

University of Nevada, Reno

INVESTIGATION OF THE MIALOME OF *IXODES SCAPULARIS* REplete TICKS
AND UNDERSTANDING THE POPULATION GENOMIC STRUCTURE OF *IXODES*
PACIFICUS TICKS

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biochemistry

by

Jeremiah Reyes

Dr. Monika Gulia-Nuss/Dissertation Advisor

May , 2022

© 2022

Jeremiah B. Reyes

All Rights Reserved



THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

Jeremiah B. Reyes

entitled

**Investigation of the Mialome of *Ixodes scapularis* replete ticks and
understanding the population genomic structure of *Ixodes pacificus*
ticks**

be accepted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

Dr. Monika Gulia Nuss
Advisor

Dr. Richard Tillett
Committee Member

Dr. Mike Teglas
Committee Member

Dr. Claudia Rückert
Committee Member

Dr. Andrew Nuss
Graduate School Representative

David W. Zeh, Ph.D., Dean
Graduate School

May, 2022

Abstract

Ticks are blood-feeding arthropod ectoparasites that transmit disease-causing pathogens to humans and animals worldwide. Research on vaccine development to protect humans, companion animals, and livestock from ticks and tick-transmitted pathogens has accelerated by using proteomic and transcriptomic analyses. Immunization of hosts with targeted anti-tick vaccines would ideally lead to a reduction in tick numbers and prevent transmission of tick-borne pathogens. Vaccines using cattle tick antigens (anti-tick vaccine) have proven to be cost-effective and environmentally friendly for the control of cattle tick infestations and pathogen infection and transmission. However, new strategies are needed to identify tick protective antigens for development of vaccines for tick species of human significance.

Ticks are obligatory blood feeders. Blood meals are stored and digested in the midguts. Blood digestion is complex, and many proteins are involved. These tick-derived proteins in the midgut may be useful for anti-tick vaccines. Therefore, midguts were examined using transcriptome and proteome of unfed, partially engorged, and replete female *Ixodes scapularis*, previously uncharacterized for this species. The function of identified midgut proteins varied from nutrient transportation, anti-coagulation, erythrocyte lysis, detoxification, lipid metabolism, and immunization. Functional characterization of promising antigen targets for an anti-tick vaccine resulted in several putative candidates for further analysis (Chapter 2). In Chapter 3, function of three highly expressed serine proteases (identified in chapter 2) in blood digestion was determined. In Chapter 3, we showed that the serine proteases are indeed active in blood digestion in fully

engorged females. These findings advance our understanding of tick digestive mechanisms and provide a panel for screening of vaccine candidates.

In addition, we began to determine the genetic structure of the primary vector for Lyme Disease in California, *Ixodes pacificus* (Chapter 4). Due to the broad host range of *I. pacificus*, especially large mammals and birds which have the capability of traveling long distances, it is proposed that the ticks will migrate and may establish in new areas. We collected ticks from 10 counties, prepped a restriction-associated DNA sequencing (RADseq) library, and conducted genome-wide single nucleotide polymorphism (SNP) analysis. Our data suggest two ancestral population of *I. pacificus* and high gene flow among populations. Our data also agrees with current pathogen prevalence data and suggests higher transmission rates in some counties than others. Once complete, these data will greatly increase our knowledge of tick migration and vector-pathogen relationships.

Dedication

To my loving parents, Juan Bautista and Lupe Del Carmen Reyes. Without my hard-working and amazing parents, I would never have thought to pursue my dreams, or even be here in the first place. And to my entire family for always being there when I needed them the most.

Acknowledgements

I would like to thank my advisor, Dr. Monika Gulia-Nuss. You are one of the most special people I know. Joining your lab was like joining your family, you took me in and gave me unlimited support, both intellectually and as a friend. Your love for science and the enthusiasm grew on me. You are one of the most knowledgeable people I know and meeting you for the first time made me realize how much of an inspiration you would be in the pursuit of my dreams.

I would like to thank Dr. Arvind Sharma for his support as well, I learned so much from you and the tireless hours you were with me to help me. I would like to thank the rest of my lab members for always listening and sharing their knowledge.

I would like to thank the members of my committee, who were there to listen to my progress and give me sound advice to get through the mental blockades I frequently had.

Finally, I would like to thank all my friends I have met along this journey. Sharing ideas, pain of being in a PhD, and making it through many milestones together made it bearable.

Table of Contents

Abstract.....	i
Dedication.....	iii
Acknowledgements.....	iv
Table of Contents.....	v
List of Tables.....	vii
List of Figures.....	viii
 Chapter 1: Review of Literature	
1.1 Part 1 - Introduction.....	1
1.2 Tick control measures.....	5
1.3 Tools for identifying anti-tick vaccine candidates.....	8
1.4 Objectives - Part 1.....	9
1.5 Gaps in literature for Part 1.....	9
1.6 <i>Ixodes pacificus</i>	10
1.7 <i>Ixodes pacificus</i> population structure.....	13
1.8 Objectives for Part 2.....	14
1.9 Gaps in literature for Part 2.....	14
1.10 References.....	15
 Chapter 2: A multi-omics approach towards the identification of novel midgut targets for control of the <i>Ixodes scapularis</i>	
2.1 Abstract.....	33
2.2 Introduction.....	34
2.3 Materials and Methods.....	36
2.4 Results.....	43
2.5 Discussion.....	53
2.6 References.....	62
2.7 Figures and Tables.....	77

Chapter 3: Blood Digestion by Trypsin-Like Serine Proteases in the Replete Lyme Disease Vector Tick, *Ixodes scapularis*

3.1 Abstract.....	105
3.2 Introduction.....	106
3.3 Materials and Methods.....	108
3.4 Results.....	113
3.5 Discussion.....	117
3.6 References.....	121
3.7 Figures and Tables.....	127

Chapter 4: Population Genetic structure of the Western black-legged tick, *Ixodes pacificus*, in the state of California

4.1 Abstract.....	134
4.2 Introduction.....	135
4.3 Materials and Methods.....	136
4.4 Results.....	141
4.5 Discussion.....	143
4.6 References.....	146
4.7 Figures and Tables.....	152

Conclusions and Future Directions

Future Research Directions.....	158
Concluding Remarks.....	158

List of Tables

Chapter 2: A multi-omics approach towards the identification of novel midgut targets for control of the *Ixodes scapularis*

Table 1. Statistics of reads from Illumina sequencing data at all timepoints.....	92
Table 2. Top 5 up or downregulated differentially expressed transcripts per timepoint.....	93
Table 3. Identification of differentially expressed immunity-related transcripts as compared to UF transcripts.....	94
Table 4. Co-expressed transcripts and proteins at 1 D PBM.....	98
Table 5. Co-expressed transcripts and proteins at 7 D PBM.....	100
Table 6. Primer sequences used with T7 promoter for RNAi. Transcript ID, sequence, and product size of transcripts.....	103
Table 7. Knockdown results of transcript and protein candidates via RNA interference.....	104

Chapter 3: Blood Digestion by Trypsin-Like Serine Proteases in the Replete Lyme Disease Vector Tick, *Ixodes scapularis*

Table 1. A list of primer pairs and the annealing temperature used for each primer pair.....	132
--	-----

Chapter 4: Population Genetic structure of the Western black-legged tick, *Ixodes pacificus*, in the state of California

Table 1. Summary of statistics across samples.....	155
Table 2. Identification of private alleles per county.....	155
Table 3. Genetic variation among populations of <i>Ixodes pacificus</i> collected from different counties in the state of California.....	156
Table 4. Genetic variation between each individual subpopulation of <i>Ixodes pacificus</i> from counties in California.....	156
Table 5. Nucleotide diversity at variant and all positions within each population..	157
Table 6. Identification of samples positive for <i>Borrelia burgdorferi</i>	157

List of Figures

Chapter 2: A multi-omics approach towards the identification of novel midgut targets for control of the *Ixodes scapularis*

Figure 1. Blood feeding and digestion timecourse.....	77
Figure 2. Venn diagram of differentially expressed transcripts identified at all timepoints and profile of differentially expressed transcripts identified.....	77
Figure 3. Identification of differentially expressed proteolytic enzyme transcripts as compared to UF.....	78
Figure 4. Identification of differentially expressed immunity-related transcripts as compared to UF transcripts.....	79
Figure 5. Time series hierarchical clustering from the midguts of UF and fed <i>I. scapularis</i> ticks collected at different blood-feeding timepoints.....	80
Figure 6. qRT-PCR relative transcript expression selected for validation of RNA-seq dataset.....	81
Figure 7. Profile of differentially expressed proteins identified.....	82
Figure 8. KEGG pathway analysis of upregulated proteins at 1 and 7 D PBM.....	83
Figure 9. KEGG pathway analysis of downregulated proteins at 1 PBM.....	84
Figure 10. KEGG pathway analysis of downregulated proteins at 1 PBM.....	85
Figure 11. Molecular function gene ontology analysis of upregulated proteins at 1 D PBM.....	86
Figure 12. Molecular function gene ontology analysis of upregulated proteins at 7 D PBM.....	87
Figure 13. Co-expression analysis of transcriptome and proteome at 1 D PBM when compared to UF samples.....	88
Figure 14. Co-expression analysis of transcriptome and proteome at 1 D PBM when compared to UF samples.....	89
Figure 15. Selection of RNAi targets and qRT-PCR for knockdown confirmation.....	90
Figure 16. Knockdown efficiency of RNAi targets.....	90
Figure 17. Visual phenotype of dsRNA injected ticks.....	91

Chapter 3: Blood Digestion by Trypsin-Like Serine Proteases in the Replete Lyme Disease Vector Tick, *Ixodes scapularis*

Figure 1. Transcript expression of proteases in the midgut of adult female <i>Ixodes scapularis</i>	127
Figure 2. I BApNA assay of trypsin activity in <i>I. scapularis</i> developmental stages.....	128
Figure 3. Effect of serine protease knockdown on adult female <i>I. scapularis</i> feeding, blood digestion, survival, and reproduction.....	129
Figure 4. Hemoglobin degradation by <i>I. scapularis</i> midgut extract.....	130
Figure 5. Revised model of blood digestion in replete females.....	131
Chapter 4: Population Genetic structure of the Western black-legged tick, <i>Ixodes pacificus</i>, in the state of California	
Figure 1. California map where ticks were collected or received for RADseq processing.....	152
Figure 2. Phylogenetic analysis from haplotype data suggests two common ancestors.....	153
Figure 3. Genome-wide allelic frequency data suggests one major and two moderate migration patterns.....	154
Figure 4. Positive samples using <i>Borrelia burgdorferi</i> ospA specific primers.....	154

Chapter 1 – Literature Review

1.1 Part 1- Introduction

Ticks are one of the most important groups of arthropod disease vectors due to their ability to transmit a variety of pathogens. These ectoparasites are solely dependent on acquiring bloodmeals which are essential for survival, growth, molting, proliferation, and oogenesis (Nava et al., 2009). During the process of blood feeding and digestion, many changes occur in ticks at a morphological, physiological, and molecular level (Franta et al., 2010). The first objective for this work was to understand the physiological process of blood digestion.

Tick Phylogeny

There are 896 tick species currently recognized (Guglielmone et al., 2010) classified into three families: the *Ixodidae* (hard ticks), the *Argasidae* (soft ticks), and the monotypic *Nuttalliellidae* (Horak et al., 2002). The *Ixodidae* are further divided into two lineages: the *Prostriata*, which contains only the genus *Ixodes*, and the *Metastricata*, which contains all other hard tick genera (Keirans, 2009). The ixodid tick fauna consists of 241 species in the genus *Ixodes* and 442 species in the genera *Amblyomma*, *Anomalohimalaya*, *Bothriocroton*, *Cosmiomma*, *Dermacentor*, *Haemaphysalis*, *Hyalomma*, *Margaropus*, *Nosomma*, *Rhipicentor* and *Rhipicephalus* in the family *Ixodidae*, with the former genus *Boophilus* now classified as a subgenus of the genus *Rhipicephalus*. The argasid tick fauna comprises 183 species in four genera, namely *Argas*, *Carios*, *Ornithodoros* and *Otobius* in the family *Argasidae*. The family *Nuttalliellidae* is represented by the monospecific genus *Nuttalliella* (Burger et al., 2012).

Ticks as disease vectors

Ticks transmit many pathogens making tick-borne diseases common in both humans and other animals (Edlow 2008, Herrmann et al., 2014). Ticks can transmit a range of viruses, bacteria, protozoans, and fungi. The most medically important pathogen in the United States is *Borrelia burgdorferi*, the causative agent for Lyme disease (LD). The Center for Disease and Control suggests approximately 470,000 new cases of LD per year in the United States (Porter et al., 2019). The primary vector for this pathogen in the Eastern and Midwestern United States is the black-legged tick, *Ixodes scapularis* (De la Fuente., 2008) and on the west coast of the USA, *I. pacificus*. Other human pathogens commonly vectored by *Ixodes* ticks are *Anaplasma phagocytophilum*, *Babesia microti*, *Ehrlichia spp.*, and Powassan virus (De la Fuente., 2008). Although LD is the leading vector-borne zoonotic disease in the United States, co-infection with *B. microti* and *A. phagocytophilum* can occur in patients infected with *B. burgdorferi* (Krause et al., 1996, Belongia, 2002). Humans co-infected with these pathogens may be at risk of a harsher, more debilitating form of Lyme, *Babesiosis* or *Anaplasmosis* (Tokarz et al., 2010).

Ixodes scapularis

Ixodes scapularis, commonly known as the deer tick or black-legged tick, is a three-host tick and the main vector of Lyme disease in the United States. It is distributed all along the East coast into Canada and is established as far as the Midwest (Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin) and southernly to Texas. A blood meal from different hosts is required at each developmental stage: larva, nymph, and adult. Immature life stages (i.e., larvae and nymphs) have a broad host range, including rodents, birds, lagomorphs, primates, and

ungulates (Sonenshine et al., 2009). The primary hosts for adult feeding are mammals, most commonly white-tailed deer (*Odocoileus virginianus*) (Telford et al., 1988, Luttrell et al., 1994). Regarding their role as vectors, larval *I. scapularis* acquire *B. burgdorferi* during feeding on an infected white-footed mouse (*Peromyscus leucopus*), the *B. burgdorferi* permissible host (Wilson et al., 1985, Spielman et al., 1985). After molting to the nymphal stage, transstadial infected ticks feed on a broad range of animals, which includes mice, becoming a new reservoir for the host and continuing the cycle. After molting from nymphs, female adults feed exclusively on larger mammals such as white-tailed deer, which are not competent hosts for *B. burgdorferi* (de la Fuente et al., 2008). Therefore, the larval and nymphal feedings are crucial to maintaining the pathogen in the environment (Piesman, 1979, Kilpatrick et al., 2017).

Tick host-seeking, feeding, and blood digestion

Although ticks feed on variety of hosts, the primary hosts for *I. scapularis* are white-tailed deer (*Odocoileus virginianus*) and the white-footed mouse (*Peromyscus leucopus*) (Wilson et al., 1985, Spielman et al., 1985).

Blood feeding is critical for ticks as it is their only source of nutrition (Sojka et al., 2013). *Ixodes scapularis* larvae attach and feed for 3-5 days, nymphs for 5-7 days, and adult females feed for approximately 7-10 days. Adult female feeding consists of a slow feeding period taking 6–9 days, which is followed by a rapid engorgement of 12–24 h, accounting for about two-thirds of the total blood volume ingested before detachment. Only mated females can rapidly engorge, indicating that an unknown mechanism, possibly hormonally mediated, triggers rapid feeding (Sonenshine, 1991).

The tick midgut is responsible for digestion and comprises a ventriculus (stomach) and several pairs of highly branched caeca filling all regions of the female body. The digestion of a blood meal is a complex process that requires the coordinated expression of many genes involved in nutrient processing and absorption, detoxification mechanisms, and defense against the pathogens that are ingested with the vertebrate blood meal (Sojka et al., 2013, Horn et al., 2009, Mulenga et al., 2011). Most tick blood digestion work has focused on the slow feeding phase (5-6 days post attachment). Models developed from these studies suggests that hemoglobinolysis is most efficient in the acidic range (pH 3.5–4.5), corresponding to the pH environment of the digestive vesicles. Cathepsins D, L (pH optimum 3.5) and Legumain/asparaginyl endopeptidase have been proposed to be the major proteolytic enzymes at this stage that cleave hemoglobin into large fragments which are further cleaved by cathepsin B and C into dipeptides and amino acids (Sojka et al., 2013). Serine proteases in the gut lumen are also important for hemoglobin degradation in fully fed females (Reyes et al., 2020)

Furthermore, several mechanisms to deactivate the innate and adaptive immune systems of the vertebrate blood must operate to protect the midgut epithelia from immune-mediated injury. Additional mechanisms that might be significant are proteolytic inactivation caused by arthropod digestive enzymes and temperature-dependent immune attenuation caused by the decrease in blood temperature that occurs once the arthropod detaches from the host (Wikel et al., 1994).

1.2 Tick control measures

Acaricides

Acaricides are pesticides commonly used in killing ticks and mites. Some of these are as follows: Chlorinated hydrocarbons (dichlorodiphenyltrichloroethane) interfere with nerve conduction but have long lifespan and high toxicity. Organophosphorus compounds are esters of phosphoric acid. They have a shorter residual effectiveness than chlorinated hydrocarbons, and have a high risk of causing acute toxicity in livestock, ticks have also become resistant to these agents. Carbamates are carbamic acid esters closely resembling organophosphates, are more toxic to mammals than organophosphates and are much more expensive. Usage of very high toxic levels of other chemical agents such as permethrin, formamidines, and avermectins proves that ticks are quickly becoming immune to these biological agents (Eisen and Stafford, 2021).

Personal protective measures

Personal protective measures are excellent barriers for preventing tick bites. These include using synthetic and natural repellents which are applied to skin and clothing. Permethrin can be used as a treatment for clothing, shoes, and gear and remains on articles for up to three washes (Mencke, 2006). Environmental Protection Agency (EPA)-registered insect repellents which contain N,N-diethyl-methyl-meta-toluamide (DEET), picaridin, ethyl butylacetylaminopropionate (EBAAP), oil of lemon eucalyptus (OLE), p-menthane-diol (PMD), or 2-undecanone are effective in repelling ticks (Buchel et al., 2015).

Vaccine development

A major tool for tick control has been application of acaricides, but problems including acaricide resistance, environmental contamination, and (in certain species of ticks which feed on livestock) residues of acaricides in dairy and meat products have prompted development of alternatives (Graf et al., 2004). Other tick management methods include bioactive agents such as entomopathogenic fungi. They have been tested in several species of ticks (Pirali-Kheirabadi et al., 2007, Hornbostel et al., 2005, Kaaya et al., 1996) however, *Metarhizium anisopliae* and *Beauveria bassiana* have shown great environmental instability and target other species. These entomopathogenic fungi only worked against *I. scapularis* when combining the reduction of the deer population significantly, which was very difficult to accomplish (Williams et al., 2018).

Vaccination is a very attractive alternative to acaricides to reduce tick populations and pathogen infection as it is more environmentally friendly than acaricides (Brossard, 1998, de la Fuente et al., 1998, 2007a, 2007b; Almazan et al., 2005). The first incidence of naturally acquired resistance against ticks dates back to 1939 when guinea pigs with multiple infestations of ticks, over time, showed resistance against tick bites (Trager, 1939). Efforts have since been made to identify the potential candidate for immunization (Li et al., 2019). The goal of arthropod vector vaccines is the control of vector infestations and subsequent tick-borne disease infections. The effect of vaccines on tick borne pathogens could lead to 1) population reduction of vectors, especially if ticks only partially feed as they won't initiate feeding again, leading to decreased cases of hosts acquiring tick-borne pathogens, 2) reducing the capacity for pathogen transmission due to host immune response, and 3) a combination of both (Merino et al., 2013). There has been success with

one “concealed antigen” based vaccine, which targets midgut-specific glycoproteins in the cattle tick, *Rhipicephalus (Boophilus) microplus* (Willadsen et al., 1986). A commercially available vaccine developed against the *R. microplus* gut antigen Bm86 induces an immunological response in the vertebrate host against tick infestation. Vaccinated host had a reduction in the weight of engorging female ticks, in the weight of the eggs laid and also in *R. microplus* viable eggs percentage (Bastos et al., 2010). Two vaccines using recombinant Bm86 were later registered in Latin American countries (Gavac) and Australia (Tick GARD), demonstrating the feasibility of controlling tick infestation through immunization of host against tick antigens (Martinez-arzate et al., 2019, Trentelman et al., 2019, Odongo et al., 2007).

A current limitation for anti-tick vaccine development is the selection of suitable antigens. Current candidate gene identifications focus on proteins present in saliva and therefore fall into the category of “exposed antigens.”. However, development of a concealed antigen-based vaccine is the most promising strategy as it has already been shown to work successfully in cattle ticks because the host and parasite will not have evolved immunological interactions related to these antigens (Gao et al., 2009).

Recently, Sajid et al. (2021) showed proof-of-concept for an mRNA-based vaccine using 19 *I. scapularis* proteins (19ISP) encapsulated in lipid nanoparticles to inhibit degradation and facilitate *in vivo* delivery. They showed that immunization with 19ISP provides robust tick immunity in guinea pigs, including early erythema after tick placement on the animals and rapid tick detachment, along with severely impaired tick feeding and low engorgement weights. However, these antigens were not effective in mice (Sajid et al., 2021). These candidates also led to an increased immune response in guinea pigs after 96

h of immunization, time when tick would already be feeding. Therefore, new targets are needed to induce a robust immune response to ticks soon after attachment.

1.3 Tools for identifying anti-tick vaccine candidates

Systems based approaches for understanding tick biology

The first genome assembly of a medically important acarid species, *I. scapularis*, paved the way to system-based approaches for understanding tick biology (Gulia-Nuss et al., 2016). Since publication of this genome, several other tick genome projects (Cramaro et al., 2017; Jia et al., 2020), as well as tick cell line genomes (Miller et al., 2018) were completed. Tick comparative genomics has now begun with the availability of seven genomes and many *de novo*-based transcriptomes (Josek et al., 2018, Schwarz et al., 2013, Cramaro et al., 2015, Narasimhan et al., 2002, Egekwu et al., 2014, Francischetti et al., 2005, Charrier et al., 2018, Perner et al., 2016, Kozelkova et al., 2021, Kotsyfakis et al., 2015). These tools allow identification of new targets for anti-tick vaccines.

Functional genetic tools for the characterization of putative targets

RNA interference is a nucleic acid-based reverse genetic approach that involves disruption of gene expression to determine the function of a particular gene (Cerutti et al., 2003). Post-transcriptional gene silencing through dsRNA induction has been found in all eukaryotes studied so far, including ticks, and it is a widely used tool for functional studies (Zhou et al., 2006). Functional genetics in ticks is a valuable tool for understanding gene function, the tick-host-pathogen interface, and can be used to screen for tick protective antigens to develop anti-tick vaccines (de la Fuente et al., 2007). Until very recently, this was the only genetic screening tool available for ticks. The clustered regularly interspaced short palindromic repeats (CRISPR), CRISPR-associated protein 9 (Cas9), has

revolutionized functional genetics studies in an array of organisms (Zhang and Reed, 2017). Sharma et al. (2022) has introduced the first successful gene disruption method using this novel genetic tool through two methods: embryo injection and Receptor-Mediated Ovary Transduction of Cargo (ReMOT Control). ReMOT delivers the Cas9 ribonucleoprotein complex directly to the germline from the hemocoel, which allows for targeted and heritable mutations which are generated by adult injection. This tool is a powerful addition to functional genetics for characterizing gene functions.

1.4 Objectives – Part 1

Use of multi-omics tools for identification of vaccine targets

In this study, we examined the blood digestion profiles of partially fed and replete females. Besides enhancing our basic understanding of the shift in proteases over the course of blood digestion and to identify genes/proteins, we sought to identify proteins that could be targeted for anti-tick vaccine (Chapter 2). Earlier studies suggested that the blood digestion in rapid phase or replete females involves both acidic environment of digestive vesicles and the alkaline environment of the gut lumen and have suggested serine proteases to be involved in blood digestion in replete females; however, no experimental data was available. Therefore, our second objective was to characterize three highly expressed serine proteases in replete females (Chapter 3).

1.5 Gaps in the literature

The current hemoglobinolytic model has been determined for partially engorged *Ixodes* females. Therefore, a full repertoire of genes involved in blood digestion in rapid feeding and after repletion has not been studied extensively yet. The research on replete females remained elusive because of the assumption that acquiring mRNA from replete

females is experimentally difficult due to the exceptional fragility of tissues and the high content of RNases (Sojka et al., 2013). Thus, it remained unclear whether digestion in detached females relies on enzymes synthesized during the slow-feeding period or whether newly synthesized ‘late’ isoenzymes are responsible for the post-feeding enzyme activities. The latter possibility was supported by the expression of cathepsin D and serine proteases at the very end of female feeding close to detachment from the host (rapid phase) (O’Donoghue et al., 2012). This is important because it may hold the key factors which may be used for controlling tick populations. Additionally, serine proteases are known to be extremely important for blood digestion in many animals. Therefore, our goal was to functionally characterize female-specific and blood-meal induced vitellogenic serine proteases identified in our omics data to understand their function in tick blood feeding success and blood digestion.

1.6 Part 2 - *Ixodes pacificus*

Ixodes pacificus

Ixodes pacificus is the primary vector of Lyme disease and other pathogens in western North America (Lane et al., 1994, Teglus and Foley, 2006, Krause et al., 2015). Their range is from British Columbia to Northern Mexico along the Pacific coast. Other populations have been reported in East Oregon, Nevada, Utah, and even Arizona (Olson et al., 1992, Kain, 1996, Kain et al., 1997). These ticks perpetuate enzootic transmission cycles together with a vertebