin, were upregulated at day 5 and were present in tick saliva either at both 2 and 5 days of feeding or at a single feeding time point (day 2 or day 5).

Protein p36 was first reported as an immunosuppressant protein from the salivary glands of *D. andersoni* ticks (Bergman et al., 2000; Ribeiro et al., 2011) as it was shown to suppress concanavalin A-stimulated splenocyte proliferation (Alarcon-Chaidez et al., 2003). Homologues

390 of immunosuppressant protein p36 have been reported for *Amblyomma variegatum* (Nene et al.,
391 2002), *Haemaphysalis longicornis* (Konnai et al., 2009) and *Rhipicephalus appendiculatus*
392 (Nene et al., 2004). Cystatins are cysteinyl protease inhibitors (Ribeiro et al., 2011). A salivary
393 cystatin that targets cathepsin L, sialostatin L, was reported to have anti-inflammatory action and
394 to reduce proliferation of cytotoxic leukocytes (Kotsyfakis et al., 2006). Serpins, on the other
395 hand, are serine protease inhibitors that have been shown to inhibit platelet aggregation and
396 inflammation (Chmelar et al., 2011; Mulenga et al., 2013). Lospins are lone star tick serpins that
397 have been reported for *Amblyomma americanum* (Mulenga et al., 2007a). Thyropins, although
398 structurally different from cystatins, are also cysteine protease inhibitors (Lenarčič and Bevec,
399 1998) and are classified as thyroglobulin type-1 domain inhibitors (Lenarčič and Bevec, 1998).
400 Catalase and thioredoxin peroxidase quench reactive oxygen species. In addition, thioredoxin
401 peroxidase converts macrophages to an alternatively activated phenotype that produces high
402 levels of the regulatory cytokine IL-10 and low levels of the pro-inflammatory cytokine IL-12
403 (Donnelly et al., 2005). Although alpha-2-macroglobulin induces T cell proliferation, and
404 activation and proliferation of macrophages in vertebrates, “tick saliva alpha-2-macroglobulin
405 may be linked to interference in inflammation and immunomodulation” (Tirloni et al., 2014).
406 Similarly, tumor rejection antigen gp96 reportedly induces T cell proliferation and secretion of
407 IFN- γ at low doses and promotes a Th2-like immunoregulatory phenotype with an increased
408 IL10:IFN- γ ratio at higher doses in rats (Mirza et al., 2006). Several enzymes present in the
409 saliva of blood-feeding arthropods including apyrases, esterases, glucosidases, phospholipase
410 A2, metalloproteases, kininases and serine proteases are said to assist blood feeding
411 (Francischetti et al., 2009). Thus some of the predicted enzymes on our protein list such as
412 salivary gland metalloprotease, serine threonine kinase and hydrolases may be involved in the

modulation of host defense mechanisms during tick feeding. The majority (80%) of the identified proteins are of unknown function. Of the 677 identified proteins, 136 (20%) have homologs in the public databases. Our list of saliva proteins will be useful in future studies to aid in the development of anti-inflammatory drug targets as well as new tick control measures.

In summary, using proteomics informed by transcriptomics, we have generated the most comprehensive set of proteins detected to date in *D. andersoni* saliva. To identify proteins involved in immunomodulation, we analyzed our data set to identify genes/proteins that are upregulated early and secreted in saliva. These candidates include both known and unknown proteins and provide a list of genes/proteins for further study of immunosuppression. The identified saliva proteins may be exploited as targets for development of anti-inflammatory drugs and tick vaccines.

Acknowledgements

The global proteomic study was performed using EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at the Pacific Northwest National Laboratory, Richland, WA, USA. Lwiindi Mudenda was supported by a grant from the Fulbright Program, USA and by the Paul G. Allen School for Global Animal Health, Washington State University, USA, and Sebastián Aguilar Pierlé was supported by Conacyt, Mexico and Poncin, USA. This work was supported in part by U.S. Department of Agriculture grant ARS 5348-32000-033-00D. We appreciate the excellent technical support of Xiaoya Cheng, James Allison, Ralph Horn, Sara Davis and Kathy Mason.

References

- Alarcon-Chaidez, F.J., Muller-Doblies, U.U., Wikel, S., 2003. Characterization of a recombinant immunomodulatory protein from the salivary glands of *Dermacentor andersoni*. *Parasite Immunol.* 25, 69 - 77.
- Bergman, D.K., Palmer, M.J., Caimano, M.J., Radolf, J.D., Wikel, S.K., 2000. Isolation and Molecular Cloning of a Secreted Immunosuppressant Protein from *Dermacentor andersoni* Salivary Gland. *J. Parasitol.* 86, 516 - 525.
- Chen, G., Severo, M.S., Sohail, M., Sakhon, O.S., Wikel, S.K., Kotsyfakis, M., Pedra, J.H.F., 2012. *Ixodes scapularis* saliva mitigates inflammatory cytokine secretion during *Anaplasma phagocytophilum* stimulation of immune cells. *Parasit. Vectors* 5, 229.
- Chmelar, J., Oliveira, C.J., Rezacova, P., Francischetti, I.M.B., Kovarova, Z., Pejler, G., Kopacek, P., Ribeiro, J.M.C., Mares, M., Kopecky, J., Kotsyfakis, M., 2011. A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* 117, 736 - 744.
- Clèries, R., Galvez, J., Espino, M., Ribes, J., Nunes, V., de Heredia, M.L., 2012. BootstRatio: A web-based statistical analysis of fold-change in qPCR and RT-qPCR data using resampling methods. *Comput. Biol. Med.* 42, 438 - 445.
- de la Fuente, J., Almazan, C., Canales, M., de la Lastra, J.M.P., Kocan, K.M., Willadsen, P., 2007. A ten-year review of commercial vaccine performance for control of tick infestations on cattle. *Anim. Health Res. Rev.* 8, 23 - 28.
- Deruaz, M., Frauenschuh, A., Alessandri, A.L., Dias, J.M., Coelho, F.M., Russo, R.C., Ferreira, B.R., Graham, G.J., Shaw, J.P., Wells, T.N.C., Teixeira, M.M., Power, C.A., Proudfoot, A.E.I., 2008. Ticks produce highly selective chemokine binding proteins with anti-inflammatory activity. *J. Exp. Med.* 205, 2019 - 2031.
- Diaz-Martin, J., Diaz-Lopez, A., Moreno-Bueno, G., Castilla, M.A., Rosa-Rosa, J.M., Cano, A., Palacios, J., 2014. A core microRNA signature associated with inducers of the epithelial-to-mesenchymal transition. *J. Pathol.* 232, 319 - 329.
- Donnelly, S., O'Neill, S.M., Sekiya, M., Mulcahy, G., Dalton, J.P., 2005. Thioredoxin Peroxidase Secreted by *Fasciola hepatica* Induces the Alternative Activation of Macrophages. *Infect. Immun.* 71, 166 - 173.
- Evans, V.C., Barker, G., Heesom, K.J., Fan, J., Bessant, C., Matthews, D.A., 2012. De novo derivation of proteomes from transcriptomes for transcript and protein identification. *Nat. Methods* 9, 1207 - 1211.
- Florio, L., Miller, M.S., Mugrage, E.R., 1950. Colorado tick fever: Isolation of the virus from *Dermacentor andersoni* in nature and a laboratory study of the transmission of the virus in the tick. *J. Immunol.* 64, 257 - 263.
- Francischetti, I.M.B., Sá-Nunes, A., Mans, B.J., Santos, I.M., Ribeiro, J.M.C., 2009. The role of saliva in tick feeding. *Front. Biosci.* 14, 2051 - 2088.
- Gnerre, S., Maccallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., Walker, B.J., Sharpe, T., Hall, G., Shea, T.P., Sykes, S., Berlin, A.M., Aird, D., Costello, M., Daza, R., Williams, L., Nicol, R., Gnirke, A., Nusbaum, C., Lander, E.S., Jaffe, D.B., 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1513 - 1518.
- Guo, X., Booth, C.J., Paley, M.A., Wang, X., DePonte, K., Fikrig, E., Narasimhan, S., Montgomery, R.R., 2009. Inhibition of neutrophil function by two tick salivary proteins. *Infect. Immun.* 77, 2320 - 2329.
- Juncadella, I.J., Garg, R., Ananthnarayanan, S.K., Yengo, C.M., Anguita, J., 2007. T-cell signaling pathways inhibited by the tick saliva immunosuppressor, Salp15. *FEMS Immunol. Med. Microbiol.* 49, 433 - 438.

