Tick proteomics: Understanding tick physiology, evolution and control

Marina Popara
Doctoral Thesis
Tick proteomics: Understanding tick physiology, evolution and control

Trabajo presentado por
Marina Popara
para optar al grado de Doctora
por la Universidad de Castilla La Mancha

V° B° de los Directores

Fdo. José de la Fuente García  Fdo. Margarita Villar Rayo

DEPARTAMENTO: Ciencia y Tecnología Agroforestal y Genética
CENTRO: Instituto de Investigación en Recursos Cinegéticos IREC (UCLM-CSIC-JCCM), Grupo de Sanidad y Biotecnología (SaBio)
PROGRAMA DE DOCTORADO: Doctorado en Investigación Básica y Aplicada en Recursos Cinegéticos

Ciudad Real, 2013
La realización de este trabajo ha sido posible gracias al siguiente proyecto:

Marina Popara was supported by the EU Marie Curie actions FP7- PEOPLE – ITN programme: POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases; EU Grant No. 238511).
# Table of contents

**Organization of the thesis** ........................................................................................................................................ 1

**Chapter I Introduction** ........................................................................................................................................ 3

1. **Characterization of the tick-pathogen interface by quantitative proteomics** .............................................. 3

   - Abstract .......................................................................................................................................................... 4
   - 1. Introduction ........................................................................................................................................... 4
   - 2. Experimental approaches in quantitative proteomics. ........................................................................ 5
   - 3. Application of two-dimensional DIGE labeling to the study of tick-pathogen interactions. ............ 7
   - 4. Application of protein one-step in gel digestion, peptide iTRAQ labeling and isoelectric focusing (IEF) fractionation to the study of tick-pathogen interactions. ........................................ 10
   - 5. Possibilities and limitations of DIGE and iTRAQ approaches to the study of tick-pathogen interactions. ........................................................................................................................................... 11
   - 6. Conclusions and future perspectives. .................................................................................................. 12
   - References ............................................................................................................................................. 13

2. **Hypothesis and objectives** .......................................................................................................................... 16

   - Hypothesis ............................................................................................................................................... 16
   - Development of the objectives ............................................................................................................. 17

**Chapter II Development and validation of methods for tick proteomics** ......................................................... 18

1. **Proteomics characterization of tick-host-pathogen interactions** .................................................................. 18

   - 1. Introduction ........................................................................................................................................ 19
   - 2. Materials............................................................................................................................................... 20
   - 2.1 Tick samples ...................................................................................................................................... 20
   - 2.2. Protein extraction buffers ............................................................................................................. 20
   - 2.3. Laemmli sample buffer ................................................................................................................ 21
   - 2.4. SDS-PAGE gel components ....................................................................................................... 21
   - 2.5. SDS-PAGE running buffer .......................................................................................................... 21
   - 2.6. Protein in gel digestion buffers .................................................................................................... 21
   - 3. Methods ............................................................................................................................................... 22
   - 3.1. General considerations .................................................................................................................. 22
   - 3.2. Tick protein extraction .................................................................................................................. 22
   - 3.3. Protein fractionation method ....................................................................................................... 24
   - 3.4. Proteome in-gel digestion ............................................................................................................. 25
   - 3.5. Reverse phase-liquid chromatography RP-LC-MS/MS analysis ............................................... 25
   - 3.6. Proteomics data analysis .............................................................................................................. 26
   - 3.7. Hemoglobin removal .................................................................................................................... 26
   - 3.8. Pathogen identification ................................................................................................................ 27
   - 3.9. De novo sequencing and homology analyses ............................................................................. 28
   - 4. Notes ................................................................................................................................................... 29
   - References ............................................................................................................................................. 30

**Chapter III Comparative proteomics approaches to the study of tick biology, development and evolution** .... 32

1. **Comparative proteomics for the characterization of the most relevant Amblyomma tick species as vectors of zoonotic pathogens worldwide** ...................................................................................... 33

   - Abstract ............................................................................................................................................... 34
   - 1. Introduction ...................................................................................................................................... 35
   - 2. Materials and methods. .................................................................................................................... 36
   - 2.1. Tick collection ............................................................................................................................... 36
   - 2.2. Protein extraction .......................................................................................................................... 37
   - 2.3. 2D-DIGE of A. americanum developmental stages ...................................................................... 37
   - 2.4. Total proteome analysis of Amblyomma spp. by protein one-step in gel digestion, LC-MS/MS and peptide identification ........................................................................................................... 39
Chapter V

Lesser protein degradation machinery correlates with higher BM86 tick vaccine efficacy in Rhipicephalus annulatus when compared to R. microplus

Abstract

1. Materials and methods
   1.1. Tick collection
   1.2. Protein extraction and proteomics analysis
   1.3. Proteomics data analysis
   1.4. Western blot analysis of BM86 and Cathepsin L

References

96
Organization of the thesis

The following work focuses on the application of proteomics tools to analyze the tick-host-pathogen relationships primarily focusing on tick proteome studies.

A general introduction on current proteomics studies conducted in the field of tick investigation is presented in Chapter I summarizing the most important proteomics tools and their relevant achievements. This chapter also describes the most relevant experimental approaches in quantitative proteomics and their application to characterize differential protein expression in ticks and cultured tick cells in response to pathogen infection.

Chapter II describes various methods for protein extraction and analysis for proteomics research, established during the course of this work to characterize tick-host-pathogen interactions. The methods particularly focus on the use of blood-fed replete ticks collected after feeding on vertebrate host.

Proteomics approaches to study tick biology, development and evolutionary relationships are described in Chapter III. First part of the study investigates proteomes of three most relevant *Amblyomma* tick species. The proteomes of *A. americanum* adults and nymphs were characterized because of their role in pathogen transmission. No differences were detected between unfed adult female and male *A. americanum* at proteomic level, however, between adults and nymphs differentially represented proteins underlined important differences between these two developmental stages. Although these ticks were unfed, over-represented host proteins were identified and may act as protein reserve to supply nutrients during off-host periods.

The study also compared the proteome of *A. americanum*, *A. cajennense* and *A. variegatum* adult ticks revealing that differentially represented tick proteins are involved in tick attachment and feeding, heat shock response, protease inhibition, blood digestion and heme detoxification therefore suggesting the existence of different adaptation processes to biotic and abiotic variables in these ticks. A novel method was applied using *de novo* sequencing of proteomics data for the analysis of the phylogenetic relationships between the three *Amblyomma* spp. in a prove of phyloproteomics concept. This result suggested that phyloproteomics might be a useful tool for the phylogenetic analysis of tick species in which sequence data is a limiting factor but limitations associated with sequence homology should be considered when selecting datasets for analysis.

In the second part of the study, a systems biology approach is applied to investigate stress response in *Dermacentor reticulatus* unfed tick larvae. Paired end sequencing and proteomics
informed by transcriptomics (PIT) were used for the analysis of transcriptomics and proteomics data, showing to be useful for analyzing organisms such as D. reticulatus where little sequence information is available. The results revealed the activation of stress response in D. reticulatus unfed larvae and identified a Rickettsia sp. similar to R. raoultii in these ticks. Stress responses likely counteract the negative effect of temperature and other stress conditions such as pathogen infection and favor tick adaptation to environmental conditions to increase tick survival. These results suggest that these mechanisms are conserved across hard tick species.

In Chapter IV, the effect of different hosts on R. microplus tick feeding and reproduction is examined by comparing the ticks fed on cattle and white tailed deer. R. microplus ticks fed on cattle show overrepresented tick proteins involved in blood digestion and reproduction when compared to ticks fed on WTD. This correlates with the observed higher tick numbers, weight and reproductive performance in ticks fed on cattle. The analysis of host proteins revealed the proteins differentially expressed between the hosts and proteins involved in host response to tick infestations.

Chapter V uses proteomics tools to further characterize protective mechanisms of tick-control vaccines. The first part of the study examines differential protein expression in Rhipicephalus (Boophilus) microplus and R. annulatus ticks fed on Bos taurus cattle vaccinated with BM86 protective antigen. The study gives a further insight into the mechanisms behind the increased susceptibility of R. annulatus to this and other tick antigens compared to R. microplus. The results reveled that tick protein degradation machinery is under represented in R. annulatus when compared to R. microplus suggesting that lesser protease activity in R. annulatus results in more efficient antibody-antigen interactions and higher vaccine efficacy. In the second part of the study, proteomics analysis of ticks subjected to the effect of two tick-control vaccines – BM86 and subolesin (SUB) showed that although both vaccines reduce tick feeding and reproduction, they act through different protective mechanisms.

Finally, conclusions including most relevant achievements of studies in this work are presented. Last chapter (Chapter VII) reviews achievements in tick-host-pathogen proteomics and discusses current difficulties and future perspectives of these studies.

As an additional material, a compact disc is included, containing supplementary tables of chapters III, IV and V and a PDF copy of this thesis.
Chapter I

Introduction

Characterization of the tick-pathogen interface by quantitative proteomics

Abstract
Ticks are vectors of pathogens that affect human and animal health worldwide. Ticks and the pathogens they transmit have co-evolved molecular interactions involving genetic traits of both the tick and the pathogen that mediate their development and survival. Proteomics and genomics studies of infected ticks are required to understand tick-pathogen interactions and identify potential vaccine antigens to control tick infestations and pathogen transmission. In this paper, the application of quantitative proteomics to characterize differential protein expression in ticks and cultured tick cells in response to pathogen infection is reviewed. Analyses using (a) two-dimensional differential in gel electrophoresis (DIGE) labeling and (b) protein one-step in gel digestion, peptide iTRAQ labeling and isoelectric focusing fractionation, both followed by peptide and protein identifications by mass spectrometry resulted in the identification of host, pathogen and tick proteins differentially expressed in response to infection. Although at its infancy, these results showed that quantitative proteomics is a powerful approach to characterize the tick-pathogen interface and demonstrated pathogen and tick-specific differences in protein expression in ticks and cultured tick cells in response to pathogen infection.

1. Introduction
Ticks are ectoparasites of animals and humans and are considered to be the most important arthropod vector of pathogens in some regions [1]. Ticks and the pathogen they transmit have co-evolved molecular interactions that affect pathogen infection, multiplication and transmission while enabling both vector and pathogen to survive [2].

The overall research goal of our group is to characterize molecular interactions at the vector-pathogen interface and to use these results to develop vaccines to control tick infestations and pathogen infection/transmission. Quantitative proteomics studies provide information on the cell protein content that may differ from results at the transcriptomics level and may be more relevant for tick vaccine antigen discovery [3]. Our working hypothesis is that tick proteins differentially expressed in response to pathogen infection would include those involved in pathogen infection, multiplication and transmission, as well as being involved in the tick protective response against infection.

Few studies have characterized tick proteome in response to pathogen infection [3-8], but only three works have used a quantitative proteomics approach [3, 4, 8]. Herein, we reviewed the application of quantitative proteomics to characterize differential protein expression in ticks and cultured tick cells in response to pathogen infection. These studies are a fundamental contribution
towards the understanding of the tick-pathogen interface and may contribute to the development of new generation pathogen transmission-blocking vaccines designed to control tick infestations and reduce pathogen transmission to vertebrate hosts.

2. Experimental approaches in quantitative proteomics.

Quantitative proteomics is intended to obtain information about differences between compared cellular states. Classically, the methodologies to quantitatively study proteomes have been divided in gel-based and gel-free approaches.

**Gel-based approaches.** Difference in gel electrophoresis (DIGE) was developed by Unlü et al. (1997) with the aim to minimize the run-to-run variability and the laborious and time-consuming process common in two-dimensional polyacrylamide gel electrophoresis 2D-PAGE. The multiplex labeling with spectrally-resolvable fluorescent dyes enables the loading of multiple protein samples in the same gel. At present, two DIGE strategies are available: (i) Minimal labeling [9], where the reactive group of dyes is a NHS ester that react with the epsilon amino group of lysine residues in proteins via an amide linkage and the experimental conditions allow labeling of only 1-3% of lysine residues of proteins, and (ii) Saturation labeling [10,11] based on dyes that have a maleimide reactive group that form a covalent bond with the thiol group of cysteine residues via a thioether linkage with a protocol that label all available cysteine residues of a protein increasing sensitivity to 0.1 ng of albumin as detection limit [10] thus allowing proteomics studies from sample quantities 10-fold lower that those required to carry out the minimal labeling approach.

**Gel-free approaches.** Although 2D-PAGE based methods have been the classical methodology in proteomics, the poor reproducibility along with the difficult automation and detection of low abundance and membrane hydrophobic proteins have led to developing a wide variety of methodologies for protein quantification without the use of gels. Most of these approaches are based on stable isotope or isobaric labeling of proteins or peptides before multidimensional chromatography coupled to mass spectrometry analysis, but actually also label-free methodologies are being applied to proteome quantification.

The labeling can be done at the cellular level (metabolic labeling) or on proteins or peptides extracted (chemical and enzymatic labelings). All these methodologies are based on the introduction of a mass difference in the samples under comparison that can be clearly detected by mass spectrometry. The determination of the relative ratio of the introduced tags enables the assessment of the differential protein expression of a sample state over another. The most common methods are described below.
Metabolic labeling. The differential incorporation of stable isotopes in the proteins involves cell growth on media containing, for example, the naturally abundant nitrogen isotopes ($^{14}$N (99.6%) and $^{15}$N (0.4%)), while a second group of cells is grown on the same medium enriched in $^{15}$N (96%) [12]. A more advanced step, stable isotope labeling by amino acids in cell culture (SILAC) developed by Ong et al. [13], labels cellular proteomes through normal metabolic processes incorporating non-radioactive, stable isotope-containing amino acids ($^{13}$C-Arg, $^{13}$C-Lys, $^{13}$C-Tyr, $^{3}$d-Leu, etc) in newly synthesized proteins. After an appropriate growing period, the cell pools are combined and, after an optional protein or peptide fractionation, analyzed by MS.

Chemical labeling. In this case, the labeling of proteomes is done by adding isotopic tags, such as the isotope-coded affinity tag (ICAT) approach that introduces a difference of 8 Da in Cys residues [14], or with multiplex isobaric reagents as in the cases of isobaric tags for relative and absolute quantitation (iTRAQ) [15] and the tandem mass tags (TMT) [16]. These isobaric reagents represent a new generation of labeling reagents particularly well suited for resolving complex dynamics because they allow multiplexed quantitative analysis from 2 to 8 different samples. TMT and iTRAQ reagents, composed by an amine-reactive NHS-ester group, a spacer arm and an MS/MS reporter, react covalently with the N-terminus of peptides and e-amine of lysine residues. For each sample, a unique low-mass reporter ion results in the MS/MS spectrum, providing the identification of the peptide by the interpretation of the fragmentation spectrum and the quantification of the relative abundance of each peptide in the samples by comparison of the intensities of the corresponding reporter ion.

Enzymatic labeling. Enzymatic labeling is performed on peptides after protein digestion by enzymatic incorporation of two $^{18}$O atoms at the C-terminus of the cleaved peptides due to the catalytic mechanism of trypsin [17]. The recent development of a general protocol [18] and an statistical model [19] for quantitative high-throughput analysis of proteomes have converted the $^{18}$O labeling in one of the most simple approaches for routine quantitative analysis of large proteomes.

Label-free quantification. The drawbacks associated to labeling approaches, mainly the time and complexity of sample preparation, requirement of large protein amounts, high costs of reagents and incomplete labeling, have increased the interest for label-free proteomics approaches with the aim to obtain faster, cleaner and simpler quantitative proteomics results [20]. Label-free strategies include approaches for relative and absolute quantification of protein samples. In the first case, the quantification is based on the comparison of peptide peak intensity with its concentration [21] or comparing the spectral count (number of identified total MS/MS spectra) of the same protein
that increase with protein concentration [22]. The absolute quantification of protein abundance in the samples under comparison can be determined by calculating the Protein Abundance Index (PAI) [23, 24] and more recently by the application of the Absolute Protein Expression (APEX) technique [25], a modified spectral counting strategy recently developed as free APEX Quantitative Proteomics Tool [26] (http://pfgr.jcvi.org/).

3. Application of two-dimensional DIGE labeling to the study of tick-pathogen interactions.

The first work to characterize molecular tick-pathogen interactions by quantitative proteomics was developed in 2007 [4]. In this study, a DIGE minimal labeling approach was applied in order to analyze the proteome of *Ixodes scapularis* IDE8 cultured cells uninfected and infected with *Anaplasma marginale*. Fifty micrograms of total proteins from uninfected and *A. marginale*-infected cells at 3 dpi (days post infection) were labeled with 400 pmol of Cy3 and 400 pmol of Cy5, respectively, according to the manufacturer’s protocol (GE Healthcare, Madrid, Spain). Combined protein samples, including an internal pool with equal amounts of proteins from each cell sample labeled with Cy2, were resolving by 2-DE using 24 cm 3–11 NL pH range IPG Strips for the first IEF dimension followed by second dimension in 12% SDS-PAGE gels. Protein spots were visualized using an Ettan DIGE Imager (GE Healthcare) and image analysis was performed with DeCyder 2 D Software, version 7.0 (GE Healthcare). This experiment showed 24 spots differentially regulated, eighteen of them with greater than 2 fold change between infected and uninfected cells, of which 3 spots were down-regulated and 15 up-regulated. The analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/TOF) allowed the identification of ten proteins, seven of which were up-regulated proteins from *A. marginale*, such as Major Surface Proteins (MSP) 2 and 4, and different transcription and translation factors involved in *Anaplasma*-tick interactions and the development of the pathogen within tick cells. Of the three differentially regulated tick proteins, only a homologue to translation elongation factor 1γ was up-regulated in infected cells whereas Glutathione S-Transferase (GST) and a putative high-mobility group-like protein were down-regulated in infected IDE8 tick cells. For GST, the transcriptomics analysis showed an increase in mRNA levels after *A. marginale* infection that could reflect the existence of a posttranscriptional mechanism induced by the pathogen to control tick stress response to infection [3,4,8]. Furthermore, the functional analysis of GST by RNA interference (RNAi) in IDE8 cells demonstrated that GST gene knockdown resulted in lower *A. marginale* infection levels, thus suggesting that while GST gene expression is activated in response to pathogen infection, it is required for *A. marginale* infection, trafficking and/or multiplication in tick cells [3,4].
The complexity of conducting proteomics studies in ticks was revealed in a work for the analysis of the differential protein expression in *Rhipicephalus* spp. ticks naturally-infected with different pathogens [8]. In this case, from 300 ticks collected in Sicily, only 16 were found infected with *Rickettsia conorii* (three *R. sanguineus*), *Ehrlichia canis* (two *R. sanguineus*), *Theileria annulata* (nine *R. bursa*) and *Anaplasma ovis* (two *R. turanicus*) [8]. This low number of infected ticks could be a common finding depending on the prevalence of tick infestations and pathogen infection [27,28,29]. In this work, ticks were processed individually with TriReagent (Sigma, St. Louis, MO, USA), a method that allows simultaneously obtaining DNA necessary for the identification of pathogen infection, proteins for quantitative proteomic analysis, and RNA for the characterization of mRNA levels of the differentially expressed proteins of interest. However, this method results in low protein yields, almost forcing the choice of DIGE saturation labelling that only need 5 micrograms of each sample for the subsequent quantitative analysis of the proteomes. The results showed differences for each infected *Rhipicephalus* tick species when compared with its respective matching uninfected control with an average of 57 spots (SD = 6) differentially expressed between control and infected samples [8]. A total of 21 spots were selected for MS identification applying three different mass spectrometry approaches: (1) Peptide Mass Fingerprinting by MALDI-TOF MS, (2) reverse phase liquid chromatography coupled with tandem mass spectrometry (RP-LC-MS/MS) and database search and finally (3) manual *de novo* interpretation of non-assigned high-quality spectra. If the first methodology failed to identify one protein spot, the second approach was attempted and, if the result was not a positive mach, the manual interpretation of spectra was done. This laborious protocol aimed to achieve the greatest number of protein identifications despite the limited amount of protein sample available. A summary of the 9 unique proteins identified after the different MS analyses is illustrated in figure 1.
Figure 1. Differentially expressed unique proteins identified in *Rhipicephalus* spp. ticks naturally infected with different pathogens. The 9 unique proteins from host, ticks and pathogens were identified after MS analysis by MALDI-TOF, RP-LC-MS/MS and de novo interpretation of good quality spectra derived from the selected spots. Data of up-regulated and down-regulated proteins are from Villar et al. (2010b). Abbreviations: *R.s./ R.c.*: *Rhipicephalus sanguineus* / *Rickettsia conorii*; *R.s./ E.c.*: *Rhipicephalus sanguineus* / *Ehrlichia canis*; *R.b./ T.a.*: *Rhipicephalus bursa* / *Theileria annulata*; *R.t./ A.o.*: *Rhipicephalus turanicus* / *Anaplasma ovis*.

Four of the identified spots corresponded to abundant plasma host proteins (albumin and hemoglobins) and proteins, such haptoglobin, related with the acute-phase stress/inflammatory host response to pathogen infection. The identification of differentially expressed tick proteins showed that infection also affects biological processes in ticks such as the remodelling of the actin cytoskeleton. In this way, the protein enolase was under-expressed in infected ticks and a recent study by Nogueira et al. [30] have demonstrated that enolase immunization of murine hosts significantly reduces the acquisition of spirochetes by feeding ticks suggesting that these protein could have a stage-specific role in *Borrelia burgdorferi* survival in feeding ticks.

In summary, despite the limitations in protein yields, host, tick and pathogen proteins were identified in this study and shown to be present in different amounts in infected and uninfected ticks. These results supported that pathogen infection affect tick protein expression and demonstrated that a proteomics approach such a saturation DIGE technology in combination with MS analysis is a powerful tool for the study of host-tick-pathogen interactions.
4. Application of protein one-step in gel digestion, peptide iTRAQ labeling and isoelectric focusing (IEF) fractionation to the study of tick-pathogen interactions.

This proteomics approach was applied in *Ixodes scapularis* ISE6 cultured cells uninfected and infected with *Anaplasma phagocytophilum*, an emerging tick-borne pathogen that causes human granulocytic anaplasmosis [3]. In this case, hundred micrograms of total protein extract from each experimental condition were concentrated in a 10% SDS-PAGE gel band stopping the run as soon as the front entered 3mm into the gel. The whole proteome from each sample were trypsin digested and the resulted peptides where labeled with iTRAQ reagents (Applied Biosystems, Madrid, Spain) following manufacturer’s protocol and combined. The peptides pool were separated by IEF on a 3100 OFFgel fractionator (Agilent, Santa Clara, CA, USA) and the recovered fractions were analyzed by RP-HPLC-LIT mass spectrometry using a Surveyor LC System coupled to a linear ion trap LTQ mass spectrometer (Thermo-Finnigan, San Jose, CA, USA). The iTRAQ approach empowers the identification of proteins by the analysis of the MS/MS fragmentation pattern and also the quantification of these proteins in the different samples by comparing the intensity of the iTRAQ reporter ions. In this case, the concentration of the whole proteome of each sample in one gel band facilitates the effective removal of detergents and others contaminants that might hinder trypsin digestion or may be difficult to eliminate from the peptide pool to avoid interferences with the subsequent MS analysis [18]. Additionally, protein concentration combined with IEF fractionation allow obtaining a similar in depth analysis that is achieved with the digestion of all bands resolved in one-dimensional-SDS-PAGE, but with less time and effort [18].

After database searches of the fragmentation MS/MS spectra, 383 tick proteins were identified with an 88% of them assigned to *Ixodes* spp. (unpublished results). The quantitative analysis of the identified proteins indicated that 86 of these ticks proteins were differentially expressed as consequence of infection, of which, 42% of them were over-expressed in infected cells affecting different biological processes (unpublished results). With respect to pathogen proteins, preliminary iTRAQ results showed unique bacterial proteins, of which 10 have been assigned to *Anaplasma* ssp., that change between early and late infections (unpublished results). This proteomics analysis also showed that Heat Shock Proteins (HSPs) are involved in the infection process but, while HSP70 was over-expressed in infected cells, other putative HSPs such as HSP20 were under-expressed after infection [3]. HSP70 participates in the disassembly of clathrin coat-mediated endocitosis [31] whereas HSP20 stabilizes actin filaments preventing its polymerization that is required for the endocytosis process [32]. Taken together, these results
could imply that *A. phagocytophilum* uses this endocytic pathway to enter into the cell, thus controlling tick response to promote infection.

In summary, these preliminary results showed that protein one-step in gel digestion, peptide iTRAQ labeling and IEF fractionation combined with MS analysis is a potent approach to the study of tick-pathogen interactions. This approach results in the quantitative analysis of tick and pathogen proteins, thus allowing a dynamic analysis of tick-pathogen interactions.

5. Possibilities and limitations of DIGE and iTRAQ approaches to the study of tick-pathogen interactions.

Both quantitative proteomics approaches described herein have shown to be powerful and complementary tools to characterize the tick-pathogen interface allowing the identification of proteins from host, ticks and pathogens in the same experiment using different ticks and tick cell cultures. However, each method has advantages and disadvantages that include methodological and equipment-based considerations that should be considered when designing experiments (Table 1).

| Table 1. Comparative 2D-DIGE and iTRAQ approaches study of tick–pathogen interaction. |
|---------------------------------|---------------------------------|
| **Advantages**                   | **Disadvantages**                |
| **2D-DIGE**                     |                                 |
| Very low amount of protein sample is necessary to carried out the analysis, especially of saturation labeling that only requires 5 µg per sample. | Not always proteins detected as differentially expressed can be identified by MS and preparative gels need to be performed. |
| The use of 2D provides a visual map of the proteomes that is useful to make a rapid comparison between samples. | Proteins with extreme pIs or molecular weights, low-abundance proteins and hydrophobic membrane proteins are poorly represented. |
| Gels can be post-DIGE stained and the identification of protein spots can be performed with different MS approaches. | Although fluorescent dye labeling reduces the number of gels necessary for the analysis with respect to 2D conventional experiments, the process is laborious and time-consuming and well-trained personnel are required to obtain the best results. |
| **iTRAQ**                       |                                 |
| Can be coupled with different methodologies to concentrate the proteome prior to labeling and with multidimensional chromatography allows obtaining a great depth of analysis and a large number of protein identifications. | The iTRAQ reporter ions can only be analyzed in some mass spectrometers such as QSTAR quadrupole-TOF MS instrument (Applied Biosystems), or in LTQ linear ion trap mass spectrometer (Thermo Fisher) with the pulsed-Q-dissociation (PQD) MS/MS operating mode. |
| Very useful for resolving complex dynamic changes in the proteome and to quantify low-abundance proteins. | Need for substantial amounts of protein sample to properly conduct the analysis. |
Other important considerations are associated with the integration of transcriptomics and proteomics data. These approaches should be viewed as complementary as in most cases they provide different results in terms of numbers, types and levels of tick mRNAs/proteins differentially expressed in response to pathogen infection [3, 4, 8]. These differences could be attributed to different factors such as (a) differences in mRNA/protein abundance due to differences in mRNA and protein half-life, (b) differences in the sensitivity of the methods employed [3, 4, 8], and (c) post-transcriptional regulation of proteins levels [3, 8]. Therefore, data from transcriptomics and proteomics should be integrated using functional genomics and systems biology approaches to identified key molecules/pathways affected by pathogen infection in ticks [4].


The results discussed in this review reinforce the importance of conducting proteomics and transcriptomics analysis together to fully characterize tick-pathogen interactions and the host response to infection. New experiments will be required to define further the role of both tick and pathogen proteins differentially expressed during pathogen life cycle in ticks. These studies should include the analysis of tick proteome during infection time course both in ticks and tick cells. Some of these differentially expressed proteins may constitute good candidates for the development of vaccines for the dual control of tick infestations and pathogen transmission.

In the post-genomics era, proteomics include promising strategies to characterize dynamic interactions that cannot be analyzed by genomic or transcriptomic approaches. In the case of tick-pathogen interactions, the application of quantitative proteomics is still in its infancy and more research is needed to achieve full potential of this technology. Other labeling methods that actually seem to be highly robust and reproducible, such as $^{18}$O labeling and label-free approaches should provide more in-depth analyses of tick proteomes. New strategies involving sample pre-fractionation to study membrane or secreted proteins separately, and the characterization of tick organ-specific proteomes in response to infection with different pathogens will provide novel and crucial data for the characterization of the complex tick-pathogen interface.

The results of quantitative proteomics studies reviewed here and those obtained at the transcriptomics level could be integrated using a systems biology approach. This holistic (instead of a reductionist) approach to the study of tick-pathogen interactions and how these interactions give rise to function and behavior are critical to advance our understanding of pathogen infection and transmission by ticks. However, to translate these results into new effective vaccines, algorithms need to be developed to allow the selection of vaccine candidates from the plethora of
differentially expressed genes/proteins that affect tick function and pathogen infection and transmission.

Acknowledgements

We thank K. M. Kocan and E. F. Blouin (Oklahoma State University, Oklahoma, USA), and A. Torina and S. Caracappa (Intituto Zooprofilattico Sperimentale della Sicilia, Palermo, Sicily, Italy) for their contribution to the experiments discussed in this review. This research was supported by the Ministerio de Ciencia e Innovación, Spain (Project BFU2008-01244/BMC). M. Villar and N. Ayllón were funded by the JAE-DOC program (CSIC-FSE) and MEC, Spain, respectively. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases) within the FP7- PEOPLE – ITN programme (EU Grant No. 238511).

References


HYPOTHESIS AND OBJECTIVES
Hypothesis and objectives

Hypothesis
Hypothesis: The application of proteomics techniques to tick research increases our understanding of tick physiology, evolution and control.

Objectives

1. To develop protocols for the characterization of tick proteome.
2. To characterize tick physiology and evolution in *Amblyomma* species.
3. To characterize stress response in *Dermacentor reticulatus* larvae.
4. To characterize the effect of the hosts, cattle and deer, on *Rhipicephalus microplus* tick proteome.
5. To characterize differences in the response of cattle ticks, *Rhipicephalus microplus* and *Rhipicephalus annulatus*, to vaccination with the tick protective antigen BM86.
6. To characterize differences in the protective mechanisms between two tick-control vaccines, BM86 and subolesin (SUB).
Development of the objectives

Objective 1 is addressed in chapter II describing the protocols developed to optimize proteomics techniques applied to tick research.

Objectives 2 and 3 are described in chapter III describing characterization of three Amblyomma tick species proteomes, comparison of A. americanum developmental stages and implications for tick phyloproteomics and a systems biology approach to characterize stress response in Dermacentor reticulatus unfed tick larvae.

Objective 4 is elaborated in Chapter IV investigating host effect on tick biology and reproduction comparing the effect of tick feeding on cattle and deer on Rhipicephalus microplus tick proteome.

Objectives 5 and 6 are elaborated in Chapter V investigating the differences in the response of cattle ticks, Rhipicephalus microplus and Rhipicephalus annulatus, to vaccination with the tick protective antigen BM86, and additionally, differences in the mechanisms of protection of BM86 and subolesin (SUB) vaccines.
Chapter II

Development and validation of methods for tick proteomics

Proteomics characterization of tick-host-pathogen interactions

Abstract

Ticks are blood-feeding arthropod ectoparasites of wild and domestic animals that transmit disease-causing pathogens to humans and animals worldwide and a good model for the characterization of tick-host-pathogen interactions. Tick-host-pathogen interactions consist of dynamic processes involving genetic traits of hosts, pathogens and ticks that mediate their development and survival. Proteomics provides information on the protein content of cells and tissues that may differ from results at the transcriptomics level and may be relevant for basic biological studies and vaccine antigen discovery. In this chapter, we describe various methods for protein extraction and analysis for proteomics research in ticks to characterize tick-host-pathogen interactions. Particularly relevant for this characterization is the use of blood-fed ticks. Therefore, we put special emphasis on working with replete ticks collected after feeding on vertebrate hosts.

1. Introduction

Ticks are blood-feeding arthropod ectoparasites of wild and domestic animals that transmit disease causing pathogens to humans and animals worldwide and a good model for the characterization of tick-host-pathogen interactions [1-3]. Tick-host-pathogen interactions consist of dynamic processes involving genetic traits of hosts, pathogens and ticks that mediate their development and survival [2-4]. In the early 1990s, a cost-effective alternative for tick control became commercially available with the development of vaccines reducing the use of acaricides and the problems associated with them such as selection of acaricide-resistant ticks and the contamination of the environment and animal products with pesticide residues [5]. However, new vaccines are needed for efficient control of vector infestations and pathogen infection and transmission [6].

In the post-genomics era, proteomics has emerged as a powerful new tool that includes strategies for the characterization of dynamic interactions that cannot be analyzed by genomics or transcriptomics approaches alone. This technique provides information on the protein content of cells and tissues that may differ from results at the transcriptomics level and may be relevant for basic biological studies and vaccine antigen discovery [2, 3, 6-9]. In the proteomics studies of blood sucking arthropods, a significant number of studies were done primarily on mosquitoes, sandflies and tsetse flies [10]. The sequence databases of tick species of agricultural and medical importance are constantly increasing which enabled the expansion of research into the field of proteomics [11]. However, the application of proteomics research on ticks is still at its
infancy [2]. The only tick genome close to completion is that of the black-legged deer tick, Ixodes scapularis [12, 13], with 47,603 proteins currently in the Uniprot database (data from June 2013). However, for other tick species of medical importance such as Rhipicephalus, Dermacentor and Amblyomma species protein information in the databases is very limited, thus making proteomics research in this area very difficult.

Few studies have covered proteomics research in ticks [2, 9-11, 14-17]. Furthermore, some of these studies were done using tick cells lines but work with ticks is more complex because in most cases proteins from the vector, vector-borne symbionts or pathogenic microorganisms and vertebrate hosts are identified [2, 7]. Therefore, proteomics techniques need to be refined to adequately address these challenges. In this chapter, we describe various methods for protein extraction and analysis for proteomics research in ticks to characterize tick-host-pathogen interactions. Particularly relevant for this characterization is the use of blood-fed ticks. Therefore, we put special emphasis on working with replete ticks collected after feeding on vertebrate hosts.

2. Materials

All reagents used for buffer preparations need to be of analytical grade. The solutions are prepared with ultrapure water and store at 4°C except for the solutions containing SDS that are store at 20°C to avoid detergent precipitation. Reagents for protein digestions and mass spectrometry analysis need to be of liquid chromatography-mass spectrometry (LC-MS) grade.

2.1 Tick samples

Ticks are collected after feeding on vertebrate hosts, including both domestic and wild animals. After repletion, ticks are collected and stored in 70% ethanol at 4°C until processed (see Note 1).

2.2. Protein extraction buffers

1. Buffer 1: 10 mM phosphate buffered saline (PBS), pH 7.4 (Sigma Chemical Company, St. Louis, MO, USA) supplemented with 1% sodium dodecyl sulfate (SDS).

2. Buffer 2: 10 mM PBS, pH 7.4 (Sigma Chemical Company, St. Louis, MO, USA) supplemented with 1% octyl phenol ethoxylate (Triton–X-100).

3. Buffer 3: 10 mM PBS, pH 7.4 (Sigma Chemical Company, St. Louis, MO, USA) supplemented with 1% polyethylene glycol sorbitan monolaurate (Tween-20).

4. Buffer 4: 7 M urea, 2 M thiourea, 4% CHAPS, 20% glycerol, 200 mM KCl and 100 mM dibasic sodium phosphate (Na₂HPO₄ x 2H₂O), pH 7.4. Weigh 4.2 g urea, 1.52 g thiourea, 0.4 g CHAPS, 0.015 g KCl, 0.36 g (Na₂HPO₄ x 2H₂O). Add 9 ml of water, mix and make up to 10 ml with water.
For preparation of Buffers 1-3, add 1 ml of 10% stock solution of chosen detergent to 9 ml of 10mM PBS pH 7.4 and mix (see Note 2). All buffers are supplemented with 1 tablet of complete mini protease inhibitor cocktail (Roche, Basel, Switzerland) per 10 ml of solution.

2.3. Laemmli sample buffer
0.125 M Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 0.004% bromophenol blue and 10% β-mercaptoethanol. Mix 0.197 g Tris-HCl, 0.4 g SDS, 2 ml glycerol, 0.004 g bromophenol blue and bring up the volume to 10 ml with water. Make aliquots of 1 ml and store at -20°C. Supplement the buffer with 5% final of β-mercaptoethanol before use.

2.4. SDS-PAGE gel components
1. Resolving gel buffer: 1.5 M Tris-HCl, pH 8.8. Weigh 181.7 g of Tris, add water to a volume of 900 ml, adjust to pH 8.8 with HCl and make up to 1 liter.
2. Stacking gel buffer: 1.5 M Tris-HCl, pH 6.5. Weigh 181.7 g of Tris, add water to a volume of 900 ml, adjust to pH 6.5 with HCl and make up to 1 liter.
3. Resolving gel: Mix 2.5 ml of resolving gel buffer, 4 ml of 30% Acrylamide/Bis-acrylamide solution (Bio-Rad, Hercules, CA, USA), 100 µl of 10% SDS, 4 ml water and mix well. Add 100 µl of 10% ammonium persulfate and 5 µl of TEMED and cast gel.
4. Stacking gel: Mix 0.84 ml of stacking gel buffer, 0.84 ml of 30% Acrylamide/Bis-acrylamide solution, 50 µl of 10% SDS, 3 ml water and mix well. Add 50 µl of 10% ammonium persulfate and 5 µl of TEMED and cast gel.
5. Make 1.5 mm-thick gels with wells of 5 mm wide for conventional one-dimensional gels and 1.2 cm wide for proteome band concentration.

2.5. SDS-PAGE running buffer
Tris Glycine SDS PAGE Buffer (10X): 0.25 M Tris-HCl, pH 8.3, 1.92 M glycine, 1% SDS (National Diagnostics, Atlanta, Georgia, USA) is diluted 10 times with water before use.

2.6. Protein in gel digestion buffers
1. Add 0.04 g of ammonium bicarbonate to 9 ml of LC-MS grade water, mix and adjust pH to 8.8. with ammonium hydroxide. Complete to 10 ml with water to obtain a 50 mM final solution.
2. 10 mM Dithiotreitol (DTT): Add 1.54 mg to 1 ml of 50 mM ammonium bicarbonate, pH 8.8.
3. 55 mM iodoacetamide: Add 10.2 mg to 1 ml of 50 mM ammonium bicarbonate, pH 8.8.
4. Sequencing Grade Trypsin (Promega, Madison, WI, USA) is dissolved in 50 mM ammonium bicarbonate, pH 8.8 to a final concentration of 60 ng/µl.

3. Methods

3.1. General considerations

Up to date, tick proteome characterization has not been widely developed and the methodologies used are based on the analysis of either unfed ticks or the specific tick tissue of interest. Herein, we focus on establishing standardized conditions to work with replete ticks after feeding on vertebrate hosts. One important consideration, especially for quantitative proteomics analysis, is to perform protein extraction as simple as possible reducing the number of steps to minimum in order to avoid protein losses during the extraction procedure. The application of detergents for protein solubilization in a sample is widely used for routine protein extraction as for the enrichment of membrane proteins that are involved in the first contact between pathogen, vector and the host. However, to date, there is no universal detergent or detergent mixtures that allow the complete solubilization of all proteins in the sample. For proteomics analysis, buffers based on the use of chaotropic reagents such as urea and thiourea, combined with detergents and salts are also widely used for protein solubilization in one single step. Therefore, the optimal method always needs to be determined empirically depending on the sample type.

3.2. Tick protein extraction

Carry out all the procedure at 4°C until the SDS buffer is used, requiring 20°C to avoid detergent precipitation.

1. Ticks are removed from ethanol storage and left for 1-2 minutes in a fume hood to evaporate excess of ethanol.
2. The cuticle is removed by dissecting ticks using a confocal microscope and adding 10 mM PBS for constant hydration of the tissues during dissection (see Note 3).
3. The tick internal organs are broken down by quick freezing in liquid nitrogen and pulverizing with a sterilized mortar and pestle.
4. Protein extraction buffer is prepared and at least 2 ml per 100 µg of tissue is added. Homogenization is done in glass homogenizer with minimum 10 strokes of a glass rod.
5. Sample is sonicated for 1 minute in an ultrasonic cooled bath followed by 10 seconds vortex and left to rest on ice one minute. These cycles are repeated 2-3 times, or until the sample becomes completely solubilized. Extracts are shaken moderately for 20
min at 4°C and then additionally sheared mechanically by passing through a syringe equipped with a 21-gauge needle to assist break down of remaining nucleic acids.

6. Samples are centrifuged for 5 minutes at 200 x g at 20°C to remove the cell debris.

7. Supernatant is collected, quantified by BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) (Table 1) and analyzed by SDS-PAGE (Fig. 1).

Table 1. Identification of proteins extracted using different extraction buffers.

<table>
<thead>
<tr>
<th>Database search</th>
<th>Buffer 1</th>
<th>Buffer 2</th>
<th>Buffer 3</th>
<th>Buffer 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ixodida</td>
<td>62,9 %</td>
<td>48,2 %</td>
<td>44,4 %</td>
<td>82,0 %</td>
</tr>
<tr>
<td>Pecora</td>
<td>72,4 %</td>
<td>44,2 %</td>
<td>40,2 %</td>
<td>72,7 %</td>
</tr>
</tbody>
</table>

The efficacy of different detergents on tick protein extraction is compared by using protein extraction buffers 1-4. We have observed that extraction with buffer 4 gives the highest yield for both tick and host proteins identification, compared to the detergents based buffers 1-3. Of the three detergents based buffers, SDS buffer 1 shows the best efficiency in proteins extraction whereas the non-ionic detergents in buffers 2 and 3, Triton-X-100 and Tween-20, show similar values for protein extraction. Proteins are expressed as percentages of the total proteins identified.

Figure 1. Representative one-dimensional SDS-PAGE gel showing protein band patterns in extracts prepared using different protein extraction buffers. 1: Proteins extracted with buffer 1. 2: Proteins extracted with buffer 2. 3: Proteins extracted with buffer 3. 4: Proteins extracted with buffer 4. MW: molecular weight markers (PageRuler Plus Prestained Protein Ladder, Scientific, San Jose, CA, USA). Twenty micrograms of soluble fractions are run on a 5 mm wide conventional 12% SDS-PAGE gel performing electrophoresis at 180 V of constant voltage. Bands are visualized by staining with Gel Code Blue Stain Reagent (Thermo Scientific, San Jose, CA, USA) by manufacturer’s protocol.
3.3. Protein fractionation method

In this method an additional step is introduced for sample processing, thus resulting in dividing the sample in two fractions: cytosolic supernatant and a crude plasma membrane enriched pellet.

1. Protein extraction is done following the same steps as in the previous method (method 3.2, steps 1 to 7).

2. After centrifugation at 200 x g, the obtained supernatant is additionally centrifuged at 12000 x g for 20 minutes obtaining two fractions: cytosolic supernatant and crude plasma membrane enriched pellet.

3. Supernatant is collected and quantified by BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) (Fig. 2A) (see Note 4).

4. The pellet containing crude plasma membrane is directly resuspended in 100100 µl of Laemli sample buffer and left on a vertical rotating shaker for 30 minutes to one hour at 20°C with vigorous shaking to enable solubilization. Sample is then centrifuged for 30 seconds on a benchtop centrifuge to remove insoluble fragments (Fig. 2B) (see Note 5).

![Figure 2. The subcellular distribution of the extracted proteins depends on the buffer used. (A) Subcellular distribution of identified tick proteins after extraction with buffer 1 and buffer 4. We observed that buffer 4, a more astringent buffer, allow the extraction of a greater number of internal organelle and membrane proteins. (B) Number of identified tick and host proteins in the different fractions obtained using the buffer 1. An average of 41% increase in the number of tick proteins and up to 72% for host proteins is detected when crude plasma membrane fraction is included in the analysis compared to the supernatant fraction only. Therefore, is necessary to process both fractions, cytoplasmatic soluble and plasma membrane pellet obtained after detergent extraction in order to characterize the entire proteome. Identification is performed using the SEQUEST algorithm of Proteome Discoverer 1.3 (Thermo Scientific) against Ixodida and Ruminantia databases for tick and host proteins identification, respectively.]
FDR < 0.05 for tick and FDR < 0.01 for host proteins identification is considered as cut-off. Abbreviations: Cyt, cytoplasmic soluble protein fraction; PM, plasma membrane protein fraction.

3.4. Proteome in-gel digestion

1. Two hundred micrograms of protein extracts to be analyzed are precipitated by adding four volumes of ice-cold acetone to one volume of sample. The mixture is vortexed, incubated at -20°C for at least 4 hours and centrifuged at 12000 x g for 15 minutes at 4°C. The supernatant is discarded and the pellet is air dried and resuspended in 100 µl of Laemli sample buffer supplemented with 5% β-mercaptoethanol.
2. Samples are applied onto 1.2-cm wide wells in a 12% SDS-PAGE gel.
3. The electrophoretic run is stopped as soon as the front enters 3 mm into the resolving gel, so that the whole proteome becomes concentrated in the stacking/resolving gel interface (see Note 6).
4. The unseparated protein bands are visualized by staining with Gel Code Blue Stain Reagent (Thermo Scientific), excised, cut into 2x2 mm cubes and digested overnight at 37°C with 60 ng/µl sequencing grade trypsin (Promega, Madison, WI, USA) at 5:1 protein:trypsin (w/w) ratio in 50 mM ammonium bicarbonate, pH 8.8 containing 10% (v/v) acetonitrile.
5. Trifluoroacetic acid is added to a final concentration of 1% to stop the digestion.
6. Samples are desalted onto OMIX Pipette tips C18 (Agilent Technologies, Santa Clara, CA, USA) following manufacturer instructions, vacuum dried and stored at -20°C until mass spectrometry analysis.

3.5. Reverse phase-liquid chromatography RP-LC-MS/MS analysis

1. Protein digests are resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS in an Easy-nLC II system coupled to an ion trap LCQ Fleet mass spectrometer (Thermo Scientific, San Jose, CA, USA).
2. The peptides are concentrated (on-line) by reverse phase chromatography using a 0.1mm × 20 mm C18 RP precolumn (Thermo Scientific), and then separated using a 0.075 mm x 100 mm C18 RP column (Thermo Scientific) operating at 0.3 µl/min.
3. Elution of peptides is done using a 180-min gradient from 5 to 35% solvent B (Solvent A: 0.1% formic acid in water, solvent B: 0.1% formic acid in acetonitrile). ESI ionization is done with Nano-bore emitters Stainless Steel ID 30 µm (Thermo Scientific) interface (see Note 7).
4. Peptides are detected in survey scans from 400 to 1600 amu (1 \( \mu \)scan), followed by three data dependent MS/MS scans (Top 3), using an isolation width of 2 in mass-to-charge ratio units, normalized collision energy of 35\%, and dynamic exclusion applied during 30 seconds periods \( \text{(see Note 8)} \).

### 3.6. Proteomics data analysis

1. Peptide identification from raw data is carried out using the SEQUEST algorithm (Proteome Discoverer 1.3, Thermo Scientific).
2. Database search is performed against Uniprot-Ixodida.fasta (40,849 entries in June 2013), Uniprot-Pecora.fasta (59,354 entries in June 2013) and Uniprot-Alphaproteobacteria.fasta (2,138,599 entries in June 2013) for tick, host and pathogen proteins identification, respectively.
3. The following constraints were used for the searches: tryptic cleavage after Arg and Lys, up to two missed cleavage sites, and tolerances of 1 Da for precursor ions and 0.8 Da for MS/MS fragment ions and the searches were performed allowing optional Met oxidation and Cys carbamidomethylation. Search against decoy database (integrated decoy approach) was done using false discovery rate (FDR) < 0.01.

### 3.7. Hemoglobin removal

In ticks collected after feeding on a host the major constrain for the successful protein identification is a large amount of the hosts, predominantly blood proteins. These proteins such as hemoglobin, serum albumin and immunoglobulins are very abundant and mask the detection of vector and pathogen proteins. In engorged tick samples, a high abundance of host hemoglobin is detected with average of 20\% of total number of identified proteins (Fig. 3A) \( \text{(see Note 9)} \). For this reason, we tested the efficacy of a method reported to remove hemoglobin from blood samples (18) with minor modifications and compared the result to a conventional precipitation method with ice-cold acetone that was used as a control.

1. Two samples of two hundred micrograms of a protein extract each are processed.
2. One sample is directly precipitated following conventional acetone precipitation method.
3. To the other sample, a mixture of methanol and chloroform is slowly added under homogenization to a final concentration of 19\% methanol and 0.6\% chloroform. This mixture is kept at 4°C and is homogenized for an additional 20 minutes to obtain hemoglobin elimination. The sample is centrifuged at 2500 x g for 10 min at 4°C and the
resulting hemoglobin containing precipitate is discarded. The resulting supernatant is precipitated with acetone precipitation (see Note 10).

4. Both resulting pellets are dissolved in Laemli buffer, concentrated on conventional SDS-PAGE gel and compared by LC-MS/MS analysis (Figs. 3B and 3C).

![Figure 3. Removal of host hemoglobin from tick protein samples.](image)

(A) Representative protein distribution in engorged tick samples. (B) Mass spectra comparing the MS/MS fragment peptides of two precipitation methods: acetone (above) and chloroform based method (below). The same profile is observed with a slight increase in the intensity of the spectra using the acetone precipitation method. (C) Presence of hemoglobins in the sample before and after treatment. Chloroform-methanol precipitation decreases the number of detected hemoglobins in a total sample where hemoglobins represented 14% of total of proteins identified compared to 23% of hemoglobins present after acetone precipitation. Additionally, the hemoglobins that remain in the sample are identified with a lower number of peptides. This method is therefore very helpful in treating the engorged tick samples but it does not completely eliminate host hemoglobins.

### 3.8. Pathogen identification

Proteomics is one of the most powerful technologies that allow simultaneous detection of proteins originating from different organisms in the same sample. Pathogen proteins present in ticks can
be successfully detected with a high level of confidence in the same protein extract where tick and host proteins are identified if the search is performed against a Database containing pathogen proteins (Fig. 4A).

### 3.9. **De novo sequencing and homology analyses**

Several software permit protein identification by generating de novo peptide sequences, which is especially useful when working with organisms such as ticks with limited sequence information available.

1. Here, the analysis is done using Peaks Studio v 6.0 software (Bioinformatics Solutions Inc.) with the same general parameters as for the routinely applied software Proteome Discoverer 1.3 (Thermo Scientific) as described in section 3.6.

2. Additionally, in Peaks a special algorithm is used to generate de novo sequences of the input spectrum and the SPIDER module is used to identify variations from sequences using a homology match query. The filtering of the scores for all identified peptides is done by assigning a -10lg P value of 30 that was established after manual analyses of the obtained peptide matches to assure quality of the identifications.

3. The application of Peaks software using the above listed parameters gives 10% to 20% average increase in the number of detected proteins by applying homology search and generating **de novo** peptide sequences (Fig. 4B). This approach enables the identification of every peptide in the dataset, whether it is in a database or not, modified or mutated.

---

**Figure 4. Identification of tick, host and pathogen proteins in a single sample with a high level of confidence.** Protein extract of engorged ticks was obtained by total protein extraction method with buffer 4. (A) Database search is performed against a database composed of Uniprot-ruminantia.fasta, Uniprot-ixodida.fasta and Uniprot-alphaproteobacteria.fasta. Data is analyzed using the SEQUEST algorithm of Proteome Discoverer 1.3 software (Thermo Scientific) applying a 1% FDR as criteria for assignments. (B) The application of **de novo** sequencing software PEAKS Studio 6.0 (Bioinformatics Solutions Inc.) significantly increase the number of identified proteins, primarily tick and pathogen proteins that are usually masked by host proteins in engorged ticks.
4. Notes

1. When working in field experiments such as epidemiology studies or vaccine trials, collected ticks are generally stored in ethanol and not deep-frozen because it is easier under field conditions and makes their shipment from one lab to another cheaper. Therefore, this work focused on optimizations on this kind of samples.

2. Triton X-100 and Tween-20 detergents are diluted to a 10% stock solution to reduce their viscosity and facilitate pipetting. SDS is also prepared in a 10% stock solution in order to work in a similar manner.

3. Removing the cuticle from a tick enables a better detection of low-abundance tick proteins. Our previous studies showed that processing the entire tick results in the detection of over 50% of abundant cuticle proteins masking other minority tick proteins that might be a key for understanding ticks metabolical processes. The removal of the cuticle of completely engorged ticks is relatively easy to perform. Care should be taken to completely remove the tissues around the mouthparts. However, if not working with freshly collected ticks, differentiating tick tissues is very difficult due to a large presence of coagulated hosts blood.

4. When using a urea-containing buffer such as buffer 4, the obtained pellet containing the protein fraction is less abundant than when using buffers 1-3 in which the solubility of the proteins is higher. However, it has to be taken into consideration that the obtained pellet protein fraction cannot be processed further since a dense and insoluble viscous pellet is formed, probably due to the accumulation of nucleic acids or coagulation of abundant host blood proteins. However, the obtained protein extracts in the soluble fraction remain stable through time and do not show signs of degradation after long-term storage. Samples of the soluble fraction, if very diluted, can be successfully concentrated using columns for protein concentration. Millipore (Billerica, MA, USA) columns give a good result and four times concentration is accomplished by centrifuging a 1 ml sample for 10 minutes at 4800 x g).

5. Due to the impossibility of direct quantification of insoluble pellet fraction, equal amount of micrograms of dried pellet is measured and the protein concentration is determined by running the same amount of the sample on a one-dimensional gel and comparing to an albumin standard. Pellet is very difficult to dissolve and needs to be left on a shaker for at least 30 minutes. It is best to store dry pellets at -20°C and directly solubilize them in Laemli buffer prior to the analysis.

6. The concentration of proteome in one single band allows quantitative comparisons between samples without using protein-labeling approaches [19].
7. The development of long gradients promotes peptides separation and increases the number of identified proteins. When an HPLC working on microliters instead of nanoliters is used, it is recommended to increase the duration of the gradient time.

8. In the case of using a mass spectrometer with scan speed and mass resolution higher than LCQ Fleet (for example, a LTQ), it is recommended to increase the numbers of dependent MS/MS scans to 15 or more, allowing the detection of low abundant proteins.

9. Even in unfed ticks, host blood proteins can persist months after feeding and molting. Vertebrate actins are found in *Ixodes ricinus* nymphs and *Rhipicephalus microplus* larvae even weeks after molting [14, 16, 17].

10. When precipitating with acetone, a special care needs to be taken to allow it to evaporate from the obtained pellet before further processing. Otherwise, the efficiency of protein detection after this method decreases significantly [20, 21].

**Acknowledgments**

We thank CMBSO proteomics facility (Centro de Biología Molecular Severo Ochoa, Madrid, Spain) for technical assistance. This research was supported by grants BFU2011-23896 and the EU FP7 ANTIGONE project number 278976. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Postgraduate training network for capacity building to control ticks and tick-borne diseases) within the FP7-PEOPLE – ITN programme (EU Grant No. 238511).

**References**


6. de la Fuente J. Vaccines for vector control: Exciting possibilities for the future. The Veterinary Journal 2012; 194: 139-140.


Chapter III

Proteomics approaches to the study of tick biology, development and evolution


Chapter IIIa

Comparative proteomics for the characterization of the most relevant *Amblyomma* tick species as vectors of zoonotic pathogens worldwide

Abstract

Arthropod obligate blood-feeder ectoparasites such as *Amblyomma americanum*, *A. cajennense* and *A. variegatum* ticks vector emerging or re-emerging zoonotic pathogens worldwide. Nevertheless, very little information is available on the genome, transcriptome and proteome of these vector ectoparasites. Herein, we characterized the proteomes of *A. americanum* adults and nymphs because of their role in pathogen transmission and compared the proteome of *A. americanum*, *A. cajennense* and *A. variegatum* adult ticks. We also applied a novel method by using de novo sequencing proteomics data for the analysis of the phylogenetic relationships between the three *Amblyomma* spp. in a prove of concept for phyloproteomics. The results showed no differences between unfed adult female and male *A. americanum*. However, host proteins and tick proteins involved in blood digestion, heme detoxification, development and innate immunity were differentially represented between adults and nymphs. Although these ticks were unfed, over-represented host proteins have been shown to be present in ticks after molting and may act as protein reserve to supply nutrients during off-host periods. Tick proteins involved in tick attachment and feeding, heat shock response, protease inhibition, blood digestion and heme detoxification were differentially represented between *Amblyomma* tick spp., suggesting adaptation processes to biotic and abiotic variables in these ticks. The results suggested that phyloproteomics might be a useful tool for the phylogenetic analysis of tick species in which sequence data is a limiting factor but limitations associated with sequence homology should be considered when selecting datasets for analysis. These results demonstrate the possibilities of proteomics studies for the characterization of relevant tick vector species and provide new relevant information to understand the physiology, development and evolution of these tick species.

Significance

This is the first report on the proteome of the most important *Amblyomma* tick species for their relevance as vectors of zoonotic pathogens worldwide. Nevertheless, very little information is available on the genome, transcriptome and proteome of these vector ectoparasites. The results reported herein provide new relevant information to understand the physiology, development and evolution of these tick species. Phyloproteomics using de novo protein sequencing was assayed as a new approach for the phylogenetic analysis of tick species in which sequence data is a limiting factor.
1. Introduction

The arthropod subphylum Chelicerata includes Acari that contains obligate blood-feeder ectoparasites such as Amblyomma spp. ticks that vector emerging or re-emerging zoonotic pathogens including those causing typhus, Human and canine ehrlichiosis, Lyme disease, Rocky Mountain Spotted Fever, African Tick-Bite Fever, Flinders Island Spotter Fever, Heartwater, theileriosis, babesiosis, Crimean-Congo Hemorrhagic Fever [1]. According to the latest revision of the tick classification and nomenclature [2], the genus Amblyomma contains 130 species with a worldwide distribution, but three of them are the most widely distributed Amblyomma spp. vectors of zoonotic pathogens. Amblyomma americanum (lone star tick) is the most abundant tick specie in the southeastern regions of North America and the main vector of agents causing of ehrlichiosis, Southern Tick-Associated Rash Illness and tularemia [3]. Amblyomma cajennense, found throughout the Americas from the southern USA to northern Argentina [4], is considered to be a tick of great public health importance in these regions because of its major role in transmitting Rickettsia rickettsii, the causal agent of Rocky Mountain spotted fever [5], high infestation rates in cattle farms and resistance to several acaricides [6]. And finally, Amblyomma variegatum, found in the tropics and subtropics, is endemic in savannas in many countries in sub-Saharan Africa as well as in southern Arabia, the Caribbean, and some islands in the Atlantic and Indian Oceans [7]. The long mouthparts of A. variegatum cause severe and painful bites, leaving large wounds on the skin, thereby producing substantial economic losses in domestic ruminants by exsanguination or by physical injury. In addition, A. variegatum is host of a number of microbial pathogens including Ehrlichia ruminantium, the agent of heartwater, and Rickettsia africæ, the agent of African tick-bite fever, which is an emerging zoonosis in rural sub-Saharan Africa and the Caribbean [8].

However, these tick species are largely understudied. The only tick genome currently close to completion is that of the black-legged deer tick, Ixodes scapularis [9, 10] and there are only a small number of proteomics studies performed on these ticks, mainly focused on specific tissues like salivary glands and midgut proteins that are determined to contain proteins holding a key role in tick physiology and pathogen transmission [11, 12]. Therefore, more comprehensive genomics, transcriptomics and proteomics data within the Acari and particularly for tick species that vector human and animal pathogens is urgently needed to aid in evolutionary studies and the identification of common pathways involved in host-vector-pathogen interactions and mechanisms for control.

In order to expand our knowledge of the Amblyomma spp., in this work we characterized the
proteomes of *A. americanum* adults and nymphs because of their role in the transmission of an increased number of disease causing pathogens. Moreover, we carried out a proteomics study of *A. americanum*, *A. cajennense* and *A. variegatum* adult ticks to deepen in the characterization of these important *Amblyomma* spp. We applied a novel method by using *de novo* sequencing proteomics data for the analysis of the phylogenetic relationships between the three *Amblyomma* spp. in a prove of concept for phyloproteomics.

2. Materials and methods

2.1. Tick collection

*Amblyomma americanum* unfed ticks were obtained from the laboratory colony maintained at the Oklahoma State University Tick Rearing Facility. Larvae and nymphs were fed on rabbits and adult ticks were fed on sheep. Off-host ticks were maintained in a 12 hr light: 12 hr dark photoperiod at 22-25 °C and 95% relative humidity. Animals were housed at the Tick Rearing Laboratory with the approval and supervision of the OSU Institutional Animal Care and Use Committee.

*Amblyomma cajennense* unfed ticks were obtained from a laboratory colony maintained at the University of Tamaulipas, Mexico. Originally, this tick strain was established at CENAPA, Cuernavaca, Mexico. Ticks were fed on cattle and collected after repletion to allow for oviposition and hatching in humidity chambers at 12 hr light: 12 hr dark photoperiod, 22-25 °C and 95% relative humidity. Ticks were maintained in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources Commission on Life Sciences, National Research Council (http://www.nal.usda.gov/awic/pubs/noawicpubs/careuse.htm). The protocol was approved by the Committee on the Ethics of Animal Experiments of FOMIX (TAMPS-2007-C13-73622) and the University of Tamaulipas (0073622).

*Amblyomma variegatum* unfed ticks were collected from cattle in three areas of Plateau State, Nigeria [13]. These ticks were brought to the Parasitology Division, National Veterinary Research Institute, Vom Plateau State, Nigeria, where they were morphologically identified using the guidelines for ticks identification [14] and independently corroborated by A. Estrada-Peña (University of Zaragoza, Spain).

All ticks were stored in 70% ethanol at -20°C until used for protein extraction.
2.2. Protein extraction

A total of sixty ticks from each adult group and one hundred and fifty *A. americanum* nymphs were divided in four different batches which were each pulverized in liquid nitrogen and homogenized with a glass homogenizer (10 strokes) in 1 ml lysis buffer (7M urea, 2M thiourea, 4% CHAPS, 25mM Tris-HCl, pH 8.0). Samples were sonicated for 1 min in an ultrasonic cooled bath followed by 10 sec of vortex. After 3 cycles of sonication-vortex, the homogenates were centrifuged at 200 x g for 5 min at 4 °C to remove cellular debris. The supernatants were collected and protein concentration was determined using the BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) using BSA as standard.

2.3. 2D-DIGE of *A. americanum* developmental stages.

Two hundred micrograms of protein from each *A. americanum* females, males and nymphs groups were purified with a 2-D Clean-Up Kit (GE Healthcare, Madrid, Spain) according to the manufacturer’s instructions and protein concentration was determined using the 2D-Quant Kit (GE Healthcare). CyDye minimal labeling was performed following the manufacturer’s protocols (GE Healthcare). Briefly, 50 µg of protein of each sample was labeled with 400 pmol of Cy3 or Cy5 fluorochromes dissolved in anhydrous DMF (Sigma, St. Louis, MO, USA) for 30 min on ice, in the dark. Reactions were quenched by adding 1 µL of 10 mM lysine followed by incubation for 10 min on ice, in the dark. For internal standardization, a pool of equal amounts of all samples (25 µg per sample) was created and labeled with Cy2 dye with the same procedure but scaling adjusting the quantities of reagents according to the amount of protein (300 µg). The twelve individual labeled samples corresponding to four biological replicates from males, females and nymphs ticks were distributed randomly across six DIGE gels with the internal standard pooled sample also present in each separation. After sample combination, an equal volume of 2x sample buffer was added (7M urea, 2M thiourea, 4% w/v CHAPS, 2% w/v DTT and 2% v/v IPG buffer, pH 3-11). The two-dimensional electrophoresis was carried out using reagents and equipment from GE Healthcare. For the first dimension, 24-cm 3-11 NL pH range IPG strips were rehydrated overnight in 450 µL of DeStreak Rehydration Solution supplemented with 0.5% IPG buffer pH 3-11 using a reswelling tray. IEF was performed at 20 °C using an Ettan IPGphor 3. Samples were applied using anodic cup loading and the isoeletrofocusing was carried out using the following conditions: 300 V for 3h, 300-1000 V for 6h, 1000-10000 V for 3h, 10000 V for 3h and 500 V for 3h. Prior second dimension, proteins present in focused IPG strips were reduced and alkylated by successive incubations in two different equilibration buffer solutions (50 mM
Tris-HCl pH 8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, 0.2% bromophenol blue, supplemented with either 0.5% w/v DTT for the first incubation or 4.5% w/v iodoacetamide for the second incubation) for 15 min each with gentle rocking. Equilibrated IPG strips were placed onto 12% homogeneous SDS-polyacrylamide gels casted in low fluorescence glass plates using an Ettan-DALT Six System. Electrophoresis was carried out at 20 °C and 0.5W/gel for 30 min followed by a second step at 15 W/gel for 4 hours.

2.3.1. Image acquisition and data analysis.

Proteins were visualized using an Ettan DIGE Imager (GE Healthcare) following the manufacturer’s instructions. Image analysis was performed with DeCyder 2D Software, version 7.0 (GE Healthcare). Eighteen images were considered for the analysis, 12 corresponded to the different samples labeled with Cy3 and Cy5 and 6 corresponded to sample pool labeled with Cy2 and acquired individually with each gel. Spot co-detection, normalization of each spot against the corresponding value of the internal pool and volume ratios calculation were carried out using Differential In-Gel Analysis (DIA) module. In the Biological Variation Analysis (BVA) module, the 18 spot maps were distributed in 4 groups, that is, standards (6 images), and the 3 different groups of samples (4 males, 4 females and 4 nymphs ticks). The most representative standard image with average quality was assigned as master. Paired comparisons were carried out between adult females and males and nymphs and the proteins with differences in their abundance patterns were compared by Student’s t-test (p=0.05) and only significant average ratios of ± 2.5-fold difference were considered for mass spectrometry analysis. All of proteins of interest were manually checked to avoid false positives.

For preparative gel, 2-D electrophoresis of 250 µg unlabeled pool proteins was carried out in the same conditions as described above for CyDye labeled samples, but in this case, after second dimension the gel was stained with Sypro Ruby (Molecular Probes, Invitrogen, Eugene, OR, USA) following the protocol recommended by the manufacturer. Gel image was matched automatically in the BVA module of DeCyder software with the DIGE images and the protein spots of interest were manually excised from the gels, dehydrated with acetonitrile and vacuum-dried in a Speed Vac.

2.3.2. Trypsin digestion and MS analysis.

After drying, spots were re-hydrated and digested overnight at 37°C with 12.5 ng/µl trypsin (Promega, Madison, WI, USA) in 50 mM ammonium bicarbonate, pH 8.8 [15]. Trifluoroacetic acid was added to a final concentration of 1% and the peptides were finally desalted onto OMIX
Pipette tips C_{18} (Agilent Technologies, Santa Clara, CA, USA), dried-down and stored at 20°C until mass spectrometry analysis. The desalted protein digests were resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS using an Easy-nLC II system coupled to an ion trap LCQ Fleet mass spectrometer (Thermo Scientific, San Jose, CA, USA). The peptides were concentrated (on-line) by reverse phase chromatography using a 0.1mm × 20 mm C18 RP precolumn (Thermo Scientific), and then separated using a 0.075mm x 100 mm C18 RP column (Thermo Scientific) operating at 0.3 μl/min. Peptides were eluted using a 40-min gradient from 5 to 35% solvent B (Solvent A: 0,1% formic acid in water, solvent B: 0,1% formic acid in acetonitrile). ESI ionization was done using a Fused-silica PicoTip Emitter ID 10μm (New Objective, Woburn, MA, USA) interface. Peptides were detected in survey scans from 400 to 1600 amu (1 μscan), followed by three data dependent MS/MS scans (Top 3), using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion applied during 30 sec periods. Protein identification was carried out using the SEQUEST algorithm (Proteome Discoverer 1.3, Thermo Scientific). The MS/MS raw files were searched against the Ixodida (40,849 entries in June 2013) and Ruminantia (66,519 entries in June 2013) Uniprot databases with the following constraints: tryptic cleavage after Arg and Lys, up to two missed cleavage sites, and tolerances of 1 Da for precursor ions and 0.8 Da for MS/MS fragment ions and the searches were performed allowing optional Met oxidation and Cys carbamidomethylation. A false discovery rate (FDR) < 0.01 was considered as condition for successful peptide assignments subsequent protein identification.

2.4. Total proteome analysis of *Amblyomma* spp. by protein one-step in gel digestion, LC-MS/MS and peptide identification.

Two hundred micrograms of protein from *A. americanum, A. cajennense* and *A. variegatum* ticks obtained following the protocol described in the section 2.2 were precipitated following the methanol/chloroform procedure [16], resuspended in 100 μl Laemmli sample buffer and applied onto 1.2-cm wide wells on a 12% SDS-PAGE gel. The electrophoretic run was stopped as soon as the front entered 3 mm into the resolving gel, so that the whole proteome became concentrated in the stacking/resolving gel interface. The unseparated protein bands were visualized by staining with GelCode Blue Stain Reagent (Thermo Scientific), excised, cut into 2x2 mm cubes and digested overnight at 37°C with 60 ng/μl sequencing grade trypsin (Promega, Madison, WI, USA) at 5:1 protein:trypsin (w/w) ratio in 50 mM ammonium bicarbonate, pH 8.8 containing 10% (v/v) acetonitrile (Shevchenko et al., 2006). The resulting tryptic peptides from each band were
extracted by 30 min-incubation in 12 mM ammonium bicarbonate, pH 8.8. Trifluoroacetic acid was added to a final concentration of 1% and the peptides were finally desalted onto OMIX Pipette tips C18 (Agilent Technologies, Santa Clara, CA, USA), dried-down and stored at -20°C until mass spectrometry analysis in the CBMSO Proteomics Facility (Madrid, Spain). The experiment was performed by duplicate with similar results.

Briefly, the desalted protein digest was resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS using an Agilent 1100 LC system (Agilent Technologies) coupled to a linear ion trap LTQ-Velos mass spectrometer (Thermo Scientific). The peptides were separated by reverse phase chromatography using a 0.18 mm x 150 mm Bio-Base C18 RP column (Thermo Scientific) at 1.8 μl/min. Peptides were eluted using a 120-min gradient from 5 to 40% solvent B in solvent A (Solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid, 80% acetonitrile in water). ESI ionization was done using a microspray metal needle kit (Thermo Scientific) interface. Peptides were detected in survey scans from 400 to 1600 amu (1 μscan), followed by fifteen data dependent MS/MS scans (Top 15), using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion applied during 30 sec periods. Database searching of MS/MS raw files was performed similar as described in 2.3.2. section, against the Ixodida-Uniprot database with the SEQUEST algorithm (Proteome Discoverer 1.3, Thermo Scientific). A false discovery rate (FDR) < 0.01 was considered as cut-off for positive peptides identification. Differential protein representation for individual proteins between different samples was determined using χ² test statistics with Bonferroni correction in the IDEG6 software (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) (p=0.05).

2.5. Phylogenetic analysis

In order to generate de novo sequences for the phylogenetic analysis, the raw MS/MS files were further processed using PEAKS Studio v 6.0 software (Bioinformatics Solutions Inc.) with the same general parameters as for the routinely applied software Proteome Discoverer 1.3, (Thermo Scientific) as previously described. Additionally, in PEAKS a special algorithm is used to generate de novo sequences of the input spectrum and the SPIDER module is used to identify variations from sequences using a homology match query. The filtering of the scores for all identified peptides is done by assigning a -10lg P value of 30 that was established after manual analyses of the obtained peptide matches to assure quality of the identifications. Sequences of 30 different peptides belonging to 25 proteins were generated de novo in all three Amblyomma spp. and used for multi locus analysis (MLA). The evolutionary history was inferred using 16S rDNA and protein sequences by the Maximum Likelihood (ML) method based on the
General Time Reversible and JTT matrix-based models, respectively [17, 18]. Reference I. scapularis sequences were used as outgroup. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) and JTT approach, respectively and then selecting the topology with superior log likelihood value. All positions containing gaps and missing data were eliminated. Stability or accuracy of inferred topologies were assessed via bootstrap analysis of 1000 replications. Evolutionary analyses were conducted in MEGA5 [19].

3. Results and discussion

3.1. 2-D DIGE analysis of *A. americanum* development stages protein profiles.

Two-dimensional DIGE minimal labeling approach in combination with MS was used to characterize differential protein expression between adult and nymphal stages of *A. americanum* ticks. Protein abundance was compared between males and females adults and nymphs. In order to generate unbiased results, a randomization of samples for labeling before running 2-D gels was performed using four replicates of each biological sample (Fig. 1A).

![Figure 1. 2-D DIGE analysis of *A. americanum* development stages.](image)

(A) Experimental design for DIGE analysis, indicating gel number, CyDye labeling and the four biological replicates per each sample. (B) Representative images of DIGE gels. Overlay image of proteins of Cy3-labeled adults (green) and Cy5-
labeled nymphs (red). (C) Representative images of DIGE gels. Pooled internal standard proteome. Proteins that were differentially represented with an average ratio of ± 2.5-fold are circled. Green and red circles indicate proteins that were over-represented in adults or nymphs, respectively.

Evaluation of protein profiles of 18 spot maps obtained were performed with the DeCyder software. An average of 630 spots (S.D.=45) were automatically detected, of which 457 spots matched with the master gel. The group to group comparisons between males, females and nymphs spot maps were done and the differences obtained were considered significant when the calculated average ratio showed a value lower than -2.5 or higher than +2.5, with Student t-test (p<0.05). With these criteria, no significant differences were found between males and females and 24 spots differed in abundance between adults and nymphs, with 9 spots (37%) more abundant in nymphs (Figs 1B and 1C). Due to similar spot patterns shown by males and females, the results and discussion were referred to adults instead of males and females separately.

The differentially expressed spots were excised from a preparative gel, trypsin-digested and analyzed by RP-LC-MS/MS as described in Materials and methods. From 24 spots, 22 proteins were identified (Table 1). Many of the identified spots corresponded to different isoforms to the same protein, which resulted in the identification of 17 unique proteins. Of these 17 proteins, 3 corresponded to host proteins (serum albumin and alpha and beta hemoglobins) and the rest were assigned to tick proteins (Table 1).

**Table 1. Differentially represented proteins in *Amblyomma americanum* adults and nymphs (identified by LC-MS/MS after DIGE analysis).**

<table>
<thead>
<tr>
<th>Spot number</th>
<th>Accession number</th>
<th>Protein ID</th>
<th>Mw / pI</th>
<th>Number of matched peptides</th>
<th>Sequence coverage (%)</th>
<th>Average -fold change</th>
<th>T-Test</th>
<th>Over-represented in:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>P02769</td>
<td>Serum albumin (Bos taurus)</td>
<td>69.2/6.2</td>
<td>8</td>
<td>13.84</td>
<td>7.09</td>
<td>1.20E-07</td>
<td>Adults</td>
</tr>
<tr>
<td>3</td>
<td>P02769</td>
<td>Serum albumin (Bos taurus)</td>
<td>69.2/6.2</td>
<td>10</td>
<td>16.97</td>
<td>10.56</td>
<td>9.10E-07</td>
<td>Adults</td>
</tr>
<tr>
<td>4</td>
<td>P02769</td>
<td>Serum albumin (Bos taurus)</td>
<td>69.2/6.2</td>
<td>7</td>
<td>12.35</td>
<td>8.23</td>
<td>1.30E-06</td>
<td>Adults</td>
</tr>
<tr>
<td>9</td>
<td>P04346</td>
<td>Hemoglobin subunit beta-A (Bos javanicus)</td>
<td>16.0/6.9</td>
<td>10</td>
<td>44.14</td>
<td>17.96</td>
<td>0.000</td>
<td>16 Adults</td>
</tr>
<tr>
<td>13</td>
<td>P02070</td>
<td>Hemoglobin subunit beta (Bos taurus)</td>
<td>15.9/7.6</td>
<td>5</td>
<td>41.38</td>
<td>12.58</td>
<td>2.90E-06</td>
<td>Adults</td>
</tr>
<tr>
<td>14</td>
<td>P01967</td>
<td>Hemoglobin subunit</td>
<td>15.0/8.4</td>
<td>4</td>
<td>24.82</td>
<td>12.09</td>
<td>4.00E-04</td>
<td>Adults</td>
</tr>
<tr>
<td>Entry</td>
<td>Accession</td>
<td>Description</td>
<td>Identities</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td>G3MRF0</td>
<td>Putative uncharacterized protein</td>
<td>(Amblyomma maculatum)</td>
<td>67.4/7.4</td>
<td>2</td>
<td>5.59</td>
<td>3.64</td>
<td>2.20E-05</td>
</tr>
<tr>
<td>5</td>
<td>Not identified</td>
<td></td>
<td></td>
<td>2.61</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>L7MIU4</td>
<td>Putative actin-binding cytoskeleton protein filamin (Rhipicephalus pulchellus)</td>
<td></td>
<td>78.6/8.3</td>
<td>4</td>
<td>6.76</td>
<td>3.19</td>
<td>0.000</td>
</tr>
<tr>
<td>7</td>
<td>Q9NB96</td>
<td>P450 CYP319A1 (Boophilus microplus)</td>
<td></td>
<td>60.9/8.7</td>
<td>6</td>
<td>16.57</td>
<td>3.46</td>
<td>1.40E-05</td>
</tr>
<tr>
<td>8</td>
<td>O99828</td>
<td>Ubiquinol-cytochrome-c reductase complex cytochrome b subunit</td>
<td>(Rhipicephalus sanguineus)</td>
<td>41.1/8.8</td>
<td>2</td>
<td>6.42</td>
<td>5.48</td>
<td>0.000</td>
</tr>
<tr>
<td>10</td>
<td>G3MPV4</td>
<td>Nucleoside diphosphate kinase (Amblyomma maculatum)</td>
<td></td>
<td>17.1/7.4</td>
<td>5</td>
<td>38.82</td>
<td>6.35</td>
<td>7.60E-05</td>
</tr>
<tr>
<td>11</td>
<td>G3MIJ7</td>
<td>Putative uncharacterized protein</td>
<td>(Amblyomma maculatum)</td>
<td>18.6/7.4</td>
<td>2</td>
<td>16.47</td>
<td>5.74</td>
<td>0.000</td>
</tr>
<tr>
<td>12</td>
<td>F0J8L0</td>
<td>Leucine aminopeptidase (Amblyomma variegatum)</td>
<td></td>
<td>25.6/8.1</td>
<td>4</td>
<td>21.43</td>
<td>2.74</td>
<td>0.000</td>
</tr>
<tr>
<td>16</td>
<td>A0MVX0</td>
<td>Heme lipoprotein (Amblyomma americanum)</td>
<td></td>
<td>176.9/6.8</td>
<td>2</td>
<td>1.39</td>
<td>2.58</td>
<td>1.70E-05</td>
</tr>
<tr>
<td>17</td>
<td>A0MVX0</td>
<td>Heme lipoprotein (Amblyomma americanum)</td>
<td></td>
<td>176.9/6.8</td>
<td>3</td>
<td>2.31</td>
<td>2.76</td>
<td>5.70E-06</td>
</tr>
<tr>
<td>18</td>
<td>B7QH37</td>
<td>Activating signal</td>
<td></td>
<td>173.6/8.3</td>
<td>3</td>
<td>4.30</td>
<td>2.80</td>
<td>5.90E-05</td>
</tr>
</tbody>
</table>

**Notes:**
- The table entries include gene IDs, descriptions, identities, and various numerical values.
- The columns represent different parameters or values associated with each entry.
- The table data seems to be related to protein sequences or similar biological molecules.
- The table appears to be from a scientific document, possibly a study or research report.
cointegrator 1 complex 06
subunit 3, helc1, putative
(Ixodes scapularis)

19 Not identified 2.55 0.000 Nymphs

20 B7QLV5 G-protein-linked
acetylcholine receptor
90.0/8.7 2 4.12 2.67 0.000 Nymphs

21 Q86RN8 Paramyosin
(Boophilus microplus)
101.9/5.7 4 5.04 2.72 0.000 Nymphs

22 G3MHD3 ATP synthase subunit
beta
59.1/5.7 8 19.17 2.91 0.000 Nymphs

23 B7QGH6 Protein Wnt
(Ixodes scapularis)
35.1/10.0 2 9.43 2.77 1.20E-05 Nymphs

24 Q4PN05 Putative salivary protein
(Ixodes scapularis)
25.0/8.0 2 12.66 6.63 7.20E-06 Nymphs

a) Spot numbers refer to the 2D gel proteins of interest that were analyzed by MS (Fig. 1).
b) Accession number and protein identity are listed according to the UniProt database for the best match.
c) Abbreviations: Mw, molecular weight (kDa); pl, isoelectric point.

The most abundant proteins over-represented in A. americanum nymphs or adults corresponded
to vertebrate host proteins (serum albumin and hemoglobin; 65 peptides), tick structural proteins
(paramyosin and actin-binding; 21 peptides) and tick hemoglobin processing proteins (heme
lipoprotein, P450, cytochrome b and leucine aminopeptidase; 20 peptides) (Table 1). Although
these ticks were unfed, over-represented host proteins have been shown to be present in ticks after
molting and may act as protein reserve to supply nutrients during off-host periods [11, 20, 21].
Interestingly, some of the proteins over-represented in both nymphs and adult ticks are involved
in hemoglobin processing and heme binding and transport. Tick carrier proteins that are able to
bind, transport, and store host-blood heme are involved in blood digestion, heme detoxification
and other functions such as innate immunity [22–24], and represent a unique evolutionary
strategy of the Chelicerata to both mitigate heme toxicity and utilize the molecule as a prosthetic
group [25]. These proteins may be also involved in metabolizing stored host proteins such as
hemoglobin.
Additionally, a G-protein-linked acetylcholine receptor gar-2A (spot 20) and a Wnt protein (spot 23) were over-represented in tick nymphs. These proteins are expressed throughout development, from embryonic to adult stage, and the increases found probably reflected the differences in the developmental process between adults and nymphs [26, 27].

3.2. Proteome profiling of *A. americanum* development stages by LC-MS/MS.

To complement the 2D-DIGE analysis, the total proteome of each of the four pools prepared from *A. americanum* females, males and nymphs were in-gel concentrated, digested with trypsin and the resulting peptides were separated by RP-LC-MS/MS. After database search, a total of 350 unique tick proteins were identified with an FDR<0.01, of which 218, 222 and 185 were present in *A. americanum* females, males and nymphs, respectively (see Supplementary Information, Table S1A). A total of 93 tick proteins were identified in all samples, but some proteins were identified in one sample only (Fig. 2A).

The most host abundant proteins identified were albumins and hemoglobins, similar to DIGE results (see Supplementary Information, Table S1B). Albumin peptides represented the 26% and 28% of the total host peptides identified in females and males, respectively, whereas peptides from hemoglobins were the 18%, 21% and 17% in females, males and nymphs, respectively, corroborating the storage and transmission of these proteins intrastadially [11, 20, 21].

Although in all samples analyzed host proteins were detected, the difference in host and tick protein identifications between *A. americanum* stages increased with aging and the number tick proteins identified represented around the 64% in adults and was a 20% lower in nymphs (Fig. 2B). The number of peptides per protein used to identify each protein was higher for host than for tick proteins, reflecting differences between protein databases (Fig. 2C) and also increased in host proteins with the development, probably reflecting the effect of previous feeding of ticks.

The statistical analysis did not show significant differences between males and females, in accordance with DIGE results, and only three proteins, Glutathione S-Transferase and two uncharacterized proteins (Uniprot accession numbers: G3MNX9, E5L874, G3MM43) were over-represented in nymphs with respect to adults. These results showed that DIGE is a powerful technique for the detection of subtle differences between similar samples that may be masked when the analysis is performed at the level of peptides. However, the LC-MS/MS approach allowed the identification of a greater number of proteins. Although most of the proteins found differentially represented by DIGE were detected by LC-MS/MS, the application of a labeling
quantitative approach before LC-MS/MS will probably detect a higher number of significant differences between samples.

![Venn diagram of tick proteins identified](image)

**Figure 2. Global proteome analysis of Amblyomma americanum ticks.** Protein extracts from *A. americanum* females, males and nymphs were in-gel concentrated, trypsin digested and the resulting peptides were analyzed by RP-LC-MS/MS. Raw data were searched against Ixodida and Ruminantia Uniprot databases as described in Materials and Methods. (A) Venn diagram of tick proteins identified. (B) Percent distribution of tick and host proteins from identified. (C) Number of peptides used for protein identification in the different samples. The number of peptides per protein on each sample was represented as Ave+S.E. and compared between groups by $\chi^2$ test (*p<0.05).

### 3.3. Proteome analysis of three Amblyomma spp. by LC-MS/MS

Because no differences were found between males and females in the *A. americanum* analysis, we decided to mix the same quantity of proteins from males and females of each *Amblyomma* species. Total proteome from *A. americanum, A. cajennense* and *A. variegatum* adult ticks were in-gel concentrated, trypsin-digested and analyzed RP-LC-MS/MS. After Ixodida-Uniprot database search, a total of 800 unique proteins were identified with an FDR<0.01, of which 337, 615 and 471 were present in *A. americanum, A. cajennense* and *A. variegatum* adult ticks respectively (see Supplementary Information, Table S2). A total of 208 proteins were identified in all samples, but some proteins were identified in one sample only (Fig. 3A).
Figure 3. Global proteome analysis of *Amblyomma* spp. ticks. Protein extracts from *A. americanum*, *A. cajennense* and *A. variegatum* ticks were in gel concentrated, trypsin digested and the resulting peptides were analyzed by RP-LC-MS/MS. Raw data were searched against the Ixodida-Uniprot database as described in Materials and Methods. (A) Venn diagram of tick proteins identified. (B) Number of peptides used for protein identification in the different samples. The number of peptides per protein on each sample was represented as Ave+S.E. and compared between groups by $\chi^2$ test (p>0.05). (C) Percent distribution of proteins identified in *Amblyomma* spp., *Ixodes* spp. and other tick spp.

In all *Amblyomma* spp. the mean number of peptides used to identify each protein with FDR < 0.01 was three or more peptides, indicating a good quality of the data obtained (Fig. 3B). Of the proteins identified, around 50% were proteins assigned to *Amblyomma* spp. (52%, 47% and 49% for *A. americanum*, *A. cajennense* and *A. variegatum*, respectively) whereas the rest of proteins were mainly assigned to *Ixodes* spp. (39%, 42% and 41% for *A. americanum*, *A. cajennense* and *A. variegatum*, respectively) (Fig. 3C). These results are in accordance with the tick proteins present in databases because the only tick genome close to completion is that of the black-legged deer tick, *I. scapularis* [9], with 23,047 proteins currently in the Uniprot database (data from June 2013) whereas protein database included only 437, 32 and 3138 sequences for *A. americanum*, *A. cajennense* and *A. variegatum*, respectively.

Differences were observed between *Amblyomma* spp. for some of the identified proteins (Table 2). For example, the content in cement proteins was greater in *A. variegatum* with respect to *A.
cajennense and A. americanum (1.7 and 13.5 fold, respectively) and also in A. cajennense with respect to A. americanum (8 fold) (Fig. 4A). These results showed a correlation between the abundance of cement proteins and the weight of replete females (approximately 530, 639 and 2950 mg for A. americanum [28], A. cajennese [29] and A. variegatum [30], respectively) (Fig. 4B), suggesting a mechanism to prepare and maintain the attachment to the host. Also, the heat shock response was more represented in A. cajennense and A. variegatum when compared to A. americanum (Fig. 4C). Moreover, the hemelipoproteins involved in blood digestion and heme detoxification were increased in 3 and 2.2 fold in A. americanum with respect to A. cajennense and A. variegatum, respectively (Fig. 4D), and also A. americanum showed higher levels of protease inhibitors than A. cajennense and A. variegatum (Fig. 4E) that may represent an evolutionary advantage to aid in protection against pathogen infection [24, 31–34].

Figure 4. Proteins affecting relevant biological processes in Amblyomma spp. ticks. (A) Number of peptides for tick Cement proteins involved in attachment to the host identified in Amblyomma spp. (B) Correlation between tick weight (mg) and the total number of peptides used to identify cement proteins in Amblyomma spp. Correlation coefficient (R^2) is shown. (C) Number of peptides for Heat shock proteins identified in Amblyomma spp. (D) Number of peptides for tick Hemelipoproteins involved in blood digestion identified in Amblyomma spp. (E) Number of peptides for Protease inhibitor proteins identified in Amblyomma spp. The number of peptides per protein on each pathway were represented as Ave+S.D. and compared between groups by χ^2 test (*p<0.05).
<table>
<thead>
<tr>
<th>Accession number</th>
<th>Protein ID</th>
<th>$A_{americanum}$ vs. $A_{cajennense}$</th>
<th>$A_{americanum}$ vs. $A_{variegatum}$</th>
<th>$A_{cajennese}$ vs. $A_{variegatum}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0MVX0</td>
<td>Heme lipoprotein ($Amblyomma americanum$)</td>
<td>Up</td>
<td>Up</td>
<td>=</td>
</tr>
<tr>
<td>A7U132</td>
<td>Lospin 17 ($Amblyomma americanum$)</td>
<td>Up</td>
<td>Up</td>
<td>=</td>
</tr>
<tr>
<td>B5M727</td>
<td>Alpha-2-macroglobulin ($Amblyomma americanum$)</td>
<td>Up</td>
<td>=</td>
<td>=</td>
</tr>
<tr>
<td>B7P1Q2</td>
<td>Myosin heavy chain, skeletal muscle or cardiac muscle ($Ixodes scapularis$)</td>
<td>Down</td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>B7P5D8</td>
<td>Paramyosin ($Ixodes scapularis$)</td>
<td>=</td>
<td>Up</td>
<td>=</td>
</tr>
<tr>
<td>B7PGA6</td>
<td>Kettin ($Ixodes scapularis$)</td>
<td>Down</td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>B7PMN6</td>
<td>Cuticular protein ($Ixodes scapularis$)</td>
<td>=</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>E1CAX9</td>
<td>Vitellogenin-1 ($Haemaphysalis longicorin$)</td>
<td>Up</td>
<td>Up</td>
<td>=</td>
</tr>
<tr>
<td>E2J6U6</td>
<td>Hypothetical glycine rich secreted cement protein ($Hyalomma marginatum$)</td>
<td>=</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>F0J8B9</td>
<td>Serine protease-like protein ($Amblyomma variegatum$)</td>
<td>=</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>F0J8K3</td>
<td>Salivary protein 313 ($Amblyomma variegatum$)</td>
<td>=</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>F0J8N3</td>
<td>Glycine-rich protein 44 ($Amblyomma variegatum$)</td>
<td>=</td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>F0J8R7</td>
<td>Porin/voltage-dependent anion-selective channel protein ($Amblyomma variegatum$)</td>
<td>=</td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>F0J8T5</td>
<td>Putative cement protein ($Amblyomma variegatum$)</td>
<td>=</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>F0J8W4</td>
<td>Heme lipoprotein ($Amblyomma variegatum$)</td>
<td>Up</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>F0J965</td>
<td>Succinyl-CoA synthetase small subunit ($Amblyomma variegatum$)</td>
<td>Down</td>
<td>=</td>
<td>Up</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Direction</td>
<td>Direction</td>
<td>Status</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>F0J9N1</td>
<td>Heme lipoprotein</td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma variegatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F0J9N8</td>
<td>Hemelipoglycoprotein (<em>Amblyomma variegatum</em>)</td>
<td></td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>F0JA65</td>
<td>Putative cement protein (<em>Amblyomma variegatum</em>)</td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td>F0JA84</td>
<td>Putative cement protein (<em>Amblyomma variegatum</em>)</td>
<td></td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>G3MG31</td>
<td>Putative uncharacterized protein</td>
<td>Up</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MG13</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td>Down</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MHB6</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td></td>
<td>Up</td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MHT9</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td></td>
<td>Up</td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3ML01</td>
<td>Putative uncharacterized protein</td>
<td>Up</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MLU9</td>
<td>Putative uncharacterized protein</td>
<td>Up</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MLX4</td>
<td>Putative uncharacterized protein</td>
<td>Up</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MMF0</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td></td>
<td>Up</td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MPE0</td>
<td>Catalase</td>
<td>Up</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MQA4</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MQW1</td>
<td>Putative uncharacterized protein</td>
<td>Up</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3MRB8</td>
<td>Adenosylhomocysteinase (<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td>G3MSN9</td>
<td>Putative uncharacterized protein</td>
<td>Up</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Amblyomma maculatum</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J7LVN2</td>
<td>Paramyosin</td>
<td>Up</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Haemaphysalis longicornis</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L7M0V0</td>
<td>Putative heat shock-related protein</td>
<td></td>
<td></td>
<td>Down</td>
</tr>
<tr>
<td></td>
<td>(<em>Rhipicephalus pulchellus</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L7M2B3</td>
<td>Putative chorion peroxidase-like protein</td>
<td></td>
<td>Down</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(<em>Rhipicephalus pulchellus</em>)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Up</td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>L7M6R1</td>
<td>Putative calcium-transporting atpase sarcoplasmic/endoplasmic reticulum type calcium pump isoform 1 (Rhipicephalus pulchellus)</td>
<td></td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>L7M840</td>
<td>Putative multifunctional chaperone 14-3-3 family (Rhipicephalus pulchellus)</td>
<td></td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>L7M921</td>
<td>Putative lysosomal &amp; prostatic acid phosphatase (Rhipicephalus pulchellus)</td>
<td></td>
<td>Up</td>
<td>=</td>
</tr>
<tr>
<td>L7MEG0</td>
<td>Putative heat shock protein (Rhipicephalus pulchellus)</td>
<td></td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>L7MFP6</td>
<td>Uncharacterized protein (Rhipicephalus pulchellus)</td>
<td>Up</td>
<td>=</td>
<td>=</td>
</tr>
<tr>
<td>L7MJ61</td>
<td>Putative heparan sulfate proteoglycan 2 (Rhipicephalus pulchellus)</td>
<td>=</td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>L7MM43</td>
<td>Putative 60s acidic ribosomal protein (Rhipicephalus pulchellus)</td>
<td></td>
<td>Down</td>
<td>=</td>
</tr>
<tr>
<td>L7MMF9</td>
<td>Putative glycine-rich cell wall structural protein (Rhipicephalus pulchellus)</td>
<td>=</td>
<td>=</td>
<td>Down</td>
</tr>
<tr>
<td>Q86RN8</td>
<td>Paramyosin (Boophilus microplus)</td>
<td>Up</td>
<td>Up</td>
<td>=</td>
</tr>
</tbody>
</table>

Protein levels in *Amblyomma* spp. were compared in pairs between the different species. Abbreviations: Up, Protein over-represented in the first specie with respect to the second specie under comparison. Down, Protein under-represented in the first specie with respect to the second specie under comparison. =, No difference between samples.

Differential protein representation for individual proteins between different samples was determined using $\chi^2$ test statistics with Bonferroni correction in the IDEG6 software (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) ($p=0.05$).

### 3.4. Phyloproteomics prove of concept

The *Amblyomma* tick species included in this study are among the most abundant vectors of zoonotic pathogens worldwide [1]. However, as of January 2013, only 7,454, 1,769 and 5,421 EST and nucleotide sequences were available on GenBank for *A. americanum*, *A. cajennense* and *A. variegatum*, respectively. Protein database included only 437, 32 and 3,138 sequences for *A. americanum*, *A. cajennense* and *A. variegatum*, respectively. Therefore, more genomics, transcriptomics and proteomics data is required for the characterization of these tick species. Despite the limitations of *de novo* protein sequencing, proteomics data is particularly relevant for
these studies because this information is not affected by sequencing and assembly problems associated with genomics and transcriptomics projects [35]. For example, recent evolutionary studies suggest that *Amblyomma* genus contains more species than those actually considered, with some species being polyphyletic [7, 36]. However, the limited sequence information available affects the completion of these studies. In this context, we proposed a phyloproteomics approach by using de novo sequencing of peptides for MLA.

To prove this concept, peptides sequences were generated de novo for 30 different peptides belonging to 25 proteins in *A. americanum*, *A. cajennense* and *A. variegatum* (*I. scapularis* UNIPROT accession numbers are listed in Table 3). As expected, these peptide sequences were more homologous among *Amblyomma* spp. (60% homology) than when compared to *I. scapularis* (52% homology) (Fig. 5A). Protein sequences generated de novo in all three *Amblyomma* spp. (Fig. 5B) were used for MLA in comparison with established 16S rDNA phylogeny (Fig. 5C). Reference *I. scapularis* sequences were used as outgroup. Protein sequences included in the analysis belonged to 4 different functional categories, muscle development, cytoskeleton, metabolism and other (data not shown). The topology of the trees obtained was similar when MLA was done separately for each category or grouping all proteins together both by ML and Neighbor-Joining (Fig. 5B and data not shown). However, an apparent discrepancy was observed between protein MLA and 16S rDNA analyses in the closer association between *A. variegatum* and neotropical ticks, *A. americanum* and *A. cajennense* using protein sequences (Figs. 5B and 5C). This discrepancy could reflect the fact that peptide sequences obtained in all three *Amblyomma* spp. likely belong to highly conserved sequences that may not be the most informative for phylogenetic studies. In fact, when only the most variable protein sequences were included in the analysis (45% homology among *Amblyomma* spp.), the topology of the tree was similar to the tree obtained using 16S rDNA sequences (Figs. 5C and 5D).

These results suggested that phyloproteomics might be a useful tool for the phylogenetic analysis of tick species in which sequence data is a limiting factor but limitations associated with sequence homology should be considered when selecting datasets for analysis.
Figure 5. Phyloproteomics prove of concept. (A) Protein sequences were generated de novo in *Amblyomma* spp. (30 peptides belonging to 25 different proteins are shown in alternate underlined and bold letters) and aligned with reference *I. scapularis* sequences (UNIPROT accession numbers are listed on top of peptide sequences and referenced in Table 3). Amino acids identical among *Amblyomma* spp. and *I. scapularis* are marked with asterisks (*) while amino acids identical among *Amblyomma* spp. only are marked by dots (.). (B) The evolutionary history was inferred using all protein sequences (60% homology among *Amblyomma* spp.) by ML. (C) The evolutionary history was inferred using 16S rDNA (Genbank accession numbers L34313, L34317, L34312, L43865). (D) The evolutionary history was inferred using the most variable protein sequences (45% homology among *Amblyomma* spp.) by ML. The trees with the highest log likelihood (B; -2984.3305), (C; -1202.1836) and (D; -449.2066) are shown. *I. scapularis* sequences were used as outgroup. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. There were a total of (B) 492, (C) 450 and (D) 89 positions in the final dataset. Stability or accuracy of inferred topologies were assessed via bootstrap analysis of 1000 replications.
Table 3. *I. scapularis* UNIPROT accession numbers and names of proteins from which de novo sequenced peptides were used in the phyloproteomics analysis.

<table>
<thead>
<tr>
<th>UNIPROT accession number</th>
<th>Protein Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>B7PSI1</td>
<td>Muscle myosin heavy chain</td>
</tr>
<tr>
<td>B7Q407</td>
<td>Heme lipoprotein</td>
</tr>
<tr>
<td>B7P1F4</td>
<td>Cuticle protein</td>
</tr>
<tr>
<td>B7P1Q2</td>
<td>Myosin heavy chain, skeletal muscle or cardiac muscle</td>
</tr>
<tr>
<td>B7P2K9</td>
<td>Putative uncharacterized protein</td>
</tr>
<tr>
<td>B7P453</td>
<td>Alpha tubulin</td>
</tr>
<tr>
<td>B7P6T6</td>
<td>PDZ domain protein</td>
</tr>
<tr>
<td>B7PBG2</td>
<td>Actin</td>
</tr>
<tr>
<td>B7PDK3</td>
<td>Cuticular protein</td>
</tr>
<tr>
<td>B7PFU3</td>
<td>Muscle LIM protein isoform A isoform</td>
</tr>
<tr>
<td>B7PGA6</td>
<td>Kettin</td>
</tr>
<tr>
<td>B7PGJ7</td>
<td>Secreted glycine rich cement protein</td>
</tr>
<tr>
<td>B7JP3</td>
<td>Transcription regulatory protein</td>
</tr>
<tr>
<td>B7Q4Q2</td>
<td>SN-RNP U1</td>
</tr>
<tr>
<td>B7Q5E0</td>
<td>Conserved hypothetical protein</td>
</tr>
<tr>
<td>B7Q9P9</td>
<td>Putative uncharacterized protein</td>
</tr>
<tr>
<td>B7QAZ8</td>
<td>Transcription initiation factor TFII-D component</td>
</tr>
<tr>
<td>B7QC21</td>
<td>Annexin</td>
</tr>
<tr>
<td>B7QCK2</td>
<td>ATP synthase subunit alpha</td>
</tr>
<tr>
<td>B7QDE8</td>
<td>Actin</td>
</tr>
<tr>
<td>B7QE46</td>
<td>ATP synthase subunit beta</td>
</tr>
<tr>
<td>B7QFQ2</td>
<td>Arginine kinase</td>
</tr>
<tr>
<td>B7QGJ2</td>
<td>Conserved hypothetical protein</td>
</tr>
<tr>
<td>B7QGU2</td>
<td>Coiled-coil protein</td>
</tr>
<tr>
<td>B7QI01</td>
<td>Heat shock protein 90</td>
</tr>
</tbody>
</table>
Conclusions

In summary, the results showed no differences between unfed adult female and male *A. americanum* ticks but differentially represented proteins between adults and nymphs underlined important differences between these two developmental stages. Although these ticks were unfed, over-represented host proteins may act as protein reserve to supply nutrients during off-host periods. Tick proteins involved in tick attachment and feeding, heat shock response, protease inhibition, blood digestion and heme detoxification were differentially represented between *Amblyomma* tick spp., suggesting adaptation processes to biotic and abiotic variables in these ticks. The results suggested that phyloproteomics might be a useful tool for the phylogenetic analysis of tick species in which sequence data is a limiting factor but limitations associated with sequence homology should be considered when selecting datasets for analysis. These results demonstrate the possibilities of proteomics studies for the characterization of relevant tick vector species and provide new relevant information to understand the physiology, development and evolution of these tick species.

Acknowledgements

We thank A. Marina and M. del Valle (Proteomics Facility, CBM-SO, Spain) and A. Estrada-Peña (University of Zaragoza, Spain) for technical assistance. We would like to acknowledge K.M. Kocan (Oklahoma State University, USA), C. Almazán (University of Tamaulipas, Mexico) and N.I. Ogo (National Veterinary Research Institute, Nigeria) for providing tick samples. This research was supported by grants BFU2011-23896 and the EU FP7 ANTIGONE project number 278976. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases) within the FP7-PEOPLE-ITN programme (EU Grant No. 238511).

Supplementary material.

*Table S1. Tick and Host proteins identified in A americanum developmental*

*Table S2. Proteins identified in Amblyomma spp*
References

genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony
20. Vennestrøm J, Jensen PM. Ixodes ricinus: the potential of two-dimensional gel electrophoresis as a
21. Wickramasekara S, Bunikis J, Wysocki V, Barbour AG. Identification of residual blood proteins in
ticks by mass spectrometry proteomics. Emerging Infectious Diseases 2008; 14: 1273–1275.
Adaptations against heme toxicity in blood-feeding arthropods. Insect Biochemistry and Molecular
Biology 2006; 36: 322–335.
23. Dupejova J, Sterba J, Vancova M, Grubhoffer L. Hemelipoglycoprotein from the ornate sheep tick,
24. Sojka D, Franta Z, Horn M, Caffrey CR, Mareš M, Kopáček P. New insights into the machinery of
25. Donohue KV, Khalil SMS, Sonenshine DE, Roe RM. Heme-binding storage proteins in the
protein-linked acetylcholine receptor from Caenorhabditis elegans. Journal of Neurochemistry 2000;
75: 1800–1809.
Identification of protective antigens by RNA interference for control of the lone star tick, Amblyomma
29. Drummond RO, Whetstone TM. Oviposition of the Cayenne tick, Amblyomma cajennense (F.), in the
Medical Biology. 2010; 708: 137-62.
tick-resistant cattle of the recombinantly expressed Rhipicephalus microplus serine protease inhibitor-3
associated cathepsin D hemoglobinase from tick Ixodes ricinus (IrCD1). The Journal of Biological
34. Valdés JJ, Schwarz A, Cabeza de Vaca I, Calvo E, Pedra JHF, Guallar V, et al. Tryptogalinin is a tick
35. Kol N, Shomron N. Assembly algorithms for deep sequencing data: basics and pitfalls. Methods in
using mitochondrial genomes and nuclear rRNA genes indicates that the genus Amblyomma is
A systems biology approach to the characterization of stress response in *Dermacentor reticulatus* tick unfed larvae

**Abstract**

**Background:** *Dermacentor reticulatus* (Fabricius, 1794) is distributed in Europe and Asia where it infests and transmits disease-causing pathogens to humans, pets and other domestic and wild animals. However, despite its role as a vector of emerging or re-emerging diseases, very little information is available on the genome, transcriptome and proteome of *D. reticulatus*. Tick larvae are the first developmental stage to infest hosts, acquire infection and transmit pathogens that are transovarially transmitted and are exposed to extremely stressing conditions. In this study, we used a systems biology approach to get an insight into the mechanisms active in *D. reticulatus* unfed larvae, with special emphasis on stress response.

**Results:** The results support the use of paired end sequencing and proteomics informed by transcriptomics (PIT) for the analysis of transcriptomics and proteomics data, particularly for organisms such as *D. reticulatus* with little sequence information available. The results showed that metabolic and cellular processes involving enzymatic reactions were the most active in *D. reticulates* unfed larvae, suggesting that ticks are very active during this life stage. The stress response was activated in *D. reticulatus* unfed larvae and a *Rickettsia* sp. similar to *R. raoultii* was identified in these ticks.

**Conclusions:** The activation of stress responses in *D. reticulatus* unfed larvae likely counteracts the negative effect of temperature and other stress conditions such as *Rickettsia* infection and favors tick adaptation to environmental conditions to increase tick survival. These results show mechanisms that have evolved in *D. reticulatus* ticks to survive under stress conditions and suggest that these mechanisms are conserved across hard tick species. Targeting some of these proteins by vaccination may increase tick susceptibility to natural stress conditions, which in turn reduce tick survival and reproduction, thus reducing tick populations and vector capacity for tick-borne pathogens.

1. **Background**

Ticks are blood-sucking ectoparasites that infest and transmit pathogens to humans and animals. *Dermacentor reticulatus* (Fabricius, 1794) is a three hosts tick (larvae, nymphs and adults feed on different hosts) distributed in Europe and Asia where it infests humans, pets and other domestic and wild animals. *D. reticulatus* transmit disease-causing pathogens such as *Rickettsia slovaca* (tick-borne lymphoadenopathy; TIBOLA), Omsk hemorrhagic fever virus (OHFV; Omsk hemorrhagic fever), tick-borne encephalitis virus (TBEV; tick-borne encephalitis), *Francisella tularensis* (tularemia) and *Babesia canis* (canine babesiosis) [1-3].
Despite its role as a vector of emerging or re-emerging diseases, very little information is available on the genome, transcriptome and proteome of *D. reticulatus* (115 nucleotide sequences of which only 15 were not of rRNA and 9 protein sequences deposited in the GenBank on June 2013).

Tick larvae are the first developmental stage to infest hosts, acquire infection and transmit pathogens that are transovarially transmitted. *D. reticulatus* larvae hatch at a temperature range of 20-34 °C and can survive for 83.5 days at 5 °C and 100% relative humidity [4]. However, under natural conditions, larvae are active within 16-20 days after hatching and survive about a month before feeding [5]. *D. reticulatus* larvae feed on small mammals and are active during the summer [6].

All these facts put tick unfed larvae under extremely stressing conditions. For example, under natural conditions only 5-15% *D. reticulatus* larvae produced from a single clutch are activated [5]. In this study, we characterized the transcriptome and proteome of *D. reticulatus* unfed larvae to get an insight into the mechanisms active at this developmental stage, with special emphasis on stress response.

2. Results and discussion

2.1 *D. reticulatus* protein clusters identified after transcriptomics analysis in unfed larvae

A total of 21,677,414 (~2.1 Gb) Illumina 101 bp paired-end reads (207 bp average insert size) were subjected to analysis. After read assembly, 18,946 transcripts were obtained and annotated [see Additional file 1]. Transcripts were clustered by encoded proteins. If two transcripts were annotated as the same protein, then these transcripts were clustered together in the same protein cluster. We considered each set of transcripts annotated by the same protein as a unigene to identify transcripts from the same locus/gene. This approach identified a set of 3,808 protein clusters with 1,231±286 (Ave±S.E) estimated counts per protein [see Additional file 1].

The analysis of Biological Process (BP) and Molecular Function (MF) gene ontology (GO) showed that the most represented BPs corresponded to unknown process (N=2,163; 57%), metabolic process (N=41; 11%) and cellular process (N=378; 10%) (Fig. 1A) while proteins with unknown function (N=2,163; 57%), catalytic activity (N=658; 17%) and binding activity (N=628; 16%) were the most represented (Fig. 1B).

A closer analysis of the most expressed genes showed that translation was the most represented BP and MF in *D. reticulates* unfed larvae (Figs. 2A and 2B). These results showed that metabolic
and cellular processes involving enzymatic reactions were the most active in *D. reticulates* unfed larvae (Figs. 1A, 1B, 2A, 2B), suggesting that tick metabolism is active during this life stage.

**Figure 1. Transcriptomics of *D. reticulatus* unfed larvae.** (A) Transcripts identified in *D. reticulatus* unfed larvae were functionally annotated and grouped according to the biological process of the encoded proteins. The number of proteins on each category is shown. (B) Transcripts identified in *D. reticulatus* unfed larvae were functionally annotated and grouped according to the molecular function of the encoded proteins. The number of proteins on each category is shown.
Figure 2. Five hundred highly expressed genes. (A) The 500 more represented unigenes (protein clusters) identified in *D. reticulatus* unfed larvae were functionally annotated and grouped according to the biological process of the encoded proteins. The number of proteins on each category is shown. (B) The 500 more represented unigenes (protein clusters) identified in *D. reticulatus* unfed larvae were functionally annotated and grouped according to the molecular function of the encoded proteins. The number of proteins on each category is shown.

2.2. *D. reticulatus* proteins identified after proteomics analysis in unfed larvae

Two approaches were used for proteomics data analysis. In the first approach, mass spectra were searched against Ixodida protein database. In the second approach, a recently described technique named proteomics informed by transcriptomics (PIT) [7] was used. This approach uses a database created from transcriptomics data to search mass spectra and has been reported to increase the number of identified proteins [7]. Herein, a total of 74 proteins were identified using the first approach [see Additional file 2]. PIT approach resulted in 104 proteins identified in unfed tick larvae [see Additional file 2], representing a 40% increase with respect to the search against Ixodida protein database. The analysis of de novo sequences increased the number of identified
proteins using both approaches for proteomics data analysis [see Additional file 2]. However, while de novo protein sequences represented 4% (N=3) of the identified proteins searching against Ixodida protein database, the number of identified proteins increased in 47% (N=49) using PIT [see Additional file 2]. These results support the use of PIT for the analysis of proteomics data, particularly for organisms such as *D. reticulatus* with little sequence information available.

After removing proteins with unknown BP and MF, transcriptomics and proteomics data correlated well with respect to the most represented BPs (Figs. 3A-3C) and MFs (Figs. 4A-4C). These results provided additional support for the identified mechanisms active in *D. reticulatus* unfed larvae.

Figure 3. Biological processes identified in *D. reticulatus* unfed larvae. (A) Transcripts identified in *D. reticulatus* unfed larvae were functionally annotated and grouped according to the biological process of the encoded proteins after removing transcripts with unknown function. (B) Proteins identified in *D. reticulatus* unfed larvae after searching against Ixodida database were functionally annotated and grouped according to their biological process. (C) Proteins identified in *D. reticulatus* unfed larvae after searching against transcripts database (PIT) were functionally annotated and grouped according to their biological process. The number of proteins on each category is shown.
2.3. *Rickettsia* sp. identified in *D. reticulatus* unfed larvae

Some reads matching *Rickettsia* spp. were identified in *D. reticulatus* unfed larvae (N=16) [see Additional file 3]. These transcripts were probably wrongly annotated as *I. scapularis* proteins in Uniprot when they are likely *Rickettsia* proteins. In these cases, the Uniref representative protein of the cluster to which belongs the *I. scapularis* protein is a *Rickettsia* protein and the rest of the members of the Uniref90 cluster are also from *Rickettsia*. Proteomics analysis corroborated the presence of *Rickettsia* proteins in *D. reticulatus* unfed larvae with the identification of 14 proteins searching against Rickettsiae database [see Additional file 3].
This *Rickettsia* sp. could be a commensal bacterium that has been described in *Dermacentor* and other tick species, but not in *D. reticulatus* [8-11] or a pathogen [3]. The *Rickettsia* proteins identified in *D. reticulatus* unfed larvae are highly conserved among *Rickettsia* spp. and thus not suitable to characterize these organisms at the species level.

To gain further information on this *Rickettsia* sp., the PCR amplification and sequencing of gene markers that have been previously used for species classification was conducted [12-14]. The results showed >99% pairwise nucleotide sequence identity to *Rickettsia* sp. sequences, especially to *R. raoultii* (Table 3). As previously shown [13], the *in silico* *Pst*I and *Rsa*I restriction analysis of *ompA* sequences was highly informative and corroborated that the *Rickettsia* sp. identified in this study is similar to *R. raoultii*. These results suggested that the *Rickettsia* sp. identified in *D. reticulatus* unfed larvae is not a symbiont but the tick-borne pathogen, *R. raoultii*. However, until this *Rickettsia* sp. is fully characterized, we cannot exclude the possibility of a symbiont closely related to *R. raoultii*. These results suggested that the pathogen could be an additional stress factor in *D. reticulatus* unfed larvae, which correlated with the activation of immune response in these ticks (Figs. 1A, 3A and 3C). *Rickettsia* sequences were deposited in the GenBank with accession numbers [GenBank: KF478838, KF478839].

### 2.4. Stress response in *D. reticulatus* unfed larvae

Transcripts and proteins mapped to stress response BP in *D. reticulatus* unfed larvae were selected for further analysis. Transcriptomics results showed that heat shock, cold shock and other stress responses were active in unfed larvae, represented by 39 protein clusters (1% of all identified protein clusters) and 27,937 counts (Table 1). Of them, the most represented functions corresponded to heat shock response (Figs. 5A and 5B).

Proteomics analysis revealed 4 and 5 stress response proteins after searching against Ixodida protein database and PIT, respectively (Table 2). In agreement with transcriptomics data, the most represented function corresponded to heat shock response (Figs. 5C and 5D).

Some transcripts mapped to stress response BP were selected for the characterization of mRNA levels in *D. reticulatus* tick unfed larvae and adult guts and salivary glands by real-time RT-PCR (Fig. 6).
Figure 5. Stress response in *D. reticulatus* unfed larvae. (A) Stress response transcripts identified in *D. reticulatus* unfed larvae were grouped according to the function of their encoded protein. The number of proteins and percent in each category is shown. (B) Number of counts per protein (Ave+S.E.) in stress response proteins identified by transcriptomics analysis in *D. reticulatus* unfed larvae. (C) Stress response proteins identified in *D. reticulatus* unfed larvae were grouped according to the function of their encoded protein. The number of proteins and percent in each category is shown. (D) Number of peptides per protein (Ave+S.D.) in stress response proteins identified by proteomics analysis in *D. reticulatus* unfed larvae.

Figure 6. mRNA levels for selected genes encoding for stress response proteins. The mRNA levels were characterized by real-time RT-PCR in *D. reticulatus* unfed larvae and adult female and male guts and salivary glands (N=3), normalized against tick ribosomal protein S4 and shown as Ave+S.D. in arbitrary units.
Table 1. Tick stress response proteins identified in *D. reticulatus* unfed larvae after transcriptomics analysis.

<table>
<thead>
<tr>
<th>Uniprot ID</th>
<th>Counts per protein</th>
<th>Protein name</th>
<th>Gene name</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>B7PAR6</td>
<td>6894</td>
<td>HSP</td>
<td>ISCW017456</td>
<td><em>Ixodes scapularis</em></td>
</tr>
<tr>
<td>B7Q101</td>
<td>5601</td>
<td>HSP90</td>
<td>ISCW014265</td>
<td><em>Ixodes scapularis</em></td>
</tr>
<tr>
<td>L7M330</td>
<td>1830</td>
<td>Heat shock-related protein</td>
<td></td>
<td><em>Rhipicephalus pulchellus</em></td>
</tr>
<tr>
<td>E4W3Z2</td>
<td>1609</td>
<td>HSP70 protein 5</td>
<td></td>
<td><em>Haemaphysalis longicornis</em></td>
</tr>
<tr>
<td>L7LXP1</td>
<td>805</td>
<td>HSP90 co-chaperone p23</td>
<td></td>
<td><em>Rhipicephalus pulchellus</em></td>
</tr>
<tr>
<td>E2J6U8</td>
<td>652</td>
<td>Mitochondrial HSP60</td>
<td></td>
<td><em>Hyalomma marginatum rufipes</em></td>
</tr>
<tr>
<td>G3MSI6</td>
<td>608</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td><em>Amblyomma maculatum</em></td>
</tr>
<tr>
<td>L7M4B9</td>
<td>593</td>
<td>Heat shock-related protein</td>
<td></td>
<td><em>Rhipicephalus pulchellus</em></td>
</tr>
<tr>
<td>E0YPC0</td>
<td>577</td>
<td>Small HSP II</td>
<td></td>
<td><em>Rhipicephalus annulatus</em></td>
</tr>
<tr>
<td>I1ZDN9</td>
<td>455</td>
<td>HSP cognate 5</td>
<td>Hsc70-5</td>
<td><em>Aeglyophlus robustus</em></td>
</tr>
<tr>
<td>L7MCC0</td>
<td>435</td>
<td>HSP</td>
<td></td>
<td><em>Rhipicephalus pulchellus</em></td>
</tr>
<tr>
<td>F1CGQ9</td>
<td>334</td>
<td>HSP90</td>
<td>hsp90</td>
<td><em>Panonychus citri</em></td>
</tr>
<tr>
<td>B7P1Z8</td>
<td>301</td>
<td>HSP</td>
<td>ISCW016090</td>
<td><em>Ixodes scapularis</em></td>
</tr>
<tr>
<td>L7MFL0</td>
<td>257</td>
<td>Heat shock transcription factor</td>
<td></td>
<td><em>Rhipicephalus pulchellus</em></td>
</tr>
<tr>
<td>L7M6S1</td>
<td>227</td>
<td>Heat shock-related protein</td>
<td></td>
<td><em>Rhipicephalus pulchellus</em></td>
</tr>
<tr>
<td>G8Z375</td>
<td>220</td>
<td>HSP70-3</td>
<td></td>
<td><em>Panonychus citri</em></td>
</tr>
<tr>
<td>Q0V9A5</td>
<td>211</td>
<td>HSP70-1</td>
<td>hspa11 hspa1b</td>
<td><em>Xenopus tropicalis</em></td>
</tr>
<tr>
<td>B5M740</td>
<td>195</td>
<td>HSP90</td>
<td></td>
<td><em>Amblyomma americanum</em></td>
</tr>
<tr>
<td>J7G3V2</td>
<td>173</td>
<td>Heat shock cognate protein 70</td>
<td></td>
<td><em>Latrodectus hesperus</em></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Accession</th>
<th>Score</th>
<th>Description</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>L7M1L7</td>
<td>137</td>
<td>Heat shock-related protein</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>B4YTU0</td>
<td>128</td>
<td>HSP70-3</td>
<td>Tetranychus cinnabarinus</td>
</tr>
<tr>
<td>L7LYK1</td>
<td>121</td>
<td>Heat shock transcription factor</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>F09J7M</td>
<td>35</td>
<td>HSP9</td>
<td>Amblyomma variegatum</td>
</tr>
<tr>
<td>D8KWR5</td>
<td>33</td>
<td>HSP70</td>
<td>Haemaphysalis longicornis</td>
</tr>
<tr>
<td>B7P8Q5</td>
<td>33</td>
<td>HSP70</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>L7M513</td>
<td>30</td>
<td>Putative ahsa1 c14orf3 hspc322: activator of 90 kDa HSP atpase log</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>L7M6W4</td>
<td>21</td>
<td>HSP60</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>L7M597</td>
<td>16</td>
<td>HSP40</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>B7PRX5</td>
<td>14</td>
<td>Heat shock transcription factor</td>
<td>Ixodes scapularis</td>
</tr>
</tbody>
</table>

**Cold shock response proteins**

<table>
<thead>
<tr>
<th>Accession</th>
<th>Score</th>
<th>Description</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>B7PD37</td>
<td>531</td>
<td>Translation initiation factor 2, alpha subunit</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>L7MEM0</td>
<td>27</td>
<td>Putative cold shock domain protein</td>
<td>Rhipicephalus pulchellus</td>
</tr>
</tbody>
</table>

**Other stress response proteins**

<table>
<thead>
<tr>
<th>Accession</th>
<th>Score</th>
<th>Description</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2XW15</td>
<td>2934</td>
<td>Glutathione peroxidase</td>
<td>PHGPX</td>
</tr>
<tr>
<td>L7M323</td>
<td>504</td>
<td>Putative nucleotide kinase/nuclear protein involved oxidative stress response</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>B7QC85</td>
<td>459</td>
<td>Tumor rejection antigen, Gp96</td>
<td>ISCW022766</td>
</tr>
<tr>
<td>B7QG63</td>
<td>419</td>
<td>Glutathione peroxidase</td>
<td>ISCW022517</td>
</tr>
<tr>
<td>B7PUM7</td>
<td>232</td>
<td>Peroxinectin</td>
<td>ISCW007552</td>
</tr>
<tr>
<td>B7PP36</td>
<td>182</td>
<td>Peroxinectin</td>
<td>ISCW006862</td>
</tr>
<tr>
<td>P62140</td>
<td>64</td>
<td>Serine/threonine-protein phosphatase PP1-beta catalytic subunit</td>
<td>PPP1CB</td>
</tr>
</tbody>
</table>

68
<table>
<thead>
<tr>
<th>Uniprot ID</th>
<th>Counts per protein</th>
<th>Protein name</th>
<th>Gene name</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>B7PAR6</td>
<td>6894</td>
<td>HSP</td>
<td>ISCW017456</td>
<td><em>Ixodes</em> <em>scapularis</em></td>
</tr>
<tr>
<td>B7QI01</td>
<td>5601</td>
<td>HSP90</td>
<td>ISCW014265</td>
<td><em>Ixodes</em> <em>scapularis</em></td>
</tr>
<tr>
<td>L7M330</td>
<td>1830</td>
<td>Heat shock-related protein</td>
<td></td>
<td><em>Rhipicephalus</em> <em>pulchellus</em></td>
</tr>
<tr>
<td>E4W3Z2</td>
<td>1609</td>
<td>HSP70 protein 5</td>
<td></td>
<td><em>Haemaphysalis</em> <em>longicornis</em></td>
</tr>
<tr>
<td>L7LXP1</td>
<td>805</td>
<td>HSP90 co-chaperone p23</td>
<td></td>
<td><em>Rhipicephalus</em> <em>pulchellus</em></td>
</tr>
<tr>
<td>E2J6U8</td>
<td>652</td>
<td>Mitochondrial HSP60</td>
<td></td>
<td><em>Hyalomma</em> <em>marginatum</em> rufipes</td>
</tr>
<tr>
<td>G3MSI6</td>
<td>608</td>
<td>Putative uncharacterized protein</td>
<td></td>
<td><em>Amblyomma</em> <em>maculatum</em></td>
</tr>
<tr>
<td>L7M4B9</td>
<td>593</td>
<td>Heat shock-related protein</td>
<td></td>
<td><em>Rhipicephalus</em> <em>pulchellus</em></td>
</tr>
<tr>
<td>E0YPC0</td>
<td>577</td>
<td>Small HSP II</td>
<td></td>
<td><em>Rhipicephalus</em> <em>annulatus</em></td>
</tr>
<tr>
<td>I1ZDN9</td>
<td>455</td>
<td>HSP cognate 5</td>
<td>Hsc70-5</td>
<td><em>Aeroglyphus</em> <em>robustus</em></td>
</tr>
<tr>
<td>L7MCC0</td>
<td>435</td>
<td>HSP</td>
<td></td>
<td><em>Rhipicephalus</em> <em>pulchellus</em></td>
</tr>
<tr>
<td>F1CGQ9</td>
<td>334</td>
<td>HSP90</td>
<td>hsp90</td>
<td><em>Panonychus</em> <em>citri</em></td>
</tr>
<tr>
<td>B7P1Z8</td>
<td>301</td>
<td>HSP</td>
<td>ISCW016090</td>
<td><em>Ixodes</em> <em>scapularis</em></td>
</tr>
<tr>
<td>L7MFL0</td>
<td>257</td>
<td>Heat shock transcription factor</td>
<td></td>
<td><em>Rhipicephalus</em> <em>pulchellus</em></td>
</tr>
<tr>
<td>L7M6S1</td>
<td>227</td>
<td>Heat shock-related protein</td>
<td></td>
<td><em>Rhipicephalus</em> <em>pulchellus</em></td>
</tr>
<tr>
<td>G8Z375</td>
<td>220</td>
<td>HSP70-3</td>
<td></td>
<td><em>Panonychus</em> <em>citri</em></td>
</tr>
<tr>
<td>Q0V9A5</td>
<td>211</td>
<td>HSP70-1</td>
<td>hspa11 hspa1b</td>
<td><em>Xenopus</em> <em>tropicalis</em></td>
</tr>
<tr>
<td>Accession</td>
<td>Value</td>
<td>Description</td>
<td>Species</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>B5M740</td>
<td>195</td>
<td>HSP90</td>
<td>Amblyomma americanum</td>
<td></td>
</tr>
<tr>
<td>J7G3V2</td>
<td>173</td>
<td>Heat shock cognate protein 70</td>
<td>Latrodectus hesperus</td>
<td></td>
</tr>
<tr>
<td>L7M1L7</td>
<td>137</td>
<td>Heat shock-related protein</td>
<td>Rhipicephalus pulchellus</td>
<td></td>
</tr>
<tr>
<td>B4YTU0</td>
<td>128</td>
<td>HSP70-3</td>
<td>Tetanychus cinnabarinus</td>
<td></td>
</tr>
<tr>
<td>L7LYK1</td>
<td>121</td>
<td>Heat shock transcription factor</td>
<td>Rhipicephalus pulchellus</td>
<td></td>
</tr>
<tr>
<td>F0J9M7</td>
<td>35</td>
<td>HSP9</td>
<td>Amblyomma variegatum</td>
<td></td>
</tr>
<tr>
<td>D8KWR5</td>
<td>33</td>
<td>HSP70</td>
<td>Haemaphysalis longicornis</td>
<td></td>
</tr>
<tr>
<td>B7P8Q5</td>
<td>33</td>
<td>HSP70</td>
<td>Ixodes scapularis</td>
<td></td>
</tr>
<tr>
<td>L7M513</td>
<td>30</td>
<td>Putative ahsa1 c14orf3 hspe322: activator of 90 kDa HSP atpase log 1</td>
<td>Rhipicephalus pulchellus</td>
<td></td>
</tr>
<tr>
<td>L7M6W4</td>
<td>21</td>
<td>HSP60</td>
<td>Rhipicephalus pulchellus</td>
<td></td>
</tr>
<tr>
<td>L7M597</td>
<td>16</td>
<td>HSP40</td>
<td>Rhipicephalus pulchellus</td>
<td></td>
</tr>
<tr>
<td>B7PRX5</td>
<td>14</td>
<td>Heat shock transcription factor</td>
<td>Ixodes scapularis</td>
<td></td>
</tr>
</tbody>
</table>

**Cold shock response proteins**

<table>
<thead>
<tr>
<th>Accession</th>
<th>Value</th>
<th>Description</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>B7PD37</td>
<td>531</td>
<td>Translation initiation factor 2, alpha subunit</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>L7MEM0</td>
<td>27</td>
<td>Putative cold shock domain protein</td>
<td>Rhipicephalus pulchellus</td>
</tr>
</tbody>
</table>

**Other stress response proteins**

<table>
<thead>
<tr>
<th>Accession</th>
<th>Value</th>
<th>Description</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2XW15</td>
<td>2934</td>
<td>Glutathione peroxidase</td>
<td>PHGPX</td>
</tr>
<tr>
<td>L7M323</td>
<td>504</td>
<td>Putative nucleotide kinase/nuclear protein involved oxidative stress response</td>
<td>Rhipicephalus pulchellus</td>
</tr>
<tr>
<td>B7QC85</td>
<td>459</td>
<td>Tumor rejection antigen, Gp96</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>B7QG63</td>
<td>419</td>
<td>Glutathione peroxidase</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>B7PUM7</td>
<td>232</td>
<td>Peroxinecin</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>B7PP36</td>
<td>182</td>
<td>Peroxinecin</td>
<td>Ixodes scapularis</td>
</tr>
<tr>
<td>Gene marker</td>
<td>Rickettsia sp. (Genbank accession no.)</td>
<td>Sequence identity</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>atpA</td>
<td>R. raoultii (KC428000)</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td>dnaK</td>
<td>R. sibirica subsp. mongolitimonae (KC428015)</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. massiliae (KC428014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. slovaca (CP003375)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S rDNA</td>
<td>R. raoultii (EU036982)</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rickettsia sp. RpA4 (AF120026)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ompB</td>
<td>R. raoultii (DQ365797)</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncultured Rickettsia sp. clone R2012 (JQ320341)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ompA</td>
<td>R. raoultii (HM161789)</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>recA</td>
<td>R. raoultii (KC428038)</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. massiliae (GQ144452)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[\text{P62140} \quad \text{64} \quad \text{Serine/threonine-protein phosphatase PP1-beta catalytic subunit} \quad \text{PPP1CB} \quad \text{Homo sapiens}

\[\text{L7M2W8} \quad \text{39} \quad \text{Putative bola bacterial stress-induced morphogen-related protein} \quad \text{Rhipicephalus pulchellus}\n
\text{\textsuperscript{a}}Identified by PIT.
\text{\textsuperscript{b}}Identified searching against Ixodida.

Table 3. Sequence identity for the Rickettsia sp. identified in D. reticulatus unfed larvae.

Table 3. Sequence identity of the Rickettsia sp. identified in D. reticulatus unfed larvae.

<table>
<thead>
<tr>
<th>Gene marker</th>
<th>Rickettsia sp. (Genbank accession no.)</th>
<th>Sequence identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>atpA</td>
<td>R. raoultii (KC428000)</td>
<td>99%</td>
</tr>
<tr>
<td>dnaK</td>
<td>R. sibirica subsp. mongolitimonae (KC428015)</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>R. massiliae (KC428014)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. slovaca (CP003375)</td>
<td></td>
</tr>
<tr>
<td>16S rDNA</td>
<td>R. raoultii (EU036982)</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Rickettsia sp. RpA4 (AF120026)</td>
<td></td>
</tr>
<tr>
<td>ompB</td>
<td>R. raoultii (DQ365797)</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Uncultured Rickettsia sp. clone R2012 (JQ320341)</td>
<td></td>
</tr>
<tr>
<td>ompA</td>
<td>R. raoultii (HM161789)</td>
<td>100%</td>
</tr>
<tr>
<td>recA</td>
<td>R. raoultii (KC428038)</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td>R. massiliae (GQ144452)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. PCR conditions and primer sequences used for real-time RT-PCR.

<table>
<thead>
<tr>
<th>Gene (Uniprot ID)</th>
<th>Forward and reverse primers</th>
<th>PCR annealing conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hsp (B7PAR6)</strong></td>
<td>GACAAGGGGCGCTCAGAAAA</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>CGACTTGATAGCCCTCCCT</td>
<td></td>
</tr>
<tr>
<td><strong>hsp (B7P1Z8)</strong></td>
<td>TTGAGGAGAAGGACAGACTGG</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>GACGACTTGCGGTGTTGTTC</td>
<td></td>
</tr>
<tr>
<td><strong>hsp70 (B7P8Q5)</strong></td>
<td>TCGATATCCACCTCGTCCGT</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>GCAGTAAGGAAGGGCGTTG</td>
<td></td>
</tr>
<tr>
<td><strong>Translation initiation factor 2 (tif2), alpha subunit (B7PD37)</strong></td>
<td>CACTGATGCGTGGCGAAAA</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>CCGGACACTTCCCTGTGCTC</td>
<td></td>
</tr>
<tr>
<td><strong>Putative cold shock domain protein, csp (L7MEM0)</strong></td>
<td>CACTACAGCCAGTTCTCGGG</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>CACCTCATCGCTAAGGACTCCT</td>
<td></td>
</tr>
<tr>
<td><strong>Tumor rejection antigen, gp96 (B7QC85)</strong></td>
<td>CGGCTGTGAAGAAGGGCCTA</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>CCCTCGTCAACCTTGAGACC</td>
<td></td>
</tr>
<tr>
<td><strong>Putative <em>bola</em> bacterial stress-induced morphogen-related protein (L7M2W8)</strong></td>
<td>TGAGCTGGAGGACGTTCAG</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>CATTCAACGAGATGCTGC</td>
<td></td>
</tr>
<tr>
<td><strong>Ribosomal protein S4, rpS4 (DQ066214)</strong></td>
<td>GGTGAAGAAGATGTCAAGCAGAG</td>
<td>60 ºC/30s</td>
</tr>
<tr>
<td></td>
<td>TGAAGGCTCAGAGGGTAGTTT</td>
<td></td>
</tr>
</tbody>
</table>
The results showed that all selected genes encoding for stress response proteins were more expressed in unfed larvae than in adult tissues, thus reinforcing the significance of this BP in *D. reticulatus* tick unfed larvae. The sequences of *D. reticulatus* genes encoding for stress response proteins were deposited in the GenBank with accession numbers [GenBank: KF478838, KF478839].

Ticks are very sensitive to temperature and their life cycle is dependent on a complex combination of climate variables for development and survival [15]. In particular, *D. reticulatus* tick unfed larvae are exposed to extremely stressing conditions that affect their survival and development [5]. The heat-shock and other stress responses are a conserved reaction of cells and organisms to different stress conditions such as extreme temperatures, toxicity and pathogen infection [16]. Crucial to cell survival is the sensitivity of proteins and enzymes to heat inactivation and denaturation. Therefore, adaptive mechanisms exist that protect cells from the proteotoxic effects of stress. The heat-shock proteins and other stress response proteins protect cells and organisms from damage, providing higher levels of tolerance to environmental stress. Recent studies demonstrated that the stress response is activated in ticks and cultured tick cells in response to *Anaplasma* spp. infection and heat shock [17, 18]. These results showed that at high temperatures and during blood feeding, when *hsp20, hsp70* and *subolesin* are over-expressed, *Ixodes scapularis* ticks are protected from stress and pathogen infection and have a higher questing speed. These results suggested a connection between tick stress response, questing behavior and pathogen infection [17, 18], which may be present also in *D. reticulatus* tick unfed larvae.

3. Conclusions

The activation of stress responses in *D. reticulatus* unfed larvae likely counteracts the negative effect of temperature and other stress conditions such as *Rickettsia* infection and favors tick adaptation to environmental conditions to increase tick survival. These results are relevant to understand how *D. reticulatus* ticks have evolved mechanisms to survive under stress conditions and suggest that these mechanisms are conserved across hard tick species. Targeting some of these proteins by vaccination may increase tick susceptibility to natural stress conditions, which in turn reduce tick survival and reproduction, thus reducing tick populations and vector capacity for tick-borne pathogens [19].
4. Methods

4.1. Ticks and sample preparation

*D. reticulatus* unfed larvae were obtained from a single female from a Dutch colony maintained at the Utrecht Centre for Tick-borne Diseases (UCTD), Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands. Total RNA and DNA were extracted from approximately 500 *D. reticulatus* larvae using the AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Valencia, CA, USA) according to manufacturer instructions. RNA was purified with the RNeasy MinElute Cleanup Kit (Qiagen, Valencia, CA, USA) and characterized using the Agilent 2100 Bioanalyzer (Santa Clara, CA, USA) in order to evaluate the quality and integrity of RNA preparations. RNA concentration was determined using the Nanodrop ND-1000 (NanoDrop Technologies Wilmington, Delaware USA). Approximately 200 *D. reticulatus* larvae were pulverized in liquid nitrogen and homogenized with a glass homogenizer (20 strokes) in 4 ml buffer (0.25 M sucrose, 1 mM MgCl2, 10 mM Tris-HCl, pH 7.4) supplemented with 4% SDS and complete mini protease inhibitor cocktail (Roche, Basel, Switzerland). Sample was sonicated for 1 min in an ultrasonic cooled bath followed by 10 sec vortex. After 3 cycles of sonication-vortex, the homogenate was centrifuged at 200xg for 5 min at room temperature to remove cellular debris. The supernatant was collected and protein concentration was determined using the BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) using BSA as standard.

*D. reticulatus* unfed female and male adults were obtained from a colony of ticks originally collected in southern Slovakia and maintained at the Biology Centre of the ASCR, Parasitology Institute, České Budějovice, Czech Republic. Ticks were dissected and total RNA was extracted from 3 guts and salivary glands using TriReagent (Sigma, St. Louis, MO, USA) following manufacturer’s recommendations.

4.2. Transcriptomics data acquisition

The RNA purified from unfed tick larvae was used for library preparation using the TruSeq RNA sample preparation kit v.1 and the standard low throughput procedure (Illumina, San Diego, CA, USA). Briefly, 0.7 µg total RNA was used as starting material for library preparation. Messenger RNA was captured using poly-dT magnetic beads and purified polyA+ RNA was chemically fragmented and reverse-transcribed. Remaining RNA was enzymatically removed and the second strand generated following an end repair procedure and preparation of double-stranded cDNA for
adaptor ligation. Adaptor oligonucleotides containing the signals for subsequent amplification and sequencing were ligated to both ends and the cDNA was washed using AMPure SPRI-based magnetic beads (Beckman Coulter, IZASA, Barcelona, Spain). Adaptors contained identifiers, which allow multiplexing in the sequencing run. An enrichment procedure based on PCR was then performed to ensure that all molecules in the library conserved the adapters at both ends. The number of PCR cycles was adjusted to 15. The final amplified library was checked again on a BioAnalyzer 2100 (Agilent, Santa Clara, CA, USA) and titrated by quantitative PCR using a reference standard to characterize molecules concentration in the library (12.44 nM). The library was denatured and seeded on the lane of the flowcell at a final concentration after renaturalization of 10-14 pM. After binding, clusters were formed in the flowcell by local amplification using a Cluster Station apparatus (Illumina). Following sequencing primer annealing, flowcell was loaded into a GAIIx equipment (Illumina) to perform sequencing using the TruSeq® system (Illumina). The sample was run under a pair-end 2x100bp protocol for de novo sequencing. After sequencing and quality filtering, reads were split according to their different identifiers and fastq files were generated to proceed to quality analysis and de novo transcript assembly and gene expression analysis.

4.3. Bioinformatics for transcriptomics data

Sequence reads were trimmed at the error probability higher than 0.05 and assembled only when two members of the pair remained after filtering at trimming. Oases [20] was used for read assembly in the mode of single (not merged) assembly because results were better in this mode. A K value of 79 was chosen, which was very close to the total length of the read (~100 bp) to avoid misassemblies since the higher the overlapping required the more accurate the transcript is. Final assembly was explored in detail using Tablet (http://bioinf.scri.ac.uk/tablet/download.shtml) [21]. Functional annotations were inferred by similarity to Uniprot reference proteins using Blast E values < 10E-10. We selected a set of 34,095 reference proteins downloaded from Uniprot on March 7, 2013, including all proteins that were representative of Uniref90 clusters belonging to the taxonomic node Chelicerata, which are 8 levels above D. reticulatus taxon. In the Uniref90 clusters, each protein belongs to only one cluster with a 90% similarity to the representative protein for all members of the cluster. It provides a more homogeneous and uniform distance between reference proteins. Reference proteins were used for transcript clusterization to obtain a protein-centred analysis of gene expression that is more useful for functional analysis in a de novo transcriptome. The eXpress algorithm was used for mapping reads to multiple targets to quantify gene expression levels [22].
4.4. Proteomics data acquisition

Protein extract (150 μg) was precipitated following the methanol/chloroform procedure [23], resuspended in 100 μl Laemmli sample buffer and applied onto 1.2-cm wide wells on a 12% SDS-PAGE gel. The electrophoretic run was stopped as soon as the front entered 3 mm into the resolving gel, so that the whole proteome became concentrated in the stacking/resolving gel interface. The unseparated protein band was visualized by staining with GelCode Blue Stain Reagent (Thermo Scientific), excised, cut into 2x2 mm cubes and digested overnight at 37°C with 60 ng/μl sequencing grade trypsin (Promega, Madison, WI, USA) at 5:1 protein:trypsin (w/w) ratio in 50 mM ammonium bicarbonate, pH 8.8 containing 10% (v/v) acetonitrile [24]. The resulting tryptic peptides from the gel band were extracted by 30 min-incubation in 12 mM ammonium bicarbonate, pH 8.8. Trifluoroacetic acid was added to a final concentration of 1% and the peptides were finally desalted onto OMIX Pipette tips C18 (Agilent Technologies, Santa Clara, CA, USA), dried-down and stored at -20°C until mass spectrometry analysis. The desalted protein digest was resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS using an Easy-nLC II system coupled to an ion trap LCQ Fleet mass spectrometer (Thermo Scientific). The peptides were concentrated (on-line) by reverse phase chromatography using a 0.1 x 20 mm C18 RP precolumn (Thermo Scientific), and then separated using a 0.075 x 100 mm C18 RP column (Thermo Scientific) operating at 0.3 μl/min. Peptides were eluted using a 180-min gradient from 5 to 35% solvent B (Solvent A: 0.1% formic acid in water, solvent B: 0.1% formic acid in acetonitrile). ESI ionization was done using a Fused-silica PicoTip Emitter ID 10μm (New Objective, Woburn, MA, USA) interface. Peptides were detected in survey scans from 400 to 1600 amu (1 μscan), followed by three data dependent MS/MS scans (Top 3), using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion applied during 30 sec periods.

4.5. Bioinformatics for proteomics data

The MS/MS raw files were searched against Ixodida (40,849 entries in June 2013) and Rickettsiaeae (58,899 entries in June 2013) Uniprot databases and against a database created from transcriptomics data (PIT) [7] using the SEQUEST algorithm (Proteome Discoverer 1.3, Thermo Scientific) with the following constraints: tryptic cleavage after Arg and Lys, up to two missed cleavage sites, and tolerances of 1 Da for precursor ions and 0.8 Da for MS/MS fragment ions and the searches were performed allowing optional Met oxidation and Cys carbamidomethylation. De novo peptide sequencing was conducted with Peaks Studio 6.0
software (Bioinformatics Solutions Inc., Waterloo, ON Canada). A false discovery rate (FDR) < 0.01 was considered as condition for successful peptide assignments.

### 4.6. Gene and protein ontology assignments

Functional data for each protein were obtained from Uniprot and included GO annotations, EC number and Interpro motifs. Assignment of GO terms to identified proteins was done by Blast2GO software (version 2.6.6; [http://www.blast2go.org/](http://www.blast2go.org/)) in three main steps: blasting to find homologous sequences, mapping to collect GO-terms associated to blast hits and annotation to assign functional terms to query sequences from the pool of GO terms collected in the mapping step [25]. Sequence data of identified proteins were uploaded as FASTA file to the Blast2GO software and the function assignment was based on GO database. The blast step was performed against NCBI public databases through blastp. Other parameters were kept at default values: e-value threshold of 1e-3, recovery of 20 hits per sequence, minimal alignment length (hsp filter) 33 (to avoid hits with matching region smaller than 100 nucleotides) and Blast mode was set to QBlast-NCBI. Configuration for annotation was an e-value-Hit-filter of 1.0E-6, annotation cut off of 55 and GO weight of 5. For visualizing the functional information (GO categories: Molecular Function and Biological process), the analysis tool of the Blast2GO software was used. The GO analysis for the 500 more represented unigenes (protein clusters) was based on the GO annotations included in the Uniprot entry of the representative protein of each cluster. The GO analysis was done using Bio4j Go Tools developed by Era7 Bioinformatics and available at [http://gotools.bio4j.com:8080/Bio4jTestServer/Bio4jGoToolsWeb.html](http://gotools.bio4j.com:8080/Bio4jTestServer/Bio4jGoToolsWeb.html). Bio4j Go Tools is a set of GO related Web Services using the open source graph bioinformatics platform Bio4j as back-end. Bio4j is a graph-based database including most data available in UniProt KB (SwissProt + Trembl), Gene Ontology (GO), UniRef (50,90,100), RefSeq, NCBI taxonomy, and Expasy Enzyme ([http://bio4j.com/](http://bio4j.com/)). Specifically designed java programs were used for the generation of the GO frequency chart data.

### 4.7. Analysis of mRNA levels by real-time RT-PCR

Real-time RT-PCR was performed on tick RNA samples with gene specific primers (Table 4) using the iScript One-Step RT-PCR Kit with SYBR Green and the iQ5 thermal cycler (Bio-Rad, Hercules, CA, USA) following manufacturer's recommendations. A dissociation curve was run at the end of the reaction to ensure that only one amplicon was formed and that the amplicons denatured consistently in the same temperature range for every sample [26]. The mRNA levels...
were normalized against tick ribosomal protein S4 using the genNorm method (ddCT method as implemented by Bio-Rad iQ5 Standard Edition, Version 2.0) [27, 28].

4.8. PCR and sequence analysis of Rickettsia amplicons

*Rickettsia* sp. DNA was characterized by PCR, cloning and sequence analysis of the amplicons. At least three clones were sequenced for each amplicon. Genes targeted by PCR included fragments of ATP synthase alpha subunit (*atpA*), heat-shock protein 70 (*dnaK*), outer membrane protein A (*ompA*), outer membrane protein B (*ompB*), 16S rRNA, and *recA* [12-14]. Nucleotide sequence identity to reference strains and *in silico* *Pst*I and *Rsa*I restriction analysis of *ompA* sequences was used to characterize *Rickettsia* sp. [13, 14].

Authors’ contribution
JF designed the study. RCG performed transcriptomics analysis. MV, MP and LM-H performed proteomics analysis. MM and RT performed the bioinformatics analysis. IGFM performed the analysis of *Rickettsia* sp. JF wrote the manuscript and all authors edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements
We thank Z. Zivkovic (Beaphar B.V., Hedel, The Netherlands) and J. Erhart and L. Grubhoffer (Biology Centre of the ASCR, Parasitology Institute, České Budějovice, Czech Republic) for providing tick samples. This research was supported by grants BFU2011-23896 and the EU FP7 ANTIGONE project number 278976. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases) within the FP7-PEOPLE-ITN programme (EU Grant No. 238511). N. Ayllón and R.C. Galindo were funded by MEC, Spain.

Supplementary material.

Table S1. Tick transcripts and encoded proteins identified in transcriptomics analysis of *D. reticulatus* unfed larvae.

Table S2. Tick proteins identified in proteomics analysis of *D. reticulatus* unfed larvae.

Table S3. *Rickettsia* sp. transcripts and proteins identified in *D. reticulatus* unfed larvae.

References
5. Filchagov AV, Lebedeva NN: The ecology of hungry larvae of *Dermacentor reticulatus* and their
Chapter IV

Proteomics approach to the study of host effect on tick feeding and reproduction

Marina Popara, Margarita Villar, Lourdes Mateos-Hernández, Isabel G. Fernández de Mera, Anabel Marina, Mercedes del Valle, Consuelo Almazán, Ana Domingos, José de la Fuente, Lesser protein degradation machinery correlates with higher BM86 tick vaccine efficacy in *Rhipicephalus annulatus* when compared to *R. microplus*. Vaccine 2013; 31:4728–4735.
Abstract

**Background:** Cattle ticks, *Rhipicephalus (Boophilus) microplus*, are a serious threat to animal health and production in many regions of the world. White tailed deer (WTD), *Odocoileus virginianus*, play a role in the maintenance and expansion of cattle tick populations. However, cattle ticks fed on WTD show lower tick numbers, weight and reproductive performance when compared to ticks fed on cattle, pointing at unknown effect of the host on tick feeding and reproduction.

**Methods:** To elucidate these factors, a proteomics approach was used to characterize tick and host proteins in *R. microplus* ticks fed on cattle and WTD.

**Results:** The results showed that *R. microplus* ticks fed on cattle have overrepresented tick proteins involved in blood digestion and reproduction when compared to ticks fed on WTD, correlating with the higher tick numbers, weight and reproductive performance observed in ticks fed on cattle. The analysis of host proteins supported these results by showing that although ticks fed on cattle probably ingested more blood as reflected by higher tick weights and Alpha-2-macroglobulin and Immunoglobulin-like protein levels, higher levels of host Hemoglobin, Haptoglobin and Albumin in ticks fed on WTD correlated with the lesser blood digestion machinery in these ticks. Higher Fibrinogen levels in ticks fed on cattle may reflect host response to tick infestations and an indicator of the lower tick infestations observed in WTD with possible implications in host-tick co-evolution. Previous results from vaccination trials in cattle suggest the possibility of using these tick proteins for the control of cattle tick infestations and provided additional support for the results presented here.

**Conclusions:** These results demonstrate the existence of host factors that affect tick feeding and reproduction that impact on vector evolution and ecology.

1. **Background**

Cattle ticks, *Rhipicephalus (Boophilus) microplus*, are a serious threat to animal health and production in many regions of the world [1]. The role of wildlife and particularly of white tailed deer (WTD), *Odocoileus virginianus*, in the maintenance of cattle tick populations has been well established [2-7]. However, although *R. microplus* can complete its developmental cycle on WTD, the weight of engorged females, oviposition and fertility are reduced by 40%, 58% and 95%, respectively when compared to ticks fed on cattle [8, 9].

These studies showed that WTD are physiologically suitable hosts for *R. microplus* [8, 9]. However, the factors responsible for the differences in tick feeding and reproduction observed
between ticks fed on cattle and WTD are unknown. The characterization of the factors affecting these differences is important to understand host effect on tick biology and the possibilities for tick control. Herein, we addressed this question by comparing the proteome of *R. microplus* ticks fed on cattle and WTD.

2. Methods

2.1. Tick collection

Adult female *R. microplus* (Susceptible Media Joya strain, CENAPA, Mexico) ticks were collected in previously reported trials after completing feeding on cattle [10] and WTD [11]. Tick infestation, data collection and analysis was similar in both experiments (Table 1; [10, 11]). Ticks were stored at -20 °C in 70% ethanol until used for protein extraction.

2.2. Protein extraction

Eight ticks from each group were dissected, cuticle removed, pulverized in liquid nitrogen and homogenized with a glass homogenizer (10 strokes) in 1 ml buffer (10 mM phosphate buffer saline (PBS), pH 7.4) supplemented with 1% SDS and complete mini protease inhibitor cocktail (Roche, Basel, Switzerland) per 50 µg sample. Samples were sonicated for 1 min in an ultrasonic cooled bath followed by 10 sec of vortex. After 3 cycles of sonication-vortex, the homogenates were centrifuged at 200xg for 5 min at room temperature to remove cellular debris. The supernatants were collected and protein concentration was determined using the BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) using BSA as standard.

2.3. Proteomics

Protein extracts (200 µg from each sample) were precipitated following the methanol/chloroform procedure [12], resuspended in 100 µl Laemmlı sample buffer and applied onto 1.2-cm wide wells on a 12% SDS-PAGE gel. The electrophoretic run was stopped as soon as the front entered 3 mm into the resolving gel, so that the whole proteome became concentrated in the stacking/resolving gel interface. The unseparated protein bands were visualized by staining with GelCode Blue Stain Reagent (Thermo Scientific), excised, cut into 2x2 mm cubes and digested overnight at 37°C with 60 ng/µl sequencing grade trypsin (Promega, Madison, WI, USA) at 5:1 protein:trypsin (w/w) ratio in 50 mM ammonium bicarbonate, pH 8.8 containing 10% (v/v) acetonitrile [13]. The resulting tryptic peptides from each band were extracted by 30 min-incubation in 12 mM ammonium bicarbonate, pH 8.8. Trifluoroacetic acid was added to a final concentration of 1% and the peptides were finally desalted onto OMIX Pipette tips C18 (Agilent...
Technologies, Santa Clara, CA, USA), dried-down and stored at -20°C until mass spectrometry analysis.

The desalted protein digest was resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS using an Agilent 1100 LC system (Agilent Technologies) coupled to a linear ion trap LTQ-Velos mass spectrometer (Thermo Scientific). The peptides were separated by reverse phase chromatography using a 0.18 mm x 150 mm Bio-Basic C18 RP column (Thermo Scientific) at 1.8 μl/min. Peptides were eluted using a 120-min gradient from 5 to 40% solvent B in solvent A (Solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid, 80% acetonitrile in water). ESI ionization was done using a microspray metal needle kit (Thermo Scientific) interface. Peptides were detected in survey scans from 400 to 1600 amu (1 μscan), followed by fifteen data dependent MS/MS scans (Top 15), using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion applied during 30 sec periods.

2.4. Proteomics data analysis

The MS/MS raw files were searched against the Ixodida (40,849 entries in June 2013) and Ruminantia (66,519 entries in June 2013) Uniprot databases with the following constraints: tryptic cleavage after Arg and Lys, up to two missed cleavage sites, and tolerances of 1.0 Da for precursor ions and 0.8 Da for MS/MS fragment ions. Due to the limited number of tick proteins included in the database, a false discovery rate (FDR) ≤ 0.05 was used for tick proteins while only host proteins with FDR ≤ 0.01 were considered. Differential protein representation for individual proteins between different samples was determined using χ² test statistics with Bonferroni correction in the IDEG6 software (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) (p=0.05).

2.5. Protein ontology assignments

Functional data for each protein were obtained from Uniprot and included gene ontology (GO) annotations, EC number and Interpro motifs. Assignment of GO terms to identified proteins was done by Blast2GO software (version 2.6.6; http://www.blast2go.org/) in three main steps: blasting to find homologous sequences, mapping to collect GO-terms associated to blast hits and annotation to assign functional terms to query sequences from the pool of GO terms collected in the mapping step [14]. Sequence data of identified proteins were uploaded as FASTA file to the Blast2GO software and the function assignment was based on GO database. The blast step was performed against NCBI public databases through blastp. Other parameters were kept at default values: e-value threshold of 1e-3, recovery of 20 hits per sequence, minimal alignment length
(hsp filter) 33 (to avoid hits with matching region smaller than 100 nucleotides) and Blast mode was set to Q Blast-NCBI. Configuration for annotation was an e-value-Hit-filter of 1.0E-6, annotation cut off of 55 and GO weight of 5. For visualizing the functional information (GO categories: Molecular Function (MF) and Biological process (BP)), the analysis tool of the Blast2GO software was used.

2.6. Western blot analysis of Cathepsin L

Total proteins (150 µg from each sample) were methanol/chloroform precipitated, resuspended in Laemmli sample buffer and separated on a 15% SDS-PAGE gel under reducing conditions. After electrophoresis, proteins were transferred to nitrocellulose membranes (Bio-Rad, Hercules, CA, USA), blocked with SuperBlock blocking buffer in TBS (Thermo Scientific) and incubated overnight at 4 ºC with rabbit polyclonal anti-Cathepsin L (Mature region No. pab0213-0; Covalab, Villeurbanne, France) antibodies. To detect the antigen-bound antibody, membranes were incubated with goat anti-rabbit IgG conjugated with horseradish peroxidase (dilution 1:10,000; Sigma-Aldrich, St. Louis, MO, USA). Immunoreactive proteins were detected by chemoluminescence using the SuperSignal West Pico chemiluminescent substrate (Thermo Scientific), visualized with an ImageQuant 350 Digital Imaging System (GE Healthcare, Pittsburgh, PA, USA), quantified using the ImageQuant TL 7.0 software (GE Healthcare) and normalized against total proteins. Normalized protein levels (N=2) were compared between samples by Student’s t-test (p=0.05).

3. Results and discussion

3.1. R. microplus tick infestations in cattle and WTD

The same strain of R. microplus was used to infest cattle and WTD under similar conditions [10, 11]. Reduced tick numbers, weight, oviposition and fertility was obtained in ticks feed on WTD when compared to ticks fed on cattle (Table 1). These results were similar to those previously reported [8, 9] and support using these ticks for the comparative proteomics analysis.

3.2. Characterization of the tick proteome

The proteomics analysis resulted in the identification of 202 and 240 tick and host proteins, respectively in R. microplus fed on cattle and WTD [see Additional files 1 and 2]. The number of peptides used for protein identification was higher in ticks fed on cattle for both tick and host proteins (Fig. 1A), pointing at the first difference between ticks fed on cattle and WTD. The GO analysis showed that the most represented BPs corresponded to metabolic (25%) and cellular.
(24%) processes (Fig. 1B) while the most represented MFs corresponded to binding (43%) and catalytic activity (37%) (Fig. 1C). Globally, differences were not observed in the composition of BPs and MFs between ticks fed on cattle and WTD (data not shown). However, the analysis of some pathways such as blood digestion and reproduction showed significant differences between ticks fed on cattle and WTD (p<0.05; Fig. 1D).

3.3. Tick proteins affecting tick infestations

Proteins involved in blood digestion and reproduction were selected trying to explain differences between ticks fed on cattle and WTD. The results of proteomics analysis showed that most of the proteins in these pathways were overrepresented in ticks fed on cattle when compared to ticks fed
on WTD (Figs. 2A and 2B; Table 2), a result that was corroborated by Western blot for Cathepsin L (Fig. 2C).

**Figure 2.** Tick and host proteins affecting blood digestion and reproduction in ticks fed on cattle and WTD. (A) Number of peptides for tick proteins involved in blood digestion in ticks fed on cattle and WTD ($\chi^2$ test; *p<0.05). (B) Number of peptides for tick proteins involved in reproduction in ticks fed on cattle and WTD ($\chi^2$ test; *p<0.05). (C) Cathepsin L protein levels were determined by Western blot in *R. microplus* fed on cattle and WTD, quantified and normalized against total proteins. Normalized protein levels (Ave+S.D. in arbitrary units) were compared between samples by Student’s t-test (*p<0.05; N=2). Abbreviation: MW, molecular weight markers. (D) Number of peptides for host proteins with significant differences ($\chi^2$ test; *p < 0.05) between ticks fed on cattle and WTD.

Lesser blood digestion machinery in ticks fed on WTD correlated with higher levels of host Hemoglobin, Haptoglobin and Albumin in feeding ticks (Fig. 2D and Table 3), showing that ticks fed on WTD digested blood poorer than ticks fed on cattle. Blood digestion is critical for tick feeding and reproduction [15, 16], and these parameters were reduced in ticks fed on WTD when compared to ticks fed on cattle (Table 1).

Furthermore, failure to properly process Hemoglobin could be toxic for the ticks [15, 16]. However, ticks fed on cattle were probably ingesting more blood as reflected by higher tick weights (Table 1) and Alpha-2-macroglobulin and Immunoglobulin-like protein levels (Fig. 2D).
and Table 3), likely because of better blood digestion in these ticks. Additionally, proteins involved in tick reproduction were also overrepresented in ticks fed on cattle (Fig. 2B), again correlating with the higher reproductive performance in these ticks when compared to ticks fed on WTD (Table 1).

These results suggested that underrepresented proteins involved in blood digestion and reproduction in ticks fed on WTD are responsible for the lower *R. microplus* tick feeding and reproduction observed in WTD when compared to cattle (Table 1; [8, 9]).

Table 1. *R. microplus* infestations in WTD and cattle.

<table>
<thead>
<tr>
<th>Experimental group</th>
<th><em>R. microplus</em> (Media Joya strain)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. ticks</td>
</tr>
<tr>
<td>WTD</td>
<td>381±195</td>
</tr>
<tr>
<td>Cattle</td>
<td>841±94</td>
</tr>
<tr>
<td>Cattle/WTD ratio</td>
<td>2.2*</td>
</tr>
<tr>
<td>(% reduction in ticks fed on WTD when compared to ticks fed on cattle)</td>
<td>-55%</td>
</tr>
</tbody>
</table>

Deer (N=4) and cattle (N=5) were infested with 10,000 *R. microplus* larvae/animal applied individually to each animal in separate cotton cells attached to the back of the animals. Adult female tick number, tick weight (mg), oviposition (egg weight (mg)/tick) and egg fertility (larvae weight/egg weight) and were compared by Chi2-test (tick numbers) or Student’s t-test with unequal variance (tick weight, oviposition and fertility) between groups (*P<0.01). Data was obtained from Canales et al. (2009) and Carreón et al. (2012) for cattle and WTD, respectively.
Table 2. Tick proteins involved in reproduction and blood digestion. Tick proteins identified in ticks fed on cattle and WTD were compared by $\chi^2$ test (*p<0.05).

<table>
<thead>
<tr>
<th>Uniprot Accession No.</th>
<th>Description</th>
<th>Ticks fed on cattle</th>
<th>Ticks fed on WTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>B0F457</td>
<td>Vitellogenin</td>
<td>9*</td>
<td>6</td>
</tr>
<tr>
<td>L7M551</td>
<td>Putative multicellular organism reproduction</td>
<td>2*</td>
<td>0</td>
</tr>
<tr>
<td>A8WAA7</td>
<td>Vitellogenin-2</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Q5EG54</td>
<td>Vitellogenin</td>
<td>11*</td>
<td>5</td>
</tr>
<tr>
<td>G9M4L6</td>
<td>Vitellogenin-B</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>I3VGB9</td>
<td>Vitellin-degrading cysteine endopeptidase</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Q56CZ1</td>
<td>Yolk cathepsin</td>
<td>3*</td>
<td>1</td>
</tr>
<tr>
<td>J9QJ79</td>
<td>Cathepsin L</td>
<td>1*</td>
<td>0</td>
</tr>
<tr>
<td>Q7YW74</td>
<td>Cathepsin L-like cysteine proteinase B</td>
<td>2*</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. Host proteins differentially represented in fed ticks. Host proteins identified in ticks fed on cattle and WTD were compared by $\chi^2$ test (*p<0.05).

<table>
<thead>
<tr>
<th>Uniprot Accession No.</th>
<th>Description</th>
<th>Ticks fed on cattle</th>
<th>Ticks fed on WTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>P21380</td>
<td>Hemoglobin subunit beta</td>
<td>10</td>
<td>28*</td>
</tr>
</tbody>
</table>
3.4. Host proteins affecting tick infestations

The analysis of tick proteome showed the presence of tick proteins involved in blood digestion and reproduction that are overrepresented in ticks fed on cattle when compared to ticks fed on WTD, correlating with the higher cattle tick infestations in cattle. However, host-tick co-evolution involves genetic traits of both the host and the vector. Several cattle proteins such as hemoglobin, immunoglobulin-associated proteins and albumin identified herein have been potentially associated with tick infestations [17-22]. However, the exact role of these proteins in tick infestations is unknown [17].
In our experiments, we found significantly higher host Fibrinogen levels in ticks fed on cattle when compared to ticks fed on WTD (Table 3). Fibrinogen is an essential component of blood coagulation [23], a process affected by tick feeding through the secretion of proteins that hydrolyze Fibrinogen and delay fibrin clot formation for successful blood pool maintenance and digestion [24, 25]. Additionally, Reck et al. [26] showed that Fibrinogen levels increase in response to *R. microplus* tick infestations in cattle. Therefore, higher Fibrinogen levels in ticks fed on cattle may reflect host response to tick infestations and an indicator of the lower tick infestations observed in WTD. Additionally, the possible role of host Fibrinogen in stimulating the production of Cathepsin and other peptidases in feeding ticks remains to be elucidated as a possible adaptation mechanism to circumvent host responses while promoting better blood digestion machinery that results in higher tick infestations [27].

### 3.5. Is it possible to target these tick proteins for the control of cattle tick infestations?

The results reported here support the possibility of using these proteins as vaccine candidates for the control of cattle tick infestations. Targeting tick proteins that are involved in blood digestion and reproduction will mimic results observed in ticks fed on WTD with lower tick infestations and reproduction that will ultimately result in control of tick populations.

Vaccines against cattle ticks became available in the early 1990’s as a cost-effective alternative for tick control that reduced the use of acaricides and the problems associated with them such as selection of acaricide-resistant ticks, environmental contamination and contamination of animal products with pesticide residues [28-31]. Currently, only two methods are available for the control of tick infestations in WTD, both involving the use of acaricides [4]. Recently, WTD vaccination with recombinant BM86 and Subolesin tick proteins proved their efficacy for the control of cattle tick infestations [11]. These results showed that deer produced an antibody response that correlated with the reduction in tick infestations similar to results in cattle vaccinated with these antigens [10, 11, 32]. Therefore, tick vaccines appear as an alternative for tick control in cattle and WTD with a possible impact on the transmission of tick-borne pathogens [33].

Recently, cattle vaccination with *R. microplus* Vitellin-degrading cysteine endopeptidase and Yolk cathepsin resulted in control of cattle tick infestations [34-36]. Additionally, preliminary experiments in sheep vaccinated with *R. microplus* Vitellin, a protein derived from the proteolytic processing of Vitellogenin, showed an effect on the control of cattle tick infestations [37]. These results support using these proteins for the control of cattle tick infestations and provide additional support for the results shown herein.
4. Conclusions

In summary, the results of the proteomics analysis showed that *R. microplus* ticks fed on cattle have overrepresented proteins involved in blood digestion and reproduction when compared to ticks fed on WTD, correlating with the higher tick numbers, weight and reproductive performance obtained in ticks fed on cattle. Higher Fibrinogen levels in ticks fed on cattle may reflect host response to tick infestations and an indicator of the lower tick infestations observed in WTD with possible implications in host-tick co-evolution. Finally, previous results from vaccination trials in cattle suggest the possibility of using these tick proteins for the control of cattle tick infestations and provided additional support for the results presented here. These results demonstrate the existence of host factors that affect tick feeding and reproduction that impact on vector evolution and ecology.

Acknowledgements

We thank Consuelo Almazán, Diana Carreón and Rodolfo Lagunes (Universidad Autónoma de Tamaulipas, Mexico) for supplying ticks. CBMSO proteomics facility (Centro de Biología Molecular Severo Ochoa, Madrid, Spain) is acknowledged for technical assistance. This research was supported by grants BFU2011-23896 and the EU FP7 ANTIGONE project number 278976. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases) within the FP7- PEOPLE – ITN programme (EU Grant No. 238511).

Authors’ contributions

JF designed the study. MP and MV performed proteomics analysis. LM-H performed the Western blot. JF wrote the manuscript and all authors edited the manuscript. All authors read and approved the final manuscript.

Supplementary material.

*Table S1. Tick proteins identified in R. microplus fed on cattle and WYD (FDR=0.05).*

*Table S2. Host proteins identified in R. microplus fed on cattle and WTD (FDR=0.01).*

References

4. Pound JM, George JE, Lohmeyer KH, Davey RB. Evidence for role of White-tailed deer (Artiodactyla:


22. Piper EK, Jackson LA, Bagnall NH, Kongsuwan KK, Lew AE, Jonsson NN. Gene expression in the skin of Bos taurus and Bos indicus cattle infested with the cattle tick, Rhipicephalus (Boophilus) microplus. Veterinary Immunology and Immunopathology 2008; 126: 110-119.


24. Francischetti IM. Platelet aggregation inhibitors from hematophagous animals. Toxiconology 2010; 56: 1130-44.


26. Reck J Jr, Berger M, Terra RM, Marks FS, da Silva Vaz I Jr, Guimarães JA et al. Systemic alterations...


33. de la Fuente J. Vaccines for vector control: Exciting possibilities for the future. The Veterinary Journal 2012; 194: 139-140.


Chapter V

Comparative proteomics for the characterization of tick vaccines protective mechanisms

Marina Popara, Margarita Villar, Lourdes Mateos-Hernández, Isabel G. Fernández de Mera, Anabel Marina, Mercedes del Valle, Consuelo Almazán, Ana Domingos, José de la Fuente. Lesser protein degradation machinery correlates with higher BM86 tick vaccine efficacy in *Rhipicephalus annulatus* when compared to *R. microplus*. Vaccine 2013; 31:4728–4735.

Chapter V a

Lesser protein degradation machinery correlates with higher BM86 tick vaccine efficacy in *Rhipicephalus annulatus* when compared to *R. microplus*

Abstract

Infestations with cattle ticks, *Rhipicephalus (Boophilus) microplus* and *R. annulatus*, economically impact cattle production in tropical and subtropical regions of the world. Vaccines containing the recombinant *R. microplus* BM86 gut antigen were developed and commercialized to induce an immunological protection in cattle against tick infestations. These vaccines demonstrated that tick control by vaccination is cost-effective, reduces environmental contamination and prevents the selection of drug resistant ticks that result from repeated acaricide applications. The protection elicited by BM86-containing vaccines against tick infestations is mediated by a collaborative action between the complement system and IgG antibodies. The efficacy of the vaccination with BM86 and other tick antigens is always higher for *R. annulatus* than against *R. microplus*, suggesting that tick genetic and/or physiological factors may affect tick vaccine efficacy. These factors may be related to BM86 protein levels or tick physiological processes such as feeding and protein degradation that could result in more efficient antibody-antigen interactions and vaccine efficacy. To test this hypothesis, we compared the proteome in *R. annulatus* and *R. microplus* female ticks after feeding on BM86-vaccinated and control cattle. The results showed that cattle proteins were under represented in *R. annulatus* when compared to *R. microplus*, suggesting that *R. annulatus* ticks ingested less blood, a difference that increased when feeding on vaccinated cattle, probably reflecting the effect of antibody-BM86 interaction on this process. The results also showed that tick protein degradation machinery was under represented in *R. annulatus* when compared to *R. microplus*. BM86 mRNA and protein levels were similar in both tick species, suggesting that lesser protease activity in *R. annulatus* results in more efficient antibody-antigen interactions and higher vaccine efficacy. These results have important implications for tick vaccine research, indicating that not only genetic differences, but also physiological factors may influence tick vaccine efficacy.

1. Introduction

The cattle ticks, *Rhipicephalus (Boophilus) microplus* and *R. annulatus*, are distributed in tropical and subtropical regions of the world [1, 2]. Cattle tick infestations economically impact cattle production by reducing weight gain and milk production, and by transmitting pathogens that cause babesiosis (*Babesia bovis* and *B. bigemina*) and anaplasmosis (*Anaplasma marginale*) [2, 3].

Acaricide application constitutes a major component of integrated tick control strategies [4]. However, use of acaricides has had limited efficacy in reducing tick infestations and is often
accompanied by serious drawbacks, including the selection of acaricide-resistant ticks, environmental contamination and contamination of milk and meat products with drug residues [4]. All of these issues reinforce the need for alternative approaches to control tick infestations and pathogen transmission, including the use of vaccines with tick antigens [5-7].

In the early 1990s, commercial vaccines containing the recombinant *R. microplus* BM86 gut antigen were developed and commercialized to induce an immunological protection in cattle against tick infestations [8-12]. These vaccines reduce the number of engorging female ticks, their weight and reproductive capacity. Thus the greatest vaccine effect was the reduction of larval infestations in subsequent generations. Vaccine controlled field trials in combination with acaricide treatments demonstrated that an integrated approach resulted in control of tick infestations while reducing the use of acaricides [11-13]. These trials demonstrated that tick control by vaccination is cost-effective, reduces environmental contamination and prevents the selection of drug resistant ticks that result from repeated acaricide applications [12]. In addition, these vaccines may also prevent or reduce transmission of pathogens by reducing tick populations [7, 10-13].

Initial experiments showed that tick vaccines protect against *R. microplus* infestations through a collaborative action between the complement proteins and IgG1 antibodies that most efficiently fix the complement in cattle [14-16]. Additional experiments with BM86 showed that antibodies alone can mediate the protective response to vaccination, but complement may enhance this effect [17].

Controlled immunization trials showed that *R. microplus* BM86-containing vaccines also protect against related tick species, *R. annulatus* and *R. decoloratus* [18-20]. Interestingly, the efficacy of the BM86 vaccine is higher against *R. annulatus* than against *R. microplus*, suggesting that tick genetic and/or physiological factors may affect the efficacy of tick vaccines [5, 18-28]. Furthermore, the fact that vaccination with both BA86, the BM86 ortholog in *R. annulatus*, and BM86 has a higher efficacy for *R. annulatus* than for *R. microplus* suggested the hypothesis that tick species-specific factors different from protein sequence diversity may condition the higher susceptibility of *R. annulatus* to vaccination. These factors may be related to BM86 levels or tick physiological processes such as feeding and protein degradation. For example, higher BM86 levels and/or lesser protease activity in *R. annulatus* would result in more efficient antibody-antigen interactions and vaccine efficacy.

Herein, we tested this hypothesis by comparing the proteome in *R. annulatus* and *R. microplus* female ticks after feeding on BM86-vaccinated and control cattle.
1. Materials and methods

1.1. Tick collection

Adult female *R. annulatus* (Mercedes strain, Texas, USA) and *R. microplus* (Susceptible Media Joya strain, CENAPA, Mexico) ticks were collected from the same animal in a previously reported trial after feeding on cattle vaccinated with BM86 (Gavac, Revetmex, Mexico City, Mexico) and control cattle injected with adjuvant/saline alone [26]. In this trial, the efficacy of BM86 was higher on *R. annulatus* (99.6%) than on *R. microplus* (85.2%) and tick weight was lower in *R. annulatus* (141±59 mg) than in *R. microplus* (253±9 mg) fed on BM86 vaccinated cattle [26]. Ticks were stored at -20 °C in 70% ethanol until used for protein and RNA extraction.

1.2. Protein extraction and proteomics analysis

Eight ticks from each species were dissected, cuticle removed, pulverized in liquid nitrogen and homogenized with a glass homogenizer (10 strokes) in 1 ml buffer (10mM phosphate buffer saline (PBS), pH 7.4) supplemented with 1% SDS and complete mini protease inhibitor cocktail (Roche, Basel, Switzerland) per 50 µg sample. Samples were sonicated for 1 min in an ultrasonic cooled bath followed by 10 sec of vortex. After 3 cycles of sonication-vortex, the homogenates were centrifuged at 200xg for 5 min at room temperature to remove cellular debris. The supernatants were collected and protein concentration was determined using the BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) using BSA as standard. Protein extracts (200 µg from each sample) were precipitated following the methanol/chloroform procedure [29], resuspended in 100 µl Laemmli sample buffer and applied onto 1.2-cm wide wells on a 12% SDS-PAGE gel. The electrophoretic run was stopped as soon as the front entered 3 mm into the resolving gel, so that the whole proteome became concentrated in the stacking/resolving gel interface. The unseparated protein bands were visualized by staining with GelCode Blue Stain Reagent (Thermo Scientific), excised, cut into 2x2 mm cubes and digested overnight at 37°C with 60 ng/µl sequencing grade trypsin (Promega, Madison, WI, USA) at 5:1 protein:trypsin (w/w) ratio in 50 mM ammonium bicarbonate, pH 8.8 containing 10% (v/v) acetonitrile [30]. The resulting tryptic peptides from each band were extracted by 30min-incubation in 12 mM ammonium bicarbonate, pH 8.8. Trifluoroacetic acid was added to a final concentration of 1% and the peptides were finally desalted onto OMIX Pipette tips C18(Agilent Technologies, Santa Clara, CA, USA), dried-down and stored at -20°C until mass spectrometry analysis.

The desalted protein digest was resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS using an Agilent 1100 LC system (Agilent Technologies) coupled to a linear ion trap
LTQ-Velos mass spectrometer (Thermo Scientific). The peptides were separated by reverse phase chromatography using a 0.18 mm x 150 mm Bio-Basic C18 RP column (Thermo Scientific) at 1.8 μl/min. Peptides were eluted using a 120-min gradient from 5 to 40% solvent B in solvent A (Solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid, 80% acetonitrile in water). ESI ionization was done using a microspray metal needle kit (Thermo Scientific) interface. Peptides were detected in survey scans from 400 to 1600 amu (1 μscan), followed by fifteen data dependent MS/MS scans (Top 15), using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion applied during 30 sec periods.

1.3. Proteomics data analysis

The MS/MS raw files were searched against the Ixodida (40,849 entries in June 2013) and Ruminantia (66,519 entries in June 2013) Uniprot databases with the following constraints: tryptic cleavage after Arg and Lys, up to two missed cleavage sites, and tolerances of 0.5 Da for precursor ions and 0.8 Da for MS/MS fragment ions. Due to the limited number of tick proteins included in the database, a false discovery rate (FDR) ≤ 0.05 was used for tick proteins while only host proteins with FDR ≤ 0.01 were considered. Functional data for each protein was obtained from Uniprot and included gene ontology (GO) biological process (BP) and molecular function (MF) annotations. The comparison between under represented and over represented proteins in all BPs and for the number of peptides/protein between tick and cattle host proteins were done using a Student’s t-test with unequal variance (p=0.05). Differential protein representation for individual proteins between different samples was determined using χ² test statistics with Bonferroni correction in the IDEG6 software (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) (p=0.05).

1.4. Western blot analysis of BM86 and Cathepsin L

Total proteins (150 μg from each sample) were methanol/chloroform precipitated, resuspended in Laemmli sample buffer and separated on a 7.5% (for BM86) or 15% (for Cathepsin L) SDS-PAGE gel under reducing conditions. After electrophoresis, proteins were transferred to nitrocellulose membranes (Bio-Rad, Hercules, CA, USA), blocked with SuperBlock blocking buffer in TBS (Thermo Scientific) and incubated overnight at 4°C with rabbit polyclonal anti-Bm86 (dilution 1:1000; [31]) or anti-Cathepsin L (Mature region No. pab0213-0; Covalab, Villeurbanne, France) antibodies. To detect the antigen-bound antibody, membranes were incubated with goat anti-rabbit IgG conjugated with horseradish peroxidase (dilution 1:10,000; Sigma-Aldrich, St. Louis, MO, USA). Immunoreactive proteins were detected by
chemoluminescence using the SuperSignal West Pico chemoluminescent substrate (Thermo Scientific), visualized with an ImageQuant 350 Digital Imaging System (GE Healthcare, Pittsburgh, PA, USA), quantified using the ImageQuant TL 7.0 software (GE Healthcare) and normalized against total proteins. Normalized protein levels (N=2) were compared between samples by $\chi^2$ test (p=0.05). Positive controls (C+) corresponded to recombinant R. microplus BM86 expressed in Pichia pastoris (5 µg) and pig tonsil cytoplasmatic protein fraction (150 µg) for BM86 and Cathepsin L Western blots, respectively.

1.5. RNA extraction and RT-PCR

Five ticks were dissected to collect the gut from individual ticks for total RNA extraction using the RNeasy kit (Qiagen Inc., Valencia, CA, USA). The bm86 mRNA levels were analyzed by RT-PCR using oligonucleotide primers bm86_F: 5´-CCATTTGCTCTGACTTCGG-3´ and bm86_R: 5´-CAGTTGCGTAGTCATTT-3´ and the Access RT-PCR system (Promega, Madison, WI, USA) with PCR conditions of 5 min at 95°C and 40 cycles of 30 sec at 95°C, 30 sec at 60°C and 30 sec at 68°C using 5 ng RNA in a iQ5 thermal cycler (Bio-Rad, Hercules, CA, USA). Amplicons (270 bp) were visualized by agarose gel electrophoresis and quantitated by densitometric analysis of methylene blue-stained gels using ImageJ 1.44p (National institute of Health, USA). The amount of bm86 amplicons was normalized against total DNA and compared between groups by $\chi^2$ test (p=0.05).

1.6. Protease activity assay

Tick gut tissues were suspended in ice-cold extraction buffer (20 mM Bis-Tris-HCl, pH 6.0) containing 0.05% Tween20, homogenized with a syringe and finally subjected to two freeze-thaw cycles, and centrifuged at 12,000xg for 10 min at 4°C. The resulting supernatant was collected and used for protease activity assay. Fluorometric assays were conducted with 20 µl of protein extract in 200 µl (100 mM PBS, 5 mM DTT,pH 7.2). Reactions were initiated by the addition of 50 µM fluorogenic substrate (Z-Leu- Leu-Arg-AMC, Bachem AG, Bubendorf, Switzerland) and activity was monitored (excitation 355 nm; emission 460 nm) for 30 min at RT on a Fluorescence Microplate Reader FLx800 (BioTek Instruments Inc., Winooski, VT,USA). For all assays, saturated substrate concentration was used in order to obtain linear fluorescence curves [32].Protease activity was normalized against total protein content to obtain protease specific activity (fu/mg).
2. Results and discussion

2.1. Global proteomics results

The proteomics analysis resulted in the identification of 314 and 339 tick and cattle proteins, respectively in *R. annulatus* and *R. microplus* fed on BM86 vaccinated and control cattle (Tables S1 and S2). The number of proteins identified in *R. microplus* was higher than in *R. annulatus* for both tick (Fig. 1A) and cattle (Fig. 1B) proteins. Although the protein database is bigger for *R. microplus* (452 proteins in the Genbank on June 2013) than for *R. annulatus* (38 proteins), this result pointed at the first difference between these two closely related tick species because sequence identity is high between them [33,34,22] and thus protein identification should not be biased by differences in the database. Furthermore, the number of identified cattle proteins was also lower in *R. annulatus* and not biased by database size (Fig. 1B). The difference in protein identification between tick species increased with vaccination, suggesting a correlation between the number of tick and cattle proteins identified in *R. annulatus* and vaccination (Fig. 1C). Interestingly, the number of cattle proteins identified in ticks fed on vaccinated cattle was lower than in ticks fed on control cattle (10% and 21% reduction for *R. microplus* and *R. annulatus*, respectively; Fig. 1D), probably reflecting the effect of the vaccine on tick feeding [26]. In fact, in the vaccine trial where ticks were collected, tick weight was reduced by 15% and 54% for *R. microplus* and *R. annulatus*, respectively [26]. A total of 56 tick proteins were identified in all samples, but some proteins were identified in one sample only (Fig. 1E). The number of peptides used to identify each protein was higher for cattle than for tick proteins, reflecting differences between protein databases (Fig. 1F).
2.2. Characterization of tick proteins

The 314 tick proteins identified after proteomics analysis of *R. annulatus* and *R. microplus* fed on vaccinated and control cattle (Table S1) were annotated and classified into 13 BPs (Fig. 2A). Of them, the most abundant BPs corresponded to unknown processes, enzymatic processes, protein metabolism and protein degradation (Fig. 2A). The comparison of differentially represented tick proteins in all BPs between ticks fed on vaccinated and control cattle showed no significant differences between *R. annulatus* and *R. microplus* (Fig. 2B). The comparisons between under represented and over represented proteins in all BPs showed that under represented proteins were
more abundant in *R. annulatus* than in *R. microplus* ticks fed on vaccinated cattle (*p*=0.02) (Fig. 2C), but not between ticks fed on control cattle (Fig. 2D).

![Figure 2. Tick proteome.](image)

(A) Tick proteins identified in *R. annulatus* and *R. microplus* were functionally annotated and grouped according to their biological process. (B) Number of differentially represented tick proteins in *R. annulatus* and *R. microplus* fed on vaccinated vs. control cattle. (C) Number of differentially represented tick proteins in *R. annulatus* vs. *R. microplus* fed on vaccinated cattle. (D) Number of differentially represented tick proteins in *R. annulatus* vs. *R. microplus* fed on control cattle.

The protein degradation BP was selected for further analysis due to its relevance to test our hypothesis. Proteins in this BP (N=35) were classified into three MFs, protein degradation, inhibition of protein degradation, and regulation of protein degradation (Fig. 3A and Table 1). The results showed that proteins with protein degradation activity and involved in regulation of protein degradation were under represented while proteins involved in the inhibition of protein degradation were over represented in *R. annulatus* when compared to *R. microplus* (Fig. 3A), thus suggesting a lesser protein degradation machinery in *R. annulatus* fed on vaccinated cattle. Some of the over represented proteins in the protein degradation BP are increased at the end of feeding.
when these ticks were collected, thus likely reflecting the dynamics of the hemoglobinolytic system [35].

Figure 3. Protein degradation machinery in ticks and cattle and putative host proteins involved in protective response to vaccination. (A) Number of tick proteins in the protein degradation BP in *R. annulatus* vs. *R. microplus* fed on vaccinated cattle. Proteins were grouped according to their MF. (B) Number of cattle proteins in the protein degradation BP in *R. annulatus* vs. *R. microplus* fed on vaccinated cattle. Proteins were grouped according to their MF. (C) Number of cattle proteins in the protein degradation BP and putatively involved in protective response to vaccination (complement system and immunoglobulins) in *R. annulatus* vs. *R. microplus* fed on vaccinated cattle. (D) Cattle hemoglobin peptides identified in *R. annulatus* and *R. microplus* fed on vaccinated and control cattle.

2.3. Characterization of cattle host proteins

Of the 339 cattle host proteins identified after proteomics analysis of *R. annulatus* and *R. microplus* fed on vaccinated and control cattle (Table S2), 25 proteins were classified into the protein degradation BP and 16 as putatively involved in protective response to vaccination, considering complement system and immunoglobulins (Ig) [14-16] (Table 2). Different to results with tick proteins, cattle host proteins with protein degradation activity showed no differences
while proteins involved in the inhibition and regulation of protein degradation were under represented in *R. annulatus* fed on vaccinated cattle when compared to *R. microplus* (Fig. 3B). However, the most active protein degradation inhibitor, alpha-2 macroglobulin, was found in *R. annulatus* under represented for Q7SIH1 but over represented for L8IE16 (Table 2), suggesting that host proteins had little contribution to tick protein degradation machinery. In general, host proteins involved in protein degradation and vaccine protection were under represented or showed no differences in *R. annulatus* fed on vaccinated cattle when compared to *R. microplus* (Fig. 3C). A closer look at these proteins evidenced that both classical (C1, C3, C4) and alternative (C3) complement pathways were under represented in *R. annulatus*, correlating with the lower protein degradation machinery in this species (Table 2). For Ig, lambda and gamma (IgG) chains were under represented in *R. annulatus* fed on vaccinated cattle while an Ig J (IgA/IgM) chain was over represented in this species when compared to *R. microplus* (Table 2). These results suggested that proteins that may be involved in the cattle protective response to BM86 vaccination were under represented in *R. annulatus*. However, one the most abundant host proteins, hemoglobin (represented by 20 entries; Table S2), was also under represented in *R. annulatus* with an additional decrease in ticks fed on vaccinated cattle (Fig. 3D), again suggesting an effect of vaccination on *R. annulatus* tick feeding.

### 2.4. Characterization of BM86 levels in *R. annulatus* and *R. microplus*

A direct correlation between antibody titers and vaccine efficacy has been demonstrated for BM86-based vaccines [13]. Therefore, one possibility is that BM86 protein levels are higher in *R. annulatus* than in *R. microplus*, thus increasing the probability for antibody-antigen interactions that mediate tick vaccine efficacy [13]. However, despite differences in tick weight between *R. annulatus* and *R. microplus* fed on vaccinated cattle [26], the characterization of BM86 protein (Fig. 4A) and *bm86* mRNA (Fig. 4B) levels demonstrated that they are similar in both tick species after feeding on vaccinated and control cattle (p>0.05).

### 2.5. Tick protein degradation machinery and vaccine efficacy

To explain why BM86 vaccine efficacy is always higher against *R. annulatus* than against *R. microplus* [18-20] and the fact that this result is similar for other antigens such as BA86 [26], Subolesin [27] and Ferritin2 [28], we used a proteomics approach to test the hypothesis that tick physiological processes such as protein degradation determine the higher susceptibility of *R. annulatus* to vaccination. Protein degradation is a mechanism for ticks to counteract host response to infestation and pathogen infection and for digestion of blood meal [35-39]. Therefore, lesser
protein degradation machinery in *R. annulatus* may result in lower protection against host responses while higher levels of protease inhibitors may represent an evolutionary advantage to aid in protection against pathogen infection [35-39].

The results of proteomics analysis suggested that *R. annulatus* ticks ingested less blood than *R. microplus*, a difference that increased when feeding on vaccinated cattle, probably reflecting the effect of antibody-BM86 interaction on this process. This fact explains why cattle proteins putatively involved in protective response to vaccination were under represented in *R. annulatus* when compared to *R. microplus* fed on vaccinated cattle. Proteomics analysis also showed that tick protein degradation machinery was under represented in *R. annulatus* when compared to *R. microplus*, a result that was corroborated by Western blot of Cathepsin L protease (Fig. 4C) and total protease activity analysis (Fig. 4D) in ticks. Considering that BM86 levels were similar in *R. annulatus* and *R. microplus*, lesser protease activity in *R. annulatus* probably results in more efficient antibody-antigen interactions and higher vaccine efficacy (Fig. 4E).

![Figure 4](image-url)
arbitrary units) were compared between samples by $\chi^2$ test (p>0.05; N=2). (B) bm86 mRNA levels were determined by PCR in *R. annulatus* and *R. microplus*. The amount of bm86 amplicons was normalized against total DNA, represented as Ave+S.D. in arbitrary units and compared between groups by $\chi^2$ test (p>0.05; N=5). An agarose gel is shown with representative samples for each group. (C) Cathepsin L protein levels were determined by Western blot in *R. annulatus* and *R. microplus* fed on BM86-vaccinated and control cattle, quantified and normalized against total proteins. Positive control (C+) corresponded to pig tonsil cytoplasmatic protein fraction (150 µg). Normalized protein levels (Ave+S.D. arbitrary units) were compared between samples by $\chi^2$ test (*p<0.001; N=2). (D) Total protease activity was determined by fluorometric assays conducted with 20 µl of protein extract in *R. annulatus* and *R. microplus* fed on BM86-vaccinated and control cattle. Protease activity was normalized against total protein content to obtain protease specific activity (fu/mg). (E) Schematic representation of the relationship between protein degradation machinery and vaccine efficacy in cattle vaccinated with BM86 to protect against cattle tick infestations. Abbreviations: R.a.c, *R. annulatus* fed on control cattle; R.a.v, *R. annulatus* fed on BM86-vaccinated cattle; R.m.c, *R. microplus* fed on control cattle; R.m.v, *R. microplus* fed on BM86-vaccinated cattle.

Table 1. Tick proteins in the protein degradation biological process.

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Description</th>
<th>Putative function in protein degradation</th>
<th>Protein levels in ticks fed on vaccinated cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>K7WSB5</td>
<td>Serpin 1</td>
<td>Inhibition</td>
<td>ND</td>
</tr>
<tr>
<td>P81162</td>
<td>Protein degradation inhibitor carrapatin</td>
<td>Inhibition</td>
<td>ND</td>
</tr>
<tr>
<td>P83606</td>
<td>Kunitz-type serine protein degradation inhibitor 6</td>
<td>Inhibition</td>
<td>Over represented*</td>
</tr>
<tr>
<td>Q7Z021</td>
<td>Serpin</td>
<td>Inhibition</td>
<td>Over represented</td>
</tr>
<tr>
<td>B7P5E3</td>
<td>F-box/leucine rich repeat protein, putative</td>
<td>Regulation</td>
<td>ND</td>
</tr>
<tr>
<td>L7LX08</td>
<td>Putative molecular chaperone mortalin</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>F0J9K1</td>
<td>Ubiquitin</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L7M227</td>
<td>Putative chaperonin</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L7M5Q1</td>
<td>Putative e3 ubiquitin ligase</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L7M7J5</td>
<td>Putative 20s proteasome regulatory subunit beta type psmb1/pre7</td>
<td>Regulation</td>
<td>ND</td>
</tr>
<tr>
<td>L7MAE4</td>
<td>Putative chaperonin protein</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
<td>Regulation</td>
<td>Degradation</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------</td>
<td>------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>L7M4U4</td>
<td>Putative ubiquitin protein ligase</td>
<td>Regulation</td>
<td>ND</td>
</tr>
<tr>
<td>L7M4S5</td>
<td>Putative 20s proteasome regulatory subunit alpha type psma1/pre5</td>
<td>Regulation</td>
<td>Over represented</td>
</tr>
<tr>
<td>L7M324</td>
<td>Putative chaperonin</td>
<td>Regulation</td>
<td>ND</td>
</tr>
<tr>
<td>L7MAC0</td>
<td>Putative chaperonin</td>
<td>Regulation</td>
<td>ND</td>
</tr>
<tr>
<td>B7P4H2</td>
<td>Regulator of the ubiquitin pathway</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q9GYX7</td>
<td>Heme-binding aspartic proteinase</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q56CZ1</td>
<td>Yolk cathepsin</td>
<td>Degradation</td>
<td>Over represented*</td>
</tr>
<tr>
<td>I3VGB9</td>
<td>Vitellin-degrading cysteine endopeptidase</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>C3UTE0</td>
<td>Aspartic proteinase</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>B7SP35</td>
<td>Putative serine protein degradation protein</td>
<td>Degradation</td>
<td>Over represented</td>
</tr>
<tr>
<td>G3MMH6</td>
<td>Proteasome subunit alpha type</td>
<td>Degradation</td>
<td>Over represented</td>
</tr>
<tr>
<td>B7Q5F6</td>
<td>Proteasome subunit alpha type</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>L7M840</td>
<td>Putative multifunctional chaperone 14-3-3 family</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>B7SP51</td>
<td>Putative mitochondrial processing peptidase beta-subunit</td>
<td>Degradation</td>
<td>Over represented</td>
</tr>
<tr>
<td>J9QJ79</td>
<td>Cathepsin L</td>
<td>Degradation</td>
<td>Over represented</td>
</tr>
<tr>
<td>Q7YW74</td>
<td>Cathepsin L-like cysteine proteinase B</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q86GZ4</td>
<td>Midgut cysteine proteinase 3</td>
<td>Degradation</td>
<td>Over represented</td>
</tr>
<tr>
<td>L7MEZ0</td>
<td>Putative chaperonin</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L7M8X6</td>
<td>Putative tick serine protein degradation</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L7LZT3</td>
<td>Putative xaa-pro dipeptidase</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>L7M2J0</td>
<td>Putative aminopeptidase of the m17 family</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>Accession No.</td>
<td>Description</td>
<td>Putative function in protein degradation and vaccine protection</td>
<td>Protein levels in ticks fed on vaccinated cattle</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>L7M002</td>
<td>Putative prolyl 4-hydroxylase subunit alpha-1</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L7M8W7</td>
<td>Putative serine carboxypeptidase</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>J9QSA7</td>
<td>Calreticulin</td>
<td>Degradation</td>
<td>Under represented</td>
</tr>
</tbody>
</table>

*Protein levels in *R. annulatus* when compared to *R. microplus* fed on vaccinated cattle. Only proteins with more than one peptide in at least one of the samples were considered. Abbreviation: ND, no difference between samples. Differential protein representation for individual proteins between different samples was determined using $\chi^2$ test statistics with Bonferroni correction in the IDEG6 software (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) (*p<0.05*).

Table 2. Cattle host proteins involved in protein degradation and the protective response to vaccination with BM86.

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Description</th>
<th>Putative function in protein degradation and vaccine protection</th>
<th>Protein levels in ticks fed on vaccinated cattle</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2I7N1</td>
<td>Serpin A3-5</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>A5PJ69</td>
<td>Serpin A10</td>
<td>Inhibition</td>
<td>ND</td>
</tr>
<tr>
<td>A6QPP2</td>
<td>Serpin D1</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>F1MSZ6</td>
<td>Serpin</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>B6UV62</td>
<td>Serpin F1</td>
<td>Inhibition</td>
<td>ND</td>
</tr>
<tr>
<td>Q9TTE1</td>
<td>Serpin S1Control-1</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>L8HW35</td>
<td>Serpin S1Control-7</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q1JPB0</td>
<td>Luekocyte elastase inhibitor</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q32T06</td>
<td>Endopin 2C</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q9TS74</td>
<td>Pancreatin elastase inhibitor</td>
<td>Inhibition</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q7M312</td>
<td>Isoaprotinin G2</td>
<td>Inhibition</td>
<td>ND</td>
</tr>
<tr>
<td>P00974</td>
<td>Pancreatic trypsin</td>
<td>Inhibition</td>
<td>ND</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
<td>Regulation</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Q7SIH1</td>
<td>Alpha-2-macroglobulin inhibitor</td>
<td>Inhibition</td>
<td>Under represented*</td>
</tr>
<tr>
<td>L8IE16</td>
<td>Alpha-2-macroglobulin inhibitor</td>
<td>Inhibition</td>
<td>Over represented*</td>
</tr>
<tr>
<td>L817D0</td>
<td>T-complex protein 1 subunit eta</td>
<td>Regulation</td>
<td>ND</td>
</tr>
<tr>
<td>A7E3Q2</td>
<td>Heat shock 70 kDa protein 1A</td>
<td>Regulation</td>
<td>Over represented*</td>
</tr>
<tr>
<td>C7EDP1</td>
<td>Heat shock 70 kDa protein 5 isoform 1</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>C7EDP2</td>
<td>Heat shock 70 kDa protein 5 isoform 2</td>
<td>Regulation</td>
<td>Under represented</td>
</tr>
<tr>
<td>L81DB8</td>
<td>Polyubiquitin-C</td>
<td>Regulation</td>
<td>Under represented*</td>
</tr>
<tr>
<td>G5E589</td>
<td>Proteasome subunit beta type</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>G5E5C3</td>
<td>Proteasome subunit alpha type</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>L812N5</td>
<td>Disintegrin and metalloproteinase domain-containing protein 23</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>L8IWM2</td>
<td>Proteasome subunit alpha type-5</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>L8J0V2</td>
<td>Proteasome subunit alpha type-4</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>P25975</td>
<td>Cathepsin L1</td>
<td>Degradation</td>
<td>ND</td>
</tr>
<tr>
<td>L8IG77</td>
<td>Complement component C1q</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Vaccine protection</td>
<td>Representation</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------</td>
<td>-------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>A5YBU9</td>
<td>Complement factor B</td>
<td>Vaccine protection</td>
<td>Under represented</td>
</tr>
<tr>
<td>P01030</td>
<td>Complement component C4</td>
<td>Vaccine protection</td>
<td>Under represented*</td>
</tr>
<tr>
<td>Q29439</td>
<td>Complement component C4</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>Q2UVX4</td>
<td>Complement C3</td>
<td>Vaccine protection</td>
<td>Under represented*</td>
</tr>
<tr>
<td>Q3SYT3</td>
<td>Complement component C1</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>Q5E9E3</td>
<td>Complement component C1q subunit A</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>L8IAF4</td>
<td>Ig lambda-1 chain C region</td>
<td>Vaccine protection</td>
<td>Under represented</td>
</tr>
<tr>
<td>L8HKR7</td>
<td>IgG gamma-3 chain C region</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>L8HP94</td>
<td>IgA alpha-1 chain C region</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>L8HPF5</td>
<td>IgG gamma-1 chain C region</td>
<td>Vaccine protection</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q1RMN8</td>
<td>Ig lambda light chain</td>
<td>Vaccine protection</td>
<td>Under represented</td>
</tr>
<tr>
<td>A5PK49</td>
<td>Ig-like</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>B0JYP6</td>
<td>Ig-like</td>
<td>Vaccine protection</td>
<td>Under represented</td>
</tr>
<tr>
<td>Q3T101</td>
<td>Ig-like</td>
<td>Vaccine protection</td>
<td>ND</td>
</tr>
<tr>
<td>Q3SYR8</td>
<td>Ig J chain</td>
<td>Vaccine protection</td>
<td>Over represented</td>
</tr>
</tbody>
</table>

*Protein levels in *R. annulatus* when compared to *R. microplus* fed on vaccinated cattle. Only proteins with more than one peptide in at least one of the samples were considered. Abbreviation: ND, no difference between samples. Differential protein representation for individual proteins between different samples was determined using $\chi^2$ test statistics with Bonferroni correction in the IDEG6 software ([http://telethon.bio.unipd.it/biinfo/IDEG6_form/](http://telethon.bio.unipd.it/biinfo/IDEG6_form/)) (*p<0.05).*

112
3. Conclusions
In summary, these results support our hypothesis that lesser protein degradation machinery in *R. annulatus* may contribute to higher vaccine efficacy in this species when compared to *R. microplus*. These results have important implications for tick vaccine research, suggesting that not only genetic differences, but also physiological differences between tick species and possibly between geographic strains of the same species may affect tick vaccine efficacy.

Acknowledgements
We thank Rodolfo Laguna (Universidad Autónoma de Tamaulipas, Mexico) and Juan A. Moreno-Cid (IREC, Spain) for technical assistance. This work was supported by project BFU2011-23896. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases) within the FP7- PEOPLE – ITN programme (EU Grant No. 238511).

Supplementary material.
*Table S1. Tick proteins identified in ticks feeding on vaccinated and control cattle (FDR=0.05).*
*Table S2. Cattle host proteins identified in ticks feeding on vaccinated and control cattle (FDR=0.01).*

References
7. de la Fuente, J. Vaccines for vector control: Exciting possibilities for the future. The Veterinary Journal2012; 194: 139-140.

34. Gunn SJ, Hilburn LR, Burbach BS. Homology within the X chromosomes of Boophilus microplus (Canestrini) and B. annulatus (Say). Journal of Heredity 1993; 84: 232-5.


Comparative proteomics of replete female ticks highlights molecular mechanisms associated with BM86 and Subolesin vaccine protection

Margarita Villar, Marina Popara, Nieves Ayllón, Lourdes Mateos-Hernández, José de la Fuente.

Comparative proteomics of replete female ticks highlights molecular mechanisms associated with BM86 and Subolesin vaccine protection. In preparation.
Abstract

Cattle ticks, Rhipicephalus (Boophilus) microplus, are a serious threat to animal health and production in many regions of the world. Tick vaccines are a cost-effective alternative for tick control. Although the antibody response has been well characterized in hosts immunized with the tick protective antigens BM86 and Subolesin (SUB), little information is available on the molecular mechanisms affected in ticks fed on vaccinated hosts that can be associated with vaccine protection. Herein, we compared the proteome of R. microplus female ticks fed on animals vaccinated with BM86, SUB and controls. The results of the proteomics analysis showed that although both vaccines reduce tick feeding and reproduction, they act through different protective mechanisms. The main protective effect of the BM86 vaccine is the effect on BM86 function and levels, which probably alters tick gut structure and function and consequently blood digestion and reproduction. For the SUB vaccine, the results showed that vaccination reduces protein levels and affects its function as a transcriptional regulator of genes involved in several biological processes important for tick feeding and reproduction. These results have important implications for tick vaccine research, supporting the study of the molecular mechanisms associated with vaccine protection to combine antigens that act through different protective mechanisms to increase vaccine efficacy.

1. Introduction

Cattle ticks, Rhipicephalus (Boophilus) microplus, are a serious threat to animal health and production in many regions of the world where they transmit pathogens that cause babesiosis (Babesia bovis and B. bigemina) and anaplasmosis (Anaplasma marginale) [1, 2]. Wildlife and particularly white tailed deer (WTD), Odocoileus virginianus, could maintain cattle tick populations alone or in combination with cattle [2-5].

Vaccines containing the R. microplus BM86 gut antigen became commercially available in the early 1990s as a cost-effective alternative for cattle tick control that reduced the use of acaricides and the problems associated with them such as selection of acaricide-resistant ticks, environmental contamination and contamination of animal products with pesticide residues [6]. Tick Subolesin (SUB) is a candidate protective antigen that has been shown to control tick infestations and infection with tick-borne pathogens of cattle such as A. marginale and B. bigemina [7].

Recently, WTD were vaccinated with recombinant BM86 or SUB and experimentally infested with R. microplus larvae to characterize the effect of the vaccines on tick infestations [8]. These
experiments showed that vaccine efficacy was similar between BM86 (E=76%) and SUB (E=83%) for the control of R. microplus infestations in WTD and suggested the possibility to use this model to gain more information on the protective mechanisms elicited by these vaccines. Although the antibody response has been well characterized in hosts immunized with tick antigens such as BM86 and SUB [6-11], little information is available on the molecular mechanisms affected in ticks fed on vaccinated hosts that can be associated with vaccine protection [12, 13]. Our hypothesis is that the characterization of the proteome in ticks after feeding on vaccinated and control hosts will provide information on the molecular mechanisms that are associated with the effect of vaccination on tick feeding and reproduction. Herein, we addressed this hypothesis by comparing the proteome of R. microplus ticks fed on WTD vaccinated with BM86, SUB and controls.

2. Materials and methods

2.1. Tick collection

Adult female R. microplus (Susceptible Media Joya strain, CENAPA, Mexico) ticks were collected in a previously reported trial after completing feeding on WTD [8]. Tick infestation results were similar for ticks fed on WTD vaccinated with BM86 (44%, 14%, 28% and 41% reduction in tick numbers, weight, oviposition and fertility, respectively) or SUB (56%, 11%, 21% and 51% reduction in tick numbers, weight, oviposition and fertility, respectively) when compared to ticks fed on control animals injected with adjuvant/saline alone [8]. Replete female ticks were collected after completing feeding on WTD and stored at -20°C in 70% ethanol until used for protein or DNA/RNA extraction.

2.2. Protein extraction

Eight ticks from each group were dissected, cuticle removed, pulverized in liquid nitrogen and homogenized with a glass homogenizer (10 strokes) in 1 ml buffer (10 mM phosphate buffer saline (PBS), pH 7.4) supplemented with 1% SDS and complete mini protease inhibitor cocktail (Roche, Basel, Switzerland) per 50 µg sample. Samples were sonicated for 1 min in an ultrasonic cooled bath followed by 10 sec of vortex. After 3 cycles of sonication-vortex, the homogenates were centrifuged at 200xg for 5 min at room temperature to remove cellular debris. The supernatants were collected and protein concentration was determined using the BCA Protein Assay (Thermo Scientific, San Jose, CA, USA) using BSA as standard.
2.3. Proteomics

Proteomics analysis was conducted as previously reported [13]. Protein extracts (200 µg from each sample) were precipitated using methanol/chloroform, resuspended in 100 µl Laemmli sample buffer and applied onto 1.2-cm wide wells on a 12% SDS-PAGE gel. The electrophoretic run was stopped as soon as the front entered 3 mm into the resolving gel, so that the whole proteome became concentrated in the stacking/resolving gel interface. The unseparated protein bands were visualized by staining with GelCode Blue Stain Reagent (Thermo Scientific), excised, cut into 2x2 mm cubes and digested overnight at 37°C with 60 ng/µl sequencing grade trypsin (Promega, Madison, WI, USA) at 5:1 protein:trypsin (w/w) ratio in 50 mM ammonium bicarbonate, pH 8.8 containing 10% (v/v) acetonitrile. The resulting tryptic peptides from each band were extracted by 30 min-incubation in 12 mM ammonium bicarbonate, pH 8.8. Trifluoroacetic acid was added to a final concentration of 1% and the peptides were finally desalted onto OMIX Pipette tips C_{18} (Agilent Technologies, Santa Clara, CA, USA), dried-down and stored at -20°C until mass spectrometry analysis. The desalted protein digest was resuspended in 0.1% formic acid and analyzed by RP-LC-MS/MS using an Agilent 1100 LC system (Agilent Technologies) coupled to a linear ion trap LTQ-Velos mass spectrometer (Thermo Scientific). The peptides were separated by reverse phase chromatography using a 0.18 mm x 150 mm Bio-Basic C18 RP column (Thermo Scientific) at 1.8 µl/min. Peptides were eluted using a 120-min gradient from 5 to 40% solvent B in solvent A (Solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid, 80% acetonitrile in water). ESI ionization was done using a microspray metal needle kit (Thermo Scientific) interface. Peptides were detected in survey scans from 400 to 1600 amu (1 µscan), followed by fifteen data dependent MS/MS scans (Top 15), using an isolation width of 2 mass-to-charge ratio units, normalized collision energy of 35%, and dynamic exclusion applied during 30 sec periods.

2.4. Proteomics data analysis

The MS/MS raw files were searched against the Ixodida (40,849 entries in June 2013) and Ruminantia (66,519 entries in June 2013) Uniprot databases with the following constraints: tryptic cleavage after Arg and Lys, up to two missed cleavage sites, and tolerances of 1.0 Da for precursor ions and 0.8 Da for MS/MS fragment ions. Due to the limited number of tick proteins included in the database, a false discovery rate (FDR) \( \leq 0.05 \) was used for tick proteins while only host proteins with FDR \( \leq 0.01 \) were considered. Differential protein representation between different samples was determined by Student’s t-test or Chi\(^2\)-test (p=0.05) depending on the
number of proteins included using the number of peptides per protein in the sample. Functional data for each protein were obtained from Uniprot and included gene ontology (GO) annotations for Biological Process (BP).

2.5. Western blot analysis of SUB and BM86

Total proteins (150 µg from each sample) were methanol/chloroform precipitated, resuspended in Laemmli sample buffer and separated on a 7.5% (for BM86) or 15% (for SUB) SDS-PAGE gel under reducing conditions. After electrophoresis, proteins were transferred to nitrocellulose membranes (Bio-Rad, Hercules, CA, USA), blocked with SuperBlock blocking buffer in TBS (Thermo Scientific) and incubated overnight at 4ºC with rabbit polyclonal anti-BM86 (dilution 1:1000; [14]) or anti-SUB (dilution 1:1000; [15]) antibodies. To detect the antigen-bound antibody, membranes were incubated with goat anti-rabbit IgG conjugated with horseradish peroxidase (dilution 1:10,000; Sigma-Aldrich, St. Louis, MO, USA). Immunoreactive proteins were detected by chemoluminescence using the SuperSignal West Pico chemiluminescent substrate (Thermo Scientific), visualized with an ImageQuant 350 Digital Imaging System (GE Healthcare, Pittsburgh, PA, USA), quantified using the ImageQuant TL 7.0 software (GE Healthcare) and normalized against total proteins. Normalized protein levels were adjusted to 100 in ticks fed on control deer and compared to values obtained in ticks fed on vaccinated deer. Two experiments were conducted with similar results.

2.6. RNA extraction and real-time RT-PCR

Some genes encoding for proteins in the BPs underrepresented in ticks fed on SUB-vaccinated WTD were selected for the analysis of mRNA levels by real-time RT-PCR (Table 1). For selection, protein identification in *Rhipicephalus* spp. was considered to increase nucleotide sequence identity with *R. microplus* and the possibilities for successful PCR. Five ticks from each group were dissected for total RNA extraction using the RNeasy kit (Qiagen Inc., Valencia, CA, USA). The mRNA levels were characterized by real-time RT-PCR using gene-specific oligonucleotide primers (Table S1) and the Access RT-PCR system (Promega, Madison, WI, USA) in an iQ5 thermal cycler (Bio-Rad, Hercules, CA, USA). Negative control (C-) reactions were done in the absence of tick RNA. Gene expression levels were normalized against tick cyclophilin using the comparative Ct method [7] and compared between groups by Student’s t-test (p=0.05).
Table 1. Tick proteins in the BPs underrepresented in ticks fed on SUB-vaccinated WTD and selected for analysis of gene expression.

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Description (BP)</th>
<th>Protein levels (No. peptides/protein)</th>
<th>mRNA levels (normalized Ct values)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Control</td>
<td>BM86</td>
</tr>
<tr>
<td>L7LVH9</td>
<td>Histone cluster 2, h3c (BP1)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>L7LUH9</td>
<td>Glycine/serine hydroxymethyltransferase, GSHMT (BP2)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>L7M6W4</td>
<td>HSP60 (BP3)</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>SUB</td>
<td></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>BM86</td>
<td></td>
<td>100</td>
<td>50</td>
</tr>
</tbody>
</table>

Some genes encoding for proteins in the BPs underrepresented in ticks fed on SUB-vaccinated WTD were selected for the analysis of mRNA levels by real-time RT-PCR. SUB and BM86 were included in the analysis to characterize the effect of vaccination on the protein levels and expression levels of genes encoding for vaccine antigens. Protein levels were determined by proteomics except for SUB and BM86 which were determined by Western blot, normalized against total proteins and adjusted to 100 in ticks fed on control deer and compared to values obtained in ticks fed on vaccinated deer. Two experiments were conducted with similar results. Gene expression levels were normalized against tick cyclophilin, presented as Ave±S.E. (arbitrary units) and compared between ticks fed on vaccinated and control WTD by Student’s t-test (N=5; *p≤0.05). Abbreviations: BP1, DNA/RNA metabolism; BP2, Enzymatic process; BP3, Stress response; Control, ticks fed on control WTD; BM86, ticks fed on BM86-vaccinated WTD; SUB, ticks fed on SUB-vaccinated WTD.

3. Results

3.1. Characterization of the tick proteome

The proteomics analysis resulted in the identification of 169 and 174 tick and host proteins, respectively in R. microplus fed on WTD (Tables S2 and S3). The number of peptides used for protein identification showed that tick proteins were underrepresented in ticks fed on WTD vaccinated with SUB while host proteins were overrepresented in ticks fed on BM86-vaccinated animals (p<0.05; Figs. 1A and 1B). The number of peptides used to identify each protein was higher for host than for tick proteins, reflecting differences between protein databases (Fig. 1B). These results showed the first link between tick proteome and vaccination and suggested the need to further explore the association between proteins represented in fed ticks and the effect of vaccination on tick feeding and reproduction.
Figure 1. Global proteomics analysis and characterization of host proteins in the tick proteome. (A) Number of peptides representing tick and host proteins identified in *R. microplus* ticks fed on vaccinated and control deer. (B) Number of peptides per tick and host proteins identified in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between groups using a Student’s t-test (*p<0.05). (C) Number of peptides representing host Hemoglobin and Albumin proteins in *R. microplus* ticks fed on vaccinated and control deer. (D) Number of peptides per protein for host Hemoglobin and Albumin proteins in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between groups using a Student’s t-test (*p<0.05).

3.2. Host proteins in the tick proteome

The 174 host proteins identified in the tick proteome included Hemoglobins, Albumins, Complement proteins and Immunoglobulins among others (Table S3). The analysis of the most abundant host blood proteins such as Hemoglobin and Albumin (N=17 representing 10% of all host proteins identified; Table S3) showed that these proteins were overrepresented in both BM86- and SUB-vaccinated WTD when compared to ticks fed on control animals (*p<0.05; Figs. 1C and 1D).
3.3. Tick proteins in the tick proteome

The 169 tick proteins identified after proteomics analysis of *R. microplus* fed on vaccinated and control WTD were annotated and classified into 11 BPs (Fig. 2 and Table S2). Of them, the most abundant BPs corresponded to unknown process, protein metabolism, enzymatic process, structural proteins, DNA/RNA metabolism and stress response (Fig. 2).

**Figure 2. Tick proteins in the tick proteome.** Tick proteins identified in *R. annulatus* and *R. microplus* were functionally annotated and grouped according to their biological process (BP). The number of peptides used to identify proteins on each BP is shown for *R. microplus* ticks fed on vaccinated and control deer. Abbreviations: C, ticks fed on control WTD; BM86, ticks fed on BM86-vaccinated WTD; SUB, ticks fed on SUB-vaccinated WTD.

The analysis of proteins differentially represented in each of the BPs showed that proteins in the DNA/RNA metabolism, enzymatic process, stress response and reproduction BPs were underrepresented in ticks fed on SUB-vaccinated WTD (p<0.05; Figs. 3A-3D). Because of the role of tick SUB in the regulation of gene expression, the mRNA levels for genes encoding for some of the proteins in these BPs together with tick *sub* and *bm86* were analyzed by RT-PCR in ticks fed on vaccinated and control WTD (Table 1). As expected due to the existence of post-
transcriptional and post-translational mechanisms affecting mRNA and/or protein levels, no direct correlation was found between mRNA and protein levels (Table 1). Furthermore, high tick-to-tick variations in mRNA levels and the small number of ticks used for the analysis (N=5) because most fed female ticks were used for oviposition and proteomics, prevented from identifying statistically significant differences in most genes analyzed (Table 1). Nevertheless, the results showed that \textit{sub}, \textit{h3c} and \textit{gshmt} average mRNA levels were 50\% lower in ticks fed on SUB-vaccinated WTD when compared to ticks fed on control animals (Table 1). Interestingly, \textit{sub}, \textit{h3c} and \textit{hsp60} average mRNA levels were at least 50\% lower in ticks fed on BM86-vaccinated WTD when compared to controls (Table 1).

Tick structural proteins were overrepresented in ticks fed on WTD vaccinated with BM86 (p<0.05; Fig. 4E) and included Actin and other cytoskeleton proteins and muscle proteins such as Myosin (Table 2). The analysis of BM86 and SUB protein levels by Western blot showed that BM86 levels were lower in ticks fed on BM86-vaccinated deer while SUB levels were lower in ticks fed on SUB-vaccinated deer when compared to controls (Table 1 and Fig. 3F).

Table 2. \textbf{Tick structural proteins overrepresented in ticks fed on BM86-vaccinated WTD}.

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Description</th>
<th>Protein levels (No. peptides/protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Control</td>
</tr>
<tr>
<td>Q5D579</td>
<td>Actin</td>
<td>12</td>
</tr>
<tr>
<td>Q6X4W3</td>
<td>Actin</td>
<td>11</td>
</tr>
<tr>
<td>E2J6Z2</td>
<td>Alpha tubulin</td>
<td>1</td>
</tr>
<tr>
<td>L7M1J3</td>
<td>Tubulin alpha 1c</td>
<td>1</td>
</tr>
<tr>
<td>F0J8M3</td>
<td>Hypothetical glycine-rich secreted cement protein</td>
<td>0</td>
</tr>
<tr>
<td>L7LX42</td>
<td>Actin-binding cytoskeleton protein filamin</td>
<td>0</td>
</tr>
<tr>
<td>B7P1Q2</td>
<td>Myosin heavy chain</td>
<td>1</td>
</tr>
<tr>
<td>L7M506</td>
<td>Lamin</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 3. Differentially represented tick proteins. (A) Number of peptides per tick protein in the DNA/RNA metabolism BP in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between ticks fed on vaccinated and control WTD using a Student’s t-test (*p<0.05). (B) Number of peptides per tick protein in the enzymatic process BP in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between ticks fed on vaccinated and control WTD using a Student’s t-test (*p<0.05). (C) Number of peptides per tick protein in the stress response BP in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between ticks fed on vaccinated and control WTD using a Student’s t-test (*p<0.05). (D) Number of peptides per tick protein in the reproduction BP in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between ticks fed on vaccinated and control WTD using a Student’s t-test (*p<0.05). (E) Number of peptides per tick protein in the structural protein BP in *R. microplus* ticks fed on vaccinated and control deer. Data was represented as Ave+S.D. and compared between ticks fed on vaccinated and control WTD using a Student’s t-test (*p<0.05). (F) BM86 and SUB protein levels were determined by Western blot in *R. microplus* fed on vaccinated and control WTD, quantified and normalized against total proteins. Normalized protein levels (in arbitrary units) were adjusted to 100 in ticks fed on control deer and compared to values obtained in ticks fed on vaccinated deer. Two experiments were conducted with similar results. Abbreviations: Control, ticks fed on control WTD; BM86, ticks fed on BM86-vaccinated WTD; SUB, ticks fed on SUB-vaccinated WTD.
Figure 4. Hypothetical models for vaccine effector mechanisms. (A) Anti-BM86 antibodies and complement proteins ingested by feeding ticks interact with the antigen in the surface of gut cells producing alterations in gut structure and function that result in less blood ingestion, which in turn affects tick reproduction. (B) Vaccination reduces SUB levels in feeding ticks by an unknown mechanism but probably mediated by antibody-antigen interactions in the cell cytoplasm and affects its function as a transcriptional regulator of its own expression and of genes involved in several BPs playing an important role in tick feeding and reproduction. Another possibility yet to be tested is that antibodies affect SUB function by blocking its translocation to the nucleus.
4. Discussion

4.1. Molecular mechanisms associated with BM86 vaccine protection

The overrepresented host proteins in ticks fed on WTD vaccinated with BM86 suggested that these ticks digested less blood than ticks fed on control animals. Lower blood digestion probably resulted in less blood ingested by ticks fed on BM86-vaccinated WTD when compared to ticks fed on control animals, a result that correlates with the 14% reduction in tick weight reported in the vaccine trial from which these ticks were collected [8] and in other trials with BM86 [6, 16]. BM86 is a glycoprotein located on the surface of tick gut cells but its function is unknown and produces very little phenotypic alterations after gene knockdown [17, 18]. Nevertheless, the synergistic effect between sub and bm86 knockdown and the alterations observed in the gut of double-knockdown ticks suggest that BM86 may be involved in maintaining the structure of the tick gut [17, 18].

The results of the proteomics analysis showed that vaccination with BM86 reduces tick feeding and increases the presence of some structural proteins in ticks fed on vaccinated WTD. The structural proteins overrepresented in ticks fed on BM86-vaccinated animals may represent alterations produced by anti-BM86 antibodies ingested by these ticks, again suggesting a role for this molecule in the maintenance of gut structure.

Taken together, these results suggested that BM86 might be involved in maintaining gut structure needed for blood digestion. Vaccination with BM86 protects against R. microplus infestations through a collaborative action between the complement proteins and IgG1 antibodies that most efficiently fix the complement in cattle, but antibodies alone can mediate the protective response to vaccination [19-22]. Thus, anti-BM86 antibodies and complement proteins ingested by feeding ticks interact with the antigen on the surface of gut cells to affect the function and levels of this protein producing alterations in gut structure and function that result in less blood ingestion, which in turn affects tick reproduction. In fact, lesser protease activity in R. annulatus when compared to R. microplus has been suggested to results in more efficient antibody-antigen interactions and higher BM86 vaccine efficacy [13]. The alterations in gut structure could explain the red color seen in ticks fed on BM86-vaccinated hosts probably due to leakage of the blood meal into the tick body cavity [23, 24].

4.2 Molecular mechanisms associated with SUB vaccine protection

The analysis of the most abundant host blood proteins showed that as for ticks fed on BM86-vaccinates WTD, ticks fed on SUB-vaccinated animals digested less blood when compared to
ticks fed on control animals. As with the BM86 vaccine, this effect probably resulted in less blood ingested by ticks fed on SUB-vaccinated WTD when compared to ticks fed on control animals, a result that correlates with the 11% reduction in tick weight reported in the vaccine trial from which these ticks were collected [8] and in other trials with SUB [10, 11].

In general, proteomics analysis showed that tick proteins were underrepresented in ticks fed on SUB-vaccinated WTD with statistically significant differences for DNA/RNA metabolism, enzymatic process, stress response and reproduction BPs. Furthermore, the analysis of the mRNA levels for selected genes showed that \textit{sub}, \textit{h3c} and \textit{gshmt} mRNA levels were lower in ticks fed on SUB-vaccinated WTD, suggesting that at least for some genes, protein underrepresentation resulted from downregulation at the transcriptional level. These results confirmed that vaccination with SUB downregulates \textit{sub} gene expression in ticks [7] due to SUB auto-regulation [25], which in turn may affect the expression of other genes regulated by SUB [25-27]. This result at the transcriptional level was also reflected in the reduction of SUB protein levels in vaccinated WTD.

These results suggested that vaccination reduces SUB levels in feeding ticks by an unknown mechanism but probably mediated by antibody-antigen interactions in the cell cytoplasm [28] which affects SUB function as a transcriptional regulator of its own expression and of genes involved in several BPs playing an important role in tick feeding and reproduction [12, 28]. Another possibility yet to be tested is that antibodies affect SUB function by blocking its translocation to the nucleus. Interestingly, proteins involved in tick reproduction such as GP80 (20 and 11 peptides in ticks fed on control and vaccinated WTD, respectively) and Vitellogenins (26 and 20 peptides in ticks fed on control and vaccinated WTD, respectively) were underrepresented in ticks fed on SUB-vaccinated WTD, correlating with the 21% reduction in the oviposition reported in the vaccine trial from which these ticks were collected [8] and corroborating the role of this molecule in tick reproduction [29].

5. Conclusions

A summary of the results reported in this paper is shown in the hypothetical models depicted in figure 4. These results support our hypothesis that the characterization of the proteome in ticks after feeding on vaccinated and control hosts provides information on the molecular mechanisms that are associated with the effect of vaccination on tick feeding and reproduction. The results of the proteomics analysis showed that although both vaccines reduce tick feeding and reproduction, they act through different protective mechanisms. The main protective effect of the BM86 vaccine is the effect on BM86 function and levels, which alters tick gut structure and function and consequently blood digestion and reproduction. For the SUB vaccine, the results showed that
vaccination reduces protein levels and affects its function as a transcriptional regulator of genes involved in several BPs important for tick feeding and reproduction. These results have important implications for tick vaccine research, supporting the study of the molecular mechanisms associated with vaccine protection to combine antigens that act through different protective mechanisms to increase vaccine efficacy.

Acknowledgements

We thank C. Almazán and D. Carreón (Universidad Autónoma de Tamaulipas, Mexico) for supplying ticks and A. Marina and M. del Valle (CBMSO Protein Chemistry Facility, Spain) for technical assistance. This work was supported by project BFU2011-23896. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Postgraduate training network for capacity building to control ticks and tick-borne diseases) within the FP7-PEOPLE-ITN programme (EU Grant No. 238511). N. Ayllón is funded by MEC, Spain.

Supplementary material.

Supplementary methods. Table S1. PCR conditions and sequences of primers used for real-time RT-PCR.
Supplementary results.
Table S2. Tick proteins identified in ticks fed on vaccinated and control deer (FDR=0.05).
Table S3. Host proteins identified in ticks fed on vaccinated and control deer (FDR=0.01).

References

Chapter VI

General discussion

Proteomics of tick-pathogen and tick-host interactions: Current status and perspectives

Abstract

Ticks are vectors of pathogens that cause disease in humans and animals worldwide. Recent results in tick-host-pathogen interaction studies enabled the expansion of research to an integrated “omics” approach that combines post-genomics tools such as transcriptomics, proteomics and metabolomics. Among them, proteomics enables the study of the proteins present in certain tissues or whole organisms, that unlike the genome, display variations in response to different conditions. Therefore, proteomics studies are dynamic and difficult to perform. This review paper gives an overview of the most important proteomics results in the study of the interactions occurring between ticks and the pathogens they transmit and between ticks and their hosts. Furthermore, results of proteomics studies performed to obtain an insight into tick biology and evolutionary relationships, stress response and tick-control vaccine research are reviewed. However, tick proteomics is still at its infancy and further studies are required. Therefore, we emphasize current drawbacks and the future approaches that should be applied to further characterize molecular interactions at the tick-host-pathogen interface.

1. Introduction

Ticks are obligate haemotophagus ectoparasites with a worldwide distribution from the Arctic to tropical regions [1]. The pathogens that ticks transmit greatly impact both human and animal health [1]. For example, *Ixodes* spp., a worldwide distributed tick species, vectors several human pathogens including *Borrelia*, *Anaplasma*, *Rickettsia*, *Coxiella*, *Francisella*, and *Babesia* species, as well as viruses such as the tick-borne encephalitis virus (TBEV) [1]. *Rhipicephalus* (*Boophilus*) spp. are major ectoparasites of veterinary importance in tropical and subtropical regions of the world [2, 3]. Infestations with the cattle tick, *R. microplus*, economically impact cattle production by reducing weight gain and milk production and by transmitting pathogens that cause babesiosis (*Babesia bovis* and *Babesia bigemina*) and anaplasmosis (*Anaplasma marginale*) [2,3].

A recent advance in the systems biology technologies enabled the expansion of research to an integrated “omics” approach combining post-genomics tools such as transcriptomics, proteomics and metabolomics [4]. Among them, proteomics emerges as a powerful tool for a large scale analysis of all the proteins produced by single cells, tissues or organisms [5]. It enables the identification of proteins, studies their localization, modifications, function and possible interactions or complexes they can form [5]. Unlike the genome, there is no single, static proteome in any organism but instead, dynamic collections of proteins in different cells and tissues that display moment-to-moment variations in response to several conditions such as
stress or infectious agents, making proteomics studies difficult to perform [5]. However, proteomics is a key tool to increase the knowledge on the biology of tick-host-pathogen interactions and mechanisms of transmission of infectious diseases.

2. Proteomics of tick-pathogen interactions

Ticks and tick-borne pathogens have co-evolved molecular interactions involving genetic traits of both the tick and the pathogen that mediate their development and survival [6–9]. Tick-pathogen interactions are the most studied aspect of tick-host-pathogen interactions [5]. Nevertheless, there are not many large-scale studies integrating transcriptomics with proteomics to define the mechanisms of these interactions. Pathogen must be able to adapt to various environments during its life cycle and manipulate tick physiology to avoid immune responses and establish infection [2]. Tick compartments primarily involved in immune response and/or pathogen transmission such as salivary glands, midgut, ovaries and hemolymph are targeted in most studies [10].

Salivary glands and saliva have a central role in pathogen transmission, feeding and subversion of host-immune responses [6]. These proteomes, ’’sialomes’’, have been shown to contain an array of biologically active molecules, including anti-hemostatic, angiogenic, anti-inflammatory, and immunomodulatory proteins involved in interactions between pathogens, vectors and vertebrate hosts [11]. Some immunosuppressive proteins in the tick saliva have coevolved with tick-borne pathogens. For example, tick salivary protein, Salp15 is specifically induced by Borrelia burgdorferi to facilitate its transmission to the vertebrate host [12]. Sialome composition can vary within tick families and even within certain species suggesting that these proteins evolve quickly, possibly due to the selection pressure exerted by host immune responses [11]. Salivary transcriptomes have been characterized in most hard tick species identifying more than 25 protein families expressed in their sialomes, but little has been done in pathogen detection [13–18]. The sialome has been also characterized in most medically relevant tick species (Table 1) enabling the construction of a protein database for this tissue [13,14,19]. Additional studies of excreted saliva are currently available in Amblyomma americanum, Amblyomma maculatum and Ixodes scapularis ticks [13,19].

The midgut is of primary importance in tick feeding and disease acquisition being the first site of blood meal digestion and pathogen interaction [10]. It is the only part of the digestive tract without cuticle and anti-microbial peptides are released there in response to pathogens [25]. The midgut proteome has been investigated in R. microplus ticks [20,21] (Table 1). Up-regulation of metabolic enzymes in Babesia bovis infected ticks indicated parasite-induced changes in electron and proton transport, protein processing and retinoic acid metabolism. Most of the up-
regulated pathogen molecules are involved in signaling or metabolism [21]. Furthermore, differentially expressed tick cytoskeletal proteins suggested that modulation of the cytoskeleton during the interplay between arthropod vectors and pathogens could determine the success of infection [21]. Additional studies of midgut proteome from partially fed adult female cattle ticks identified novel proteins such as clathrin-adaptor protein and membrane-associated trafficking proteins such as Syntaxin 6 and Surfeit 4 [20]. Midgut infection could induce changes in the expression of proteins involved in pathogen invasion such as immune, stress-response and cytoskeleton proteins, and glycolytic and metabolic enzymes [25].

Pathogen transmission in ticks can be transstadial, when pathogen is transmitted between tick life stages, or transovarial, when it is transmitted from parent ticks to offspring. The proteomics analyses of soluble and membrane proteins from ovaries of Babesia bovis-infected and uninfected Rhipicephalus microplus initiated a proteome database for studies on transovarial pathogen transmission [22] (Table 1).

Tick hemolymph has important functions in nutrient and hormone transport, amino acid storage, water balance, responses to injury and immunity and at some stages of their life cycle pathogens can be found in this fluid [23]. Up to date, proteomics analysis of tick hemolymph has been conducted only on soft ticks. A differential proteomics study of systemic innate immunity of the soft tick Ornithodoros savignyi challenged with Candida albicans identified five proteins in the hemolymph specifically regulated by pathogen infection [24]. However, they could not be definitively identified due to the limited number of sequences available in databases for these ticks and a limited amount of protein sample that can generally be obtained from the hemolymph.

Whole-tick body proteome has been analyzed in few occasions [30-33]. However, these studies are relevant in cases when the effect of a specific environmental condition or stimuli needs to be investigated. The first study of entire R. microplus ticks was done on unfed larvae giving the first, however limited database of proteins produced at this developmental stage [25]. Recent studies investigated further the proteome of these ticks, together with R. annulatus and showed that proteomics approaches to analyze ticks are very different depending on the ticks feeding state [27].

For the detection of pathogens in naturally infected Rhipicephalus ticks highly sensitive proteomics methods such as differential in gel electrophoresis (DIGE) saturation labeling were successfully applied increasing the identification of lowly represented proteins and proteins from limited samples [28].
To date, most of the studies investigating the mechanisms involved in tick-pathogen interactions use tick cell cultures as an important alternative to *in vivo* studies. These models enable a controlled environment for the investigation of biological pathways involved in tick-pathogen interactions. However, recent studies showed that differences in the protein content sometimes exist between the two models [5]. Even in tick cell culture, only few studies have used proteomics approaches to characterize response to pathogen infection [33]. A good example of the successful application of tick cell cultures as models for tick-pathogen interaction studies is a comparative proteomics study of the cattle pathogen, *Anaplasma marginale* proteins overrepresented in infected ISE6 tick cells and mammalian cells where results were corroborated *in vivo* on infected *Dermacentor andersoni* ticks [99]. A similar study used a shotgun proteomics approach and identified *Ehrlichia chaffeensis* proteins involved in the interactions between the bacterium and tick cells [30]. Functional genomics studies of tick cells in response to infection with *A. marginale* combined the application of suppression-subtractive hybridization and differential in-gel electrophoresis analyses of cultured IDE8 tick cells and demonstrated that a molecular mechanism occurs by which tick cell gene expression mediates the *A. marginale* developmental cycle and trafficking through ticks [31].

3. Proteomics of tick stress response

Heat-shock and other stress responses are conserved reactions of cells and organisms to different stress conditions such as extreme temperatures, toxicity and pathogen infection [32,33]. The expression of these proteins protects cells and organisms from damage providing higher levels of tolerance to environmental stress and is crucial for cell survival [32–34]. Stress responses to tick attachment, blood feeding and pathogen infection have been characterized in ticks and cultured tick cells [20,21,29,35–42]. Proteins such as glutathione-S-transferase, selenoproteins, metallothioneins, and ferritin have been shown to be involved in the cellular response to different stress conditions [6,20,21,29,31,35–38,40–44]. Combined proteomics and transcriptomics analyses demonstrated that the stress response is activated in ticks and cultured tick cells after *Anaplasma* spp. infection and heat shock [29]. However, studies of natural vector-pathogen relationships, showed that heat-shock proteins (HSPs) and other stress response proteins were not strongly activated, probably reflecting tick-pathogen co-evolution [29]. Nevertheless, proteomics analysis of *I. scapularis* ISE6 tick cells in response to *A. phagocytophilum* infection showed that some HSPs such as the HSP70 family are overrepresented while other putative heat-shock proteins (HSPs) such as HSP20 are underrepresented in infected cells. Tick responses to pathogen infection may also be regulated at the transcriptional and/or post-transcriptional levels.
Recent proteomics and transcriptomics studies of *Dermacentor reticulatus* unfed larvae detected the activation of the stress response in this developmental stage [34]. This likely counteracts the negative effect of temperature and other stress conditions such as *Rickettsia* infection and favors tick adaptation to environmental conditions to increase tick survival. Overall, current studies show that tick stress response is highly complex process, indicating that the tick response to pathogen infection involves several molecules, such as subolesin and HSPs, but shows tick and pathogen-specific factors, which may vary among tick tissues and developmental stages [28,29,45,46].

### Table 1. Most relevant proteomics studies on major tick species and tick cell lines

<table>
<thead>
<tr>
<th>Tick species (or tick cell line)</th>
<th>Pathogen (if infected)</th>
<th>Tick tissue</th>
<th>Proteomics methods</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Salivary gland and sialome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Ixodes scapularis</em></td>
<td>Saliva and salivary glands of partially engorged ticks</td>
<td>1DE followed by Edman degradation</td>
<td>[13]</td>
<td></td>
</tr>
<tr>
<td><em>Amblyomma americanum</em>, <em>Amblyomma maculatum</em></td>
<td>Saliva of partially engorged ticks</td>
<td>1DE and 2DE followed by MS</td>
<td>[19]</td>
<td></td>
</tr>
<tr>
<td><em>Ornithodoros moubata</em>, <em>Ornithodoros erraticus</em></td>
<td>Salivary glands of adult unfed ticks</td>
<td>2-DE followed by MS, MS/MS and LC-MS/MS</td>
<td>[70]</td>
<td></td>
</tr>
<tr>
<td><em>Ornithodoros coriaceus</em></td>
<td>Salivary glands from adult ticks</td>
<td>Edman degradation, MS/MS and LC-MS/MS</td>
<td>[72]</td>
<td></td>
</tr>
<tr>
<td><em>Ixodes ricinus</em></td>
<td>Salivary glands of unfed and fed adult female ticks</td>
<td>1DE followed by LC-MS/MS</td>
<td>[14]</td>
<td></td>
</tr>
<tr>
<td><em>Ornithodoros parkeri</em></td>
<td>Salivary glands from nymphs and adult ticks</td>
<td>1DE and 2DE followed by Edman degradation, MS/MS and LC-MS/MS</td>
<td>[71]</td>
<td></td>
</tr>
<tr>
<td><em>Argas monolakensis</em></td>
<td>Salivary glands of unfed and fed adult female ticks</td>
<td>1DE followed by Edman degradation</td>
<td>[69]</td>
<td></td>
</tr>
<tr>
<td><em>Amblyomma variegatum</em></td>
<td>Salivary glands of female adult ticks</td>
<td>1DE followed by LC-MS/MS</td>
<td>[15]</td>
<td></td>
</tr>
<tr>
<td><strong>Midgut and peritrophic matrix proteome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Rhipicephalus microplus</em></td>
<td>Babesia bovis</td>
<td>Midgut of partially fed adult ticks</td>
<td>1DE and 2DE followed by LC-MS/MS</td>
<td>[21]</td>
</tr>
<tr>
<td><em>Rhipicephalus microplus</em></td>
<td></td>
<td>Midgut of partially fed females</td>
<td>2-DE and gel-free LC-MS/MS</td>
<td>[20]</td>
</tr>
<tr>
<td><strong>Hemolymph proteome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Organ/Protein Source</td>
<td>Methodology</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Amblyomma hebraeum</td>
<td>Hemolymph of fed female ticks</td>
<td>SELDI-TOF MS</td>
<td>[74]</td>
<td></td>
</tr>
<tr>
<td>Ornithodoros savignyi</td>
<td>Hemolymph of adult female ticks</td>
<td>2-DE followed by MS and MS/MS</td>
<td>[24]</td>
<td></td>
</tr>
<tr>
<td>Candida albicans</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole-body proteome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhipicephalus microplus</td>
<td>Unfed larvae</td>
<td>2DE followed by MS and MS/MS</td>
<td>[25]</td>
<td></td>
</tr>
<tr>
<td>Ixodes ricinus</td>
<td>30-weak-old nymphs</td>
<td>2DE followed by MS and MS/MS</td>
<td>[26]</td>
<td></td>
</tr>
<tr>
<td>Ixodes scapularis, Amblyomma americanum</td>
<td>Adult unfed ticks</td>
<td>Off-gel digestion followed by LC–MS/MS</td>
<td>[51]</td>
<td></td>
</tr>
<tr>
<td>Rickettsia conorii, Ehrlichia canis, Theileria annulata, Anaplasma ovis</td>
<td>Whole internal organs of infected and uninfected ticks</td>
<td>2-DE followed by MS, MS/MS and LC-MS/MS</td>
<td>[28]</td>
<td></td>
</tr>
<tr>
<td>Rhipicephalus sanguineus, Rhipicephalus bursa, Rhipicephalus turanicus</td>
<td>whole repleted engorged female ticks</td>
<td>1DE followed by LC-MS/MS</td>
<td>[27]</td>
<td></td>
</tr>
<tr>
<td>Rhipicephalus microplus, Rhipicephalus annulatus</td>
<td>Rickettsia sp.</td>
<td>Unfed larvae</td>
<td>1DE followed by LC-MS/MS</td>
<td>[unpublished results]</td>
</tr>
<tr>
<td>Dermacentor reticulatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amblyomma americanum, Amblyomma cajenennse, Amblyomma variegatum</td>
<td>Unfed females, males and nymphs</td>
<td>1DE followed by LC-MS/MS</td>
<td>[unpublished results]</td>
<td></td>
</tr>
<tr>
<td>Babesia bovis</td>
<td>Ovaries</td>
<td>2DE followed by LC-MS/MS</td>
<td>[22]</td>
<td></td>
</tr>
<tr>
<td>Tick cell lines</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISE6 tick cell line surface protein</td>
<td>Anaplasma marginale</td>
<td>1DE followed by LC-MS/MS</td>
<td>[75]</td>
<td></td>
</tr>
<tr>
<td>ISE6 and AAE2 tick cell lines, canine macrophage cellline DH82 derived E. chaffeensis</td>
<td>Ehrlichia chaffeensis</td>
<td>1DE followed by MS and LC-MS/MS</td>
<td>[30]</td>
<td></td>
</tr>
<tr>
<td>ISE6 tick cell lines and bovine erythrocytes; Dermacerator andersoni</td>
<td>Anaplasma marginale</td>
<td>Tick and bovine cells; midgut and salivary glands of infected tick</td>
<td>2-DE followed by LC-MS/MS</td>
<td>[76]</td>
</tr>
</tbody>
</table>

4. Proteomics of tick-host interactions

Identifying the mechanisms by which hosts prevent heavy tick infestations is an essential step for the development of predictive markers for resistance and vaccines designed to promote an appropriate immune response to control tick infestations [47]. The effect of host response to tick infestation has been investigated in several studies of skin samples from resistant and susceptible cattle breeds after being challenged with *R. microplus* ticks [47–49]. Differential expression of
various calcium signaling genes and proteins revealed their importance in host immunological response to ticks. These studies were further expanded to proteomics level by application of isobaric tag for relative and absolute quantification (iTRAQ) labeling method, that revealed that epidermal physical barrier has an essential role in conferring greater resistance to tick infestation in cattle and the physical structure of epidermal layers of the skin may represent the first line of defense against ectoparasites [50]. The presence of host proteins was also analyzed within ticks [51]. Host blood proteins were detected even in unfed ticks months after molting [59]. Ticks might use these proteins as a reserve during the off-host periods.

Proteomics analyses of host effect on tick proteome were investigated on *R. microplus* ticks fed on cattle and white tailed deer (WTD) showing that ticks fed on cattle have overrepresented proteins involved in blood digestion, correlating with the higher tick numbers, weight and reproductive performance obtained in ticks fed on cattle (unpublished results). Higher Fibrinogen levels in ticks fed on cattle may reflect host response to tick infestations and an indicator of the lower tick infestations observed in WTD with possible implications in host-tick co-evolution. These results demonstrate the existence of host factors that affect tick feeding and reproduction impacting vector evolution and ecology and suggest the possibility of using these tick proteins for the control of cattle tick infestations.

5. Proteomics and tick vaccine research

Control of tick infestations is difficult because ticks have few natural enemies [52]. Presently, tick control consists of integrated strategies that combine chemical acaricides with other control methods such as immunization with tick vaccines. In the early 1990s, two commercial vaccines containing the *R. microplus* BM86 gut glycoprotein’s became available for cost-effective control of cattle tick infestations and the prevalence on some tick-borne pathogens while reducing environmental contamination and preventing the selection of drug resistant ticks that result from repeated acaricide applications [7, 61-68]. Proteomics has recently proven its value to find new vaccine candidates against a variety of bacterial species [56]. However, the application of proteomics to tick vaccine research is still at its infancy. For example, the efficacy of the BM86 vaccine is always higher against the related tick species *R. annulatus* than against *R. microplus* [57–59]. This result is similar for other antigens such as BA86 [60], Subolesin [61] and Ferritin2 [62] suggesting that tick genetic and/or physiological factors may affect the efficacy of tick vaccines [2, 39, 58–60, 62–64]. The comparison of the proteomes of *R. annulatus* and *R. microplus* female ticks after feeding on BM86-vaccinated and control cattle revealed the existence of lesser protein degradation machinery in *R. annulatus*, a mechanism used by ticks to counteract host
response to infestation, pathogen infection and for digestion of blood meal \[27,65–68\]. This process may contribute to higher vaccine efficacy in \textit{R. annulatus} when compared to \textit{R. microplus} resulting in lower protection against host responses \[27\]. Higher levels of protease inhibitors were also detected in \textit{R. annulatus} and may represent an evolutionary advantage to aid in protection against pathogen infection \[65–68\].

6. Proteomics for the characterization of tick biology and evolution

Proteomics knowledge can also be applied in tick phylogenetic studies \[11\]. However, there are only few studies of tick evolutionary relationships using proteomics \[69, unpublished results\]. Transcriptomics and proteomics studies have been conducted in soft ticks to investigate the phylogenetic relationships between soft and hard ticks and understand the mechanism of hematophagy \[69–72\]. These studies revealed that major protein families are conserved in hard and soft ticks, establishing a catalogue of putative secreted salivary proteins from various tick species \[69\]. Identified proteins involved in host hemostatic and immune defenses showed that soft and hard ticks evolved anti-haemostatic mechanisms independently and suggested that ancestral tick salivary gland repertoire contained various protein families but with only protein members in each family \[69\]. The proteomics study of the \textit{Amblyomma} tick species is the first study that applied the phyloproteomics concept for the phylogenetic analysis of tick species (unpublished results). This study applied de novo protein sequencing obtaining the information that is not affected by a limiting sequence data available for these ticks and assembly problems associated with genomics and transcriptomics projects. Protein sequences generated de novo in three \textit{Amblyomma} spp., \textit{A. americanum}, \textit{A. variegatum} and \textit{A. cajennense}, were used for multi locus analysis in comparison with established 16S rDNA phylogeny using \textit{I. scapularis} sequences as outgroup. The results showed similar topology for both phylogenetic trees, thus suggesting that phyloproteomics may be a useful tool for the phylogenetic analysis of tick species in which sequence data is a limiting factor.

7. Final considerations

Studies of tick-host-pathogen interactions at the genomics level provide comprehensive sequence information. However, they cannot predict the molecular mechanisms affecting these interactions. This limitation pushed forward progresses in post-genomic approaches such as transcriptomics and proteomics. Proteomics studies have identified proteins and biochemical pathways that may be involved in molecular crosstalk at the tick-host-pathogen interface and candidate protective antigens. However, key information is still lacking.
The application of proteomics can sometimes be difficult. One of the limitations for proteomics research is that unlike DNA, proteins cannot be amplified to increase the sensitivity of detection techniques. Furthermore, in vivo studies give a limited amount of proteins that can be obtained from the tick or tick-borne pathogens. Therefore, most studies still rely on gene expression as an indicator of the presence of active proteins. Nevertheless, the relationship between mRNA and protein is complex and recent studies revealed a relatively weak correlation between mRNA expression levels and protein abundance due to existence of post-transcriptional and post-translational modifications.

Figure 1. Schematic representation of the most important results, gaps and perspectives in proteomic studies of tick-host and tick-pathogen interactions.

One of the major constraints when working with obligate intracellular parasites is the excess of proteins of host or vector origin that can interfere with pathogen protein detection [4]. To overcome this problem, equalizer technologies combined with fractionation techniques can be used to eliminate highly abundant interfering proteins and concentrate low-abundance proteins. Nevertheless, in some cases, both the proteome of the parasite and infected host cells are of interest, since infection processes can cause changes in the host cells but host can also induce
changes in the parasite. Therefore, an optimal preparation method needs to be determined empirically to enable the detection of both, parasite and host/vector proteins.

For accurate protein identification, the knowledge of nucleotide and protein sequence databases is very important [73]. The annotated genome sequences are available for most of the tick-borne pathogens but only the genome of the tick, *I. scapularis* is close to completion, thus limiting the impact of proteomics studies. In addition, many proteins are specific to a specific tick and some homologous molecules may only be found in ticks with a very close phylogenetic relationship. Therefore, there is an urgent need for more genomic information and knowledge on tick proteins present in a variety of tissues, life stages and species. This problem can be overcome using *de novo* sequencing software packages designed to extract amino acid sequence information without the use of MS databases and the combination of transcriptome and proteome analyses.

Even with all current limitations, an integrated, constantly developing methods of systems biology hold a key to understanding the molecular mechanisms at the tick-host and tick-pathogen interfaces and the discovery of novel target molecules for the control of tick infestations and tick-borne diseases.

**Acknowledgments**

This work was supported by project BFU2011-23896. M. Popara is an Early Stage Researcher supported by the POSTICK ITN (Post-graduate training network for capacity building to control ticks and tick-borne diseases) within the FP7-PEOPLE-ITN programme (EU Grant No. 238511).

**References**


Conclusions

1. The optimization of different proteomics methodologies to study of ticks has allowed the development of new protocols suitable for the characterization of tick proteome.

2. The characterization of *Amblyomma* ticks proteomes showed specific differences between developmental stages and species and although ticks analyzed were unfed, host proteins were over-represented in adults acting as protein reserve to supply nutrients during periods off-host.

3. Phylogenomics approach by using *de novo* sequencing from proteomics data appears to be a useful tool for the phylogenetic analysis of tick species in which protein sequence data is a limiting factor.

4. The activation of stress responses in *D. reticulatus* unfed larvae likely counteracts the negative effect of temperature and other stress conditions such as *Rickettsia* infection on tick questing behavior and favors tick adaptation to environmental conditions to increase their survival. These results show mechanisms that have evolved in *D. reticulatus* ticks to survive under stress conditions and suggest that they are conserved across hard tick species.

5. *R. microplus* ticks fed on cattle and deer show host-specific differences in their proteome. *R. microplus* ticks fed on cattle have overrepresented tick proteins involved in blood digestion and reproduction when compared to ticks fed on WTD, correlating with the higher tick numbers, weight and reproductive performance observed in ticks fed on cattle. These results demonstrate the existence of host factors that affect tick feeding and reproduction that impact on vector evolution and ecology.

6. Protein degradation machinery is under represented in *R. annulatus* when compared to *R. microplus*. BM86 mRNA and protein levels are similar in both tick species, suggesting that lesser protease activity in *R. annulatus* results in more efficient antibody-antigen interactions and higher vaccine efficacy. These results suggest that not only genetic differences, but also physiological differences between tick species and possibly between geographic strains of the same species may affect tick vaccine efficacy.

7. The results of the proteomics analysis showed that although BM86 and Subolesin (SUB) vaccines reduce tick feeding and reproduction, they act through different protective mechanisms. The main protective effect of the BM86 vaccine is the effect on BM86 function, which alters tick gut structure and function and consequently blood digestion and reproduction. For the SUB vaccine, the results showed that vaccination reduces protein levels and affects its function as a transcriptional regulator of genes involved in several biological processes important for tick feeding and reproduction.
Summary

In the post-genomics era, proteomics includes promising strategies to characterize dynamic interactions that cannot be analyzed by genomic or transcriptomic approaches. The application of proteomics on tick research is, however, still in its infancy and more research is needed to achieve full potential of this technology. These works therefore focused on developing new proteomic strategies and apply proteomics to get a further insight into tick-host –pathogen interactions. During the course of this thesis systems biology studies, primarily based on proteomics and complemented with transcriptomics and immunological studies, were performed identifying differentially expressed proteins in different ticks species that enabled a further insight into tick biology, development and evolution.

The review of the current results of quantitative tick proteomics presented in Chapter I outlined the necessity of integrating the investigations on transcriptomics with proteomics level using a systems biology approach. This holistic approach to the study of tick-host-pathogen interactions is critical to advance the understanding of tick-pathogen and tick-host interactions.

Due to a limited number of studies applying proteomics in tick research, new protocols for protein extraction and analysis were developed during the course of this work (Chapter II) with a special emphasis is put on working with replete ticks collected after feeding on vertebrate hosts.

The application of proteomics to study of tick biology, development and evolution was presented in Chapter III. The first part of the study gave a first report on the proteome of the most important Amblyomma tick species. These ticks are relevant vectors of zoonotic pathogens worldwide. Nevertheless, very little information is available on their genome, transcriptome and proteome. The obtained results provided new information to understand the physiology, development and evolution of these tick species. Phyloproteomics using de novo protein sequencing was assayed as a new approach for the phylogenetic analysis of tick species in which sequence data is a limiting factor. In the second part of the study, the investigations done on D. reticulatus unfed larvae supported the use of paired end sequencing approach for the analysis of transcriptomics data. Furthermore, the proteomics informed by transcriptomics (PIT) approach showed a high efficiency in protein identification using a database created from transcriptomics data to search mass spectra thereby confirming to be a good alternative for organisms such as D. reticulatus with little sequence information available. Metabolic and cellular processes involving enzymatic reactions were identified as the most active processes, suggesting that ticks are very active during this life stage. The study revealed that the stress responses were activated in D.
unfed larvae. These response likely counteract the negative effect of temperature and other stress conditions such as *Rickettsia* infection and favor tick adaptation to environmental conditions to increase tick survival. These results showed mechanisms that have evolved in *D. reticulatus* ticks to survive under stress conditions and suggested they are conserved across hard tick species.

There are very few studies investigating the effect of a specific host on ticks biology. The existence of host factors that affect tick feeding and reproduction and impact on vector evolution and ecology was discovered in a comparison of *Rhipicephalus microplus* tick proteomes fed on cattle with the same tick species fed on white tailed deer (WTD) ([Chapter IV](#)). *R. microplus* ticks fed on cattle have overrepresented tick proteins involved in blood digestion and reproduction when compared to ticks fed on WTD, correlating with the higher tick numbers, weight and reproductive performance observed in ticks fed on cattle. These ticks probably ingested more blood as reflected by higher tick weights and Alpha-2-macroglobulin and Immunoglobulin-like protein levels. However, higher levels of host Hemoglobin, Haptoglobin and Albumin in ticks fed on WTD correlated with the lesser blood digestion machinery in these ticks. Fibrinogen levels in ticks fed on cattle were higher and may reflect host response to tick infestations explaining lower tick infestations observed in WTD possibly due to host-tick co-evolution.

The potential of proteomics approaches to characterize protective mechanisms of tick vaccines was presented in [Chapter V](#). The efficacy of the vaccination with BM86 and other tick antigens has shown to be always higher for *R. annulatus* than against *R. microplus* ticks, suggesting that tick genetic and/or physiological factors may affect vaccine efficacy. To investigate this, the proteomes of *R. annulatus* and *R. microplus* female ticks were compared after feeding on BM86-vaccinated and control cattle. The cattle proteins showed to be under represented in *R. annulatus* when compared to *R. microplus*, suggesting that *R. annulatus* ticks ingested less blood, a difference that increased when feeding on vaccinated cattle, probably reflecting the effect of antibody-BM86 interaction on this process. Tick protein degradation machinery was under represented in *R. annulatus* when compared to *R. microplus*. BM86 mRNA and protein levels were similar in both tick species, suggesting that lesser protease activity in *R. annulatus* results in more efficient antibody-antigen interactions and higher vaccine efficacy. These results indicated that not only genetic differences, but also physiological factors influence tick vaccine efficacy. The second part of the study compared proteomes of *R. microplus* ticks fed on BM86-vaccinated and subolesin vaccinated deer to investigate the differences in the mechanisms of effect of two
tick protective antigens. The results showed that although both vaccines reduce tick feeding and reproduction, they act through different protective mechanisms. The main protective effect of the BM86 vaccine is the effect on BM86 function and levels, which alters tick gut structure and function and consequently blood digestion and reproduction. For the SUB vaccine, the results showed that vaccination reduces protein levels and affects its function as a transcriptional regulator of genes involved in several BPs important for tick feeding and reproduction. These results have important implications for tick vaccine research, supporting the study of the molecular mechanisms associated with vaccine protection to combine antigens that act through different protective mechanisms to increase vaccine efficacy.

Finally, an overview of most relevant proteomics studies focused on tick-host, tick-pathogen interactions, tick vaccine study, tick biology and evolution presented in Chapter VI served as a general discussion to a work presented in this thesis. This work emphasized the current limitations of tick proteomics studies and suggested further approaches that should lead to an integrated, systems biology method holding a key to understanding the molecular mechanisms at the tick-host and tick-pathogen interfaces and the discovery of novel target molecules for the control of tick infestations and tick-borne diseases.

**Resúmen**

En la era post-genómica, la proteómica incluye estrategias prometedoras para caracterizar las interacciones dinámicas que no pueden ser analizados mediante un enfoque genómico o transcriptómico. La aplicación de la proteómica en la investigación de garrapatas se encuentra en sus comienzos y es necesario continuar en la misma para alcanzar el pleno potencial de esta tecnología. Debido a esto, el presente trabajo se ha centrado el desarrollo de nuevas estrategias proteómicas y su aplicación para profundizar en la comprensión de las interacciones garrapata-hospedador-patógeno. En el transcurso de esta tesis doctoral, los estudios de biología de sistemas, basados principalmente en aproximaciones proteómicas, complementados con transcriptómica y estudios inmunológicos, que se han llevado a cabo han permitido la identificación de proteínas expresadas diferencialmente en diferentes especies de garrapatas que han conducido a una mejor comprensión tanto de la biología de la garrapata, como de su desarrollo y la evolución.

La revisión de los resultados actuales sobre proteómica cuantitativa aplicada al estudio de garrapatas presentada en el capítulo I indica la necesidad de integrar las investigaciones transcriptómicas y proteómicas utilizando un enfoque desde la biología de sistemas. Este enfoque holístico para el estudio de las interacciones garrapata-hospedador-patógeno es fundamental para
avanzar en la comprensión de las interacciones tanto garrapata-patógeno como garrapata-
hospedador.

Debido al limitado número de estudios que se han llevado a cabo aplicando la proteómica a la
investigación de garrapatas, tuvieron que desarrollarse nuevos protocolos para la extracción y el
análisis de proteínas durante el curso de este trabajo (capítulo II), poniendo especial énfasis en el
trabajo con muestras de garrapatas repletas recogidas después de alimentarse sobre hospedadores
vertebrados.

La aplicación de la proteómica al estudio de la biología, el desarrollo y la evolución de las
garrapatas se presenta en el Capítulo III. En la primera parte se recoge el primer estudio realizado
sobre el proteoma de las especies más importantes de garrapatas del género *Amblyomma*. Estas
garrapatas son vectores relevantes de patógenos zoonóticos en todo el mundo. Sin embargo, hay
muy poca información disponible en su genoma, transcriptoma y proteoma. Los resultados
obtenidos proporcionan nueva información para entender la fisiología, el desarrollo y la
evolución de estas especies de garrapatas. La filoproteómica llevada a cabo utilizando la
secuenciación de novo de proteínas se ensayó como un nuevo enfoque para el análisis
filogenético de las especies de garrapatas en las que los datos de secuencia son un factor
limitante. En la segunda parte del capítulo, se recogen las investigaciones realizadas sobre larvas
no alimentadas de *D. reticulatus* cuyos resultados apoyan el uso de la secuenciación paired-end
para el análisis de los datos de transcriptómica. Por otra parte, la aproximación PIT (proteomics
informed by transcriptomics) mostró una alta eficiencia en la identificación de proteínas
utilizando una base de datos creada a partir de los datos de transcriptómica para buscar los
espectros de masas, confirmando ser una buena alternativa para organismos con poca información
de secuencia disponible, tales como *D. reticulatus*. Los procesos que se identificaron como más
activos fueron procesos metabólicos y celulares que implican reacciones enzimáticas, lo que
sugiere que las garrapatas son muy activas durante esta etapa de la vida. El estudio también
reveló que las respuestas de estrés estaban activas en *D. reticulatus* larvas sin alimentar,
probablemente para contrarrestar el efecto negativo de la temperatura y otras condiciones de
estrés tales como la infección por *Rickettsia* y favoreciendo la adaptación de la garrapata a las
condiciones ambientales para aumentar su supervivencia. Estos resultados muestran los
mecanismos que se han desarrollado en las garrapatas *D. reticulatus* para sobrevivir bajo
condiciones de estrés y sugieren que se conservan en las especies de garrapatas duras.

Hay muy pocos estudios que han investigado el efecto de un hospedador específico sobre la
biología garrapatas. La existencia de factores del hospedador que afectan a la alimentación y la

151
reproducción de la garrapata, así como el impacto sobre la evolución y la ecología del vector se puso al descubierto en estudio realizado comparando los proteomas de *Rhipicephalus microplus* alimentados sobre ganado bovino o sobre el venado cola blanca (WTD) (capítulo IV). Las garrapatas *R. microplus* alimentadas en bovinos presentaban una sobre-expresión de proteínas implicadas en la digestión de la sangre y la reproducción comparadas con las garrapatas alimentadas sobre WTD, correlacionándose con valores mayores respecto al número de garrapatas, peso y rendimiento reproductivo observados. Estas garrapatas probablemente ingieren más sangre, como se refleja en mayores pesos de las garrapatas y mayor cantidad de alfa-2-macroglobulina e inmunoglobulina. Sin embargo, los niveles más altos de proteínas del hospedador, tales como hemoglobina, haptoglobina y albúmina, presentes en las garrapatas alimentadas en WTD, se correlacionan con una menor maquinaria para la digestión de la sangre presente en estas garrapatas. Los niveles de fibrinógeno en las garrapatas alimentadas en bovinos fueron mayores y pueden reflejar la respuesta del hospedador ante una infección explicando las menores infestaciones de garrapatas que se observan en ciervos, posiblemente debido a una co-evolución garrapata-hospedador.

El potencial de la proteómica para caracterizar los mecanismos de protección de las vacunas de garrapatas se presenta en el capítulo V. La eficacia de la vacunación con BM86 y otros antígenos de garrapata ha demostrado ser siempre superior contra *R. annulatus* que contra *R. microplus*, sugiriendo que la genética de la garrapata y/o factores fisiológicos pueden afectar la eficacia de la vacuna. Para investigar esto, se compararon los proteomas de hembras de garrapatas de las especies *R. microplus* y *R. annulatus* después de alimentarse sobre ganado bovino control y vacunado con BM86. *R. annulatus* mostró tener una menor cantidad de proteínas del hospedador bovino en comparación con *R. microplus*, sugiriendo que las garrapatas *R. annulatus* ingirieron una menor cantidad de sangre. Esta diferencia aumenta cuando se alimentan sobre animales vacunados, reflejando probablemente el efecto de las interacciones anticuerpo-BM86 en este proceso. La maquinaria de degradación protéica se mostró disminuida en *R. annulatus* respecto a *R. microplus*. Tanto los niveles de ARNm como de proteína para BM86 fueron similares en ambas especies de garrapatas. Estos resultados sugieren una menor actividad de proteasas en *R. annulatus*, interacciones anticuerpo-antígeno más eficientes y una mayor eficacia de la vacuna indicando que no sólo las diferencias genéticas, sino también factores fisiológicos influyen en eficacia de la vacuna. En la segunda parte del estudio se compararon los proteomas de *R. microplus* alimentados sobre ciervos vacunados con BM86 y con subolesina para investigar los mecanismos responsables del diferente efecto de dos antígenos protectores de garrapatas. Los
resultados mostraron que, aunque ambas vacunas reducen la alimentación de la garrapata y la reproducción, actúan a través de diferentes mecanismos de protección. El principal efecto protector de la vacuna BM86 es el efecto sobre la función y los niveles de BM86, que altera la estructura y la función intestinal garrapata y, en consecuencia, la digestión de la sangre y la reproducción. Para la vacuna SUB, los resultados mostraron que la vacunación reduce los niveles de proteína y afecta a su función como un regulador transcripcional de genes implicados en varios procesos biológicos importantes para la alimentación de la garrapata y su reproducción. Estos resultados tienen implicaciones importantes para la investigación en vacunas de garrapata, apoyando el estudio de los mecanismos moleculares asociados a la protección de la vacuna para combinar antígenos que actúan a través de diferentes mecanismos de protección para aumentar la eficacia de la vacuna.

Por último, en el Capítulo VI se presenta una revisión general de los principales estudios de proteómica realizados en garrapatas y centrados en las interacciones garrapata-hospedador, garrapata-patógeno, vacunas y biología y evolución. Este trabajo sirve como discusión general del trabajo presentado en esta tesis. Se hace hincapié en las limitaciones actuales de los estudios de proteómica en garrapatas y se sugieren nuevos enfoques que deberían dar lugar a un método integrado basado en la biología de sistemas clave para la comprensión de los mecanismos moleculares de las interacciones garrapata-hospedador y garrapata-patógeno y al descubrimiento de nuevas moléculas diana para el control de infestaciones por garrapatas y las enfermedades transmitidas por éstas.
Illustrations in this work:

Hypothesis and objectives, Introduction and Chapter III b cover:
POSTICK Spring Workshop, Budweis, Czech Republic

Chapter VI:
photo by Dr. Andras Tuberg- POSTICK Bayer Course, Monheim, Germany.

Chapters II, III, III a, IV and V:
illustrations done by Marina Popara