- Keller, A., Nesvizhskii, A.I., Kolker, E., Aebersold, R., 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74, 5383 - 5392.
- Kelly, R.T., Page, J.S., Luo, Q., Moore, R.J., Orton, D.J., Tang, K., Smith, R.D., 2006. Chemically Etched Open Tubular and Monolithic Emitters for Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* 78, 7796 - 7801.
- Kim, S., Gupta, N., Pevzner, P.A., 2008. Spectral Probabilities and Generating Functions of Tandem Mass Spectra: A Strike against Decoy Databases. *J. Proteome Res.* 7, 3354 - 3363.
- Kocan, K.M., Baron, S.J., Ewing, S.A., Hair, J.A., 1985. Transmission of *Anaplasma marginale* by adult *Dermacentor andersoni* during feeding on calves. *Am. J. Vet. Res.* 46, 1565 - 1567.
- Koči, J., Šimo, L., Park, Y., 2013. Validation of Internal Reference Genes for Real-Time Quantitative Polymerase Chain Reaction Studies in the Tick, *Ixodes scapularis* (Acari: Ixodidae). *J. Med. Entomol.* 50, 70 - 84.
- Konnai, S., Nakajima, C., Imamura, S., Yamada, S., Nishikado, H., Kodama, M., Onuma, M., Ohashi, K., 2009. Suppression of cell proliferation and cytokine expression by HL-p36, a tick salivary gland-derived protein of *Haemaphysalis longicornis*. *Immunology* 126, 209 - 219.
- Kotsyfakis, M., Sá-Nunes, A., Francischetti, I.M.B., Mather, T.N., Andersen, J.F., Ribeiro, J.M.C., 2006. Antiinflammatory and immunosuppressive activity of Sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. *J. Biol. Chem.* 281, 26298 - 26307.
- Kramer, C., Nahmias, Z., Norman, D.D., Mulvihill, T.A., Coons, L.B., Cole, J.A., 2008. *Dermacentor variabilis*: regulation of fibroblast migration by tick salivary gland extract and saliva. *Exp. Parasitol.* 119, 391 - 397.
- Kunz, S.E., Kemp, D.H., 1994. Insecticides and acaricides: resistance and environmental impact. *Rev. Sci. Tech. Off. Int. Epiz.* 13, 1249 - 1286.
- Łabaj, P.P., Leparć, G.G., Linggi, B.E., Markillie, L.M., Wiley, S.H., Kreil, D.P., 2011. Characterization and improvement of RNA-Seq precision in quantitative transcript expression profiling. *Bioinformatics* 27, i383 - i391.
- Lambson, B., Nene, V., Obura, M., Shah, T., Pandit, P., Ole-MoiYoi, O., Delroux, K., Welburn, S., Skilton, R., de Villiers, E., Bishop, R., 2005. Identification of candidate sialome components expressed in ixodid tick salivary glands using secretion signal complementation in mammalian cells. *Insect Mol. Biol.* 14, 403 - 414.
- Lenarčič, B., Bevec, T., 1998. Thyropins – New structurally related proteinase inhibitors. *Biol. Chem.* 379, 105 - 111.
- Lloyd, S.P., 1982. Least Squares Quantization in PCM. *IEEE Trans. Inform. Theory* 28, 129 - 137.
- Ma, K., Vitek, O., Nesvizhskii, A.I., 2012. A statistical model-building perspective to identification of MS/MS spectra with PeptideProphet. *BMC Bioinform.* 13, S1.
- Maiolica, A., Borsotti, D., Rappsilber, J., 2005. Self-made frits for nanoscale columns in proteomics. *Proteom.* 5, 3847 - 3850.
- Mamidala, P., Rajarapu, S.P., Jones, S.C., Mittapalli, O., 2011. Identification and validation of reference genes for quantitative real-time polymerase chain reaction in *Cimex lectularius*. *J. Med. Entomol.* 48, 947 - 951.
- McNally, K.L., Mitzel, D.N., Anderson, J.M., Ribeiro, J.M.C., Valenzuela, J.G., Myers, T.G., Godinez, A., Wolfenbarger, J.B., Best, S.M., Bloom, M.E., 2012. Differential salivary gland transcript expression profile in *Ixodes scapularis* nymphs upon feeding or flavivirus infection. *Ticks Tick Borne Dis.* 3, 18 - 26.
- Mercado-Curiel, R.F., Palmer, G.H., Guerrero, F.D., Brayton, K.A., 2011. Temporal characterization of the organ-specific *Rhipicephalus microplus* transcriptional response to *Anaplasma marginale*. *Int. J. Parasitol.* 41, 851 - 860.

- Mirza, S., Muthana, M., Fairburn, B., Slack, L.K., Hopkinson, K., Pockley, A.G., 2006. The stress protein gp96 is not an activator of resting rat bone marrow-derived dendritic cells, but is a costimulator and activator of CD3+ T cells. *Cell Stress Chaperones* 11, 364 - 378.
- Misner, I., Bicep, C., Lopez, P., Halary, S., Baptiste, E., Lane, C.E., 2013. Sequence comparative analysis using networks: software for evaluating de novo transcript assembly from next-generation sequencing. *Mol. Biol. Evol.* 30, 1975 - 1986.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621 - 628.
- Mulenga, A., Blandon, M., Khumthong, R., 2007a. The molecular basis of the *Amblyomma americanum* tick attachment phase. *Exp. Appl. Acarol.* 41, 267 - 287.
- Mulenga, A., Khumthong, R., Blandon, M.A., 2007b. Molecular and expression analysis of a family of the *Amblyomma americanum* tick Lospins. *J. Exp. Biol.* 210, 3188 - 3198.
- Mulenga, A., Kim, T., Ibelli, A.M.G., 2013. *Amblyomma americanum* tick saliva serine protease inhibitor 6 is a cross-class inhibitor of serine proteases and papain-like cysteine proteases that delays plasma clotting and inhibits platelet aggregation. *Insect Mol. Biol.* 22, 306 - 319.
- Murphy, K., Travers, P., Walport, M., 2008. Innate Immunity. In: Lawrence, E. (Ed.), *Janeway's Immunobiology*, Garland Science, New York and London, pp. 39-108.
- Narisimhan, S., Deponte, K., Marcantonio, N., Liang, X., Royce, T.E., Nelson, K.F., Booth, C.J., Koski, B., Anderson, J.F., Kantor, F., Fikrig, E., 2007. Immunity against *Ixodes scapularis* salivary proteins expressed within 24 hours of attachment thwarts tick feeding and impairs *Borrelia* transmission. *PLoS ONE* 2, e451.
- Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J., Bishop, R., 2002. AvGI, an index of genes transcribed in the salivary glands of the ixodid tick *Amblyomma variegatum*. *Int. J. Parasitol.* 32, 1447 - 1456.
- Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylor, D., Quackenbush, J., Bishop, R., 2004. Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect Biochem. Mol. Biol.* 34, 1117 - 1128.
- Niebylski, M.L., Peacock, M.G., Schwan, T.G., 1999. Lethal effect of *Rickettsia rickettsii* on its tick vector (*Dermacentor andersoni*). *Appl. Environ. Microbiol.* 65, 773-778.
- Nijhof, A.M., Balk, J.A., Postigo, M., Jongejan, F., 2009. Selection of reference genes for quantitative RT-PCR studies in *Rhipicephalus (Boophilus) microplus* and *Rhipicephalus appendiculatus* ticks and determination of the expression profile of Bm86. *BMC Mol. Biol.* 10, 112.
- Oliveira, C.J., Anatriello, E., de Miranda-Santos, I.K., Francischetti, I.M., Sa-Nunes, A., Ferreira, B.R., Ribeiro, J.M., 2013. Proteome of *Rhipicephalus sanguineus* tick saliva induced by the secretagogues pilocarpine and dopamine. *Ticks Tick Borne Dis.* 4, 469 - 477.
- Pfaffl, M.W., Horgan, G.W., Demfle, L., 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* 30, e36.
- Pierlé, S.A., Hammac, G.K., Palmer, G.H., Brayton, K.A., 2013. Transcriptional pathways associated with the slow growth phenotype of transformed *Anaplasma marginale* BMC Genomics 14, 272.
- Ramamoorthi, N., Narasimhan, S., Pal, U., Bao, F., Yang, X.F., Fish, D., Anguita, J., Norgard, M.V., Kantor, F.S., Anderson, J.F., Koski, R.A., Fikrig, E., 2005. The Lyme disease agent exploits a tick protein to infect the mammalian host. *Nature* 436, 573 - 577.
- Reif, K.E., Palmer, G.H., Ueti, M.W., Scoles, G.A., Margolis, J.J., Monack, D.M., Noh, S.M., 2011. *Dermacentor andersoni* transmission of *Francisella tularensis* subsp. novicida reflects bacterial colonization, dissemination, and replication coordinated with tick feeding. *Infect. Immun.* 79, 4941 - 4946.

- Ribeiro, J.M.C., Francischetti, I.M.B., 2003. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annu. Rev. Entomol.* 48, 73 - 88.
- Ribeiro, J.M.C., Alarcon-Chaidez, F., Francischetti, I.M.B., Mans, B.J., Mather, T.N., Valenzuela, J.G., Wikel, S.K., 2006. An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. *Insect Biochem. Mol. Biol.* 36, 111 - 129.
- Ribeiro, J.M.C., Anderson, J.M., Manoukis, N.C., Meng, Z., Francischetti, I.M.B., 2011. A further insight into the sialome of the tropical bont tick, *Amblyomma variegatum*. *BMC Genomics* 12, 136.
- Scharlaken, B., de Graaf, D.C., Goossens, K., Brunain, M., Peelman, L.J., Jacobs, F.J., 2008. Reference gene selection for insect expression studies using quantitative real-time PCR: the head of the honeybee, *Apis mellifera*, after a bacterial challenge. *J. Insect Science* 8, 33.
- Schuijt, T.J., Hovius, J.W.R., van Burgel, N.D., Ramamoorthi, N., Fikrig, E., van Dam, A.P., 2008. The tick salivary protein Salp15 inhibits the killing of serum-sensitive *Borrelia burgdorferi* sensu lato isolates. *Infect. Immun.* 76, 2888 - 2894.
- Schwalie, P.C., Schultz, J., 2009. Positive selection in tick saliva proteins of the salp 15 family. *J. Mol. Evol.* 68, 186 - 191.
- Sonenshine, D.E., 1991. *Biology of ticks* Vol. 1, Oxford University Press. Inc, New York.
- Steen, N.A., Barker, S.C., Alewood, P.F., 2006. Proteins in the saliva of the *Ixodida* (ticks): pharmacological features and biological significance. *Toxicon* 47, 1 - 20.
- Tirloni, L., Reck, J., Terra, R.M., Martins, J.R., Mulenga, A., Sherman, N.E., Fox, J.W., Yates, J.R., 3rd, Termignoni, C., Pinto, A.F., Vaz Ida, S., Jr., 2014. Proteomic analysis of cattle tick *Rhipicephalus (Boophilus) microplus* saliva: a comparison between partially and fully engorged females. *PLoS One* 9, e94831.
- Villar, M., Popara, M., Ayllon, N., Fernandez de Mera, I.G., Mateos-Hernandez, L., Galindo, R.C., Manrique, M., Tobes, R., de la Fuente, J., 2014. A systems biology approach to the characterization of stress response in *Dermacentor reticulatus* tick unfed larvae. *PLoS One* 9, e89564.
- Zeng, V., Extavour, C.G., 2012. ASGARD: an open-access database of annotated transcriptomes for emerging model arthropod species. *Database (Oxford)* 2012, bas048.

Figure legends

Fig. 1. Venn diagram representing the number of salivary gland transcripts observed at different feeding time points in *Dermacentor andersoni*.

Fig. 2. Cluster analysis of transcript data for *Dermacentor andersoni*. K means clusters were generated using normalized transcript expression values (RPKM). The Y-axis represents normalized expression (RPKM) values whilst the X-axis represents the three different feeding time points designated as unfed (day 0), early (2-day fed) and late (5-day fed). The cluster number is shown in the top left-hand corner of each panel. Clusters 1, 4 and 5 contained transcripts with relative expression patterns of interest, whilst genes falling into clusters 2 and 3 contained transcripts that were downregulated relative to day 0 and were not of interest.

Fig. 3. Bootstrap median gene expression ratios for *Dermacentor andersoni*. The whisker box plots show the comparison between RNAseq data and quantitative PCR (qPCR) data for the 0 to day 2 comparisons. The Y-axis represents the median expression ratios while the X-axis represents each gene tested. The boxes and whiskers show the distribution of the permuted ratios. The ratios for seven out of 10 of the genes are very similar as can be observed from the tight boxes and whiskers

626 **Table 1.** Genes and primers for quantitative PCR (qPCR) used in this study.

Gene ^a	Forward Primer	Reverse Primer	Size (bp)
DaSG107	ATCAGCACAGAGCTACAGCATAGTGC	TTCATGCACACTTTGAGTGGAGGTGC	173
DaSG6997	TACCATAGCGACAGCCGTTCTGT	AACGAAGACACAGCACAGCGAAG	201
DaSG7894	ATGCCGTCTCAACAAGTGAAGCAGT	TTGGTCGTCTCATCCCTGTCTGC	199
DaSG7979	TTCGGCGTCTTCTCGACAGGA	ATCGTGCGATGGCAATAACGACTG	163
DaSG8252	TTCAGCAACAACGTCCAGGAGTG	AAGCCTTGATGATGCCCTCTTGC	142
DaSG11272	TTGAAACCATCGCCCCCATTTGC	TTGTGTCGTTGAGCGTACTCTGTGC	121
DaSG12043	AACGCCAGGAACATCGAGCCAA	TTCGTGAAGCCGTAAAGTCGAGC	157
DaSG13281	AAGAGGCTGCTGAGGAGACGGAT	TTCTGCTGTGGCGTCATCGG	224
DaSG18522	TTCAGCACACACTTCCTCCAGGT	AACTGGCATCACAGCTTCAGCAG	215
DaSG12172	AGTCTGGAAGCCACTCAGCGTC	TGCTCGGAATGCGACATGGATG	97

627 ^aReference identification number.

628 Table 2. Pump methods for the first and second dimension separations of peptides.

SCX ^a fractionation in first dimension			Reversed-phase separation in second dimension		
Fraction	% B ^b Start	% B End	Time (min)	% B	Event
1	0	0	-40	0	Trap/Wash
2	0.0	1.5	0	0	
3	1.5	3.0	2	8	
4	3.0	4.0	10	8	Start Acquisition
5	4.0	5.0	20	12	
6	5.0	6.0	75	35	
7	6.0	7.0	97	60	
8	7.0	8.0	100	85	
9	8.0	9.0	103	0	
10	9.0	10.0	110	0	End Acquisition
11	10.0	11.0			
12	11.0	14.0			
13	14.0	17.0			
14	17.0	20.0			
15	17.0	70.0			

^aSCX, strong cation exchange; ^bB, semi-continuous gradient with buffer B.

636 Table 3. Differential gene transcription in *Dermacentor andersoni*.

Feeding Length	2 versus 0 days	5 versus 0 days	2 versus 5 days	2 and 5 versus 0 days
Number of genes upregulated	7132	7976	8056	2357
Number of genes downregulated	7949	7602	8579	5081
Number of genes unchanged ^a	6688	6191	5134	731

637 ^aIncludes genes with higher calculated fold changes that that were not significant.

638

639 Table 4. *Dermacentor andersoni* saliva proteins homologous to previously reported
640 immunomodulators.

641

Reference			
No.	Protein Description (BLAST identity)	Accession No.	E-value
DaSG5780	catalase (<i>Palaemon carinicauda</i>)	<u>AGV76059.1</u>	5.00E-118
DaSG497	serpin-8 precursor (<i>Ixodes scapularis</i>)	<u>XP_002415308.1</u>	3.00E-75
DaSG5735	alpha-2-macroglobulin (<i>Ixodes scapularis</i>)	<u>XP_002405338.1</u>	1.00E-52
DaSG4595	tumor rejection antigen (gp96) (<i>Ixodes scapularis</i>)	<u>XP_002413149.1</u>	0.00E+00
DaSG10832	metalloprotease (<i>Haemaphysalis longicornis</i>)	<u>BAE00066.1</u>	6.00E-25
DaSG10926	immunosuppressant protein p36 (<i>Dermacentor andersoni</i>)	<u>AAF03683.1</u>	2.00E-34
DaSG2850	lospin 9 (<i>Amblyomma americanum</i>)	<u>ABS87361.1</u>	3.00E-53
DaSG5228	lospin 8 (<i>Amblyomma americanum</i>)	<u>ABS87360.1</u>	6.00E-91
DaSG6931	serpin-1 precursor (<i>Ixodes scapularis</i>)	<u>XP_002415886.1</u>	3.00E-113
DaSG3273	protein kinase C inhibitor (<i>Hottentotta judaicus</i>)	<u>ADY39526.1</u>	7.00E-65
DaSG7266	thyropin precursor (<i>Amblyomma variegatum</i>)	<u>DAA34697.1</u>	9.00E-72
DaSG9580	intracellular cystatin (<i>Dermacentor variabilis</i>)	<u>ACF35512.1</u>	4.00E-43
DaSG8333	serine carboxypeptidase (<i>Ixodes scapularis</i>)	<u>XP_002414193.1</u>	7.00E-35
DaSG4532	protein disulfide isomerase-1 (<i>Haemaphysalis longicornis</i>)	<u>BAF63672.1</u>	1.00E-160
DaSG5417	thioredoxin peroxidase (<i>Ixodes scapularis</i>)	<u>AAV66580.1</u>	1.00E-152

642

643

644

645

646

Figure 1

Figure 2

Normalized expression values

Figure 3 colour

Figure 3 black and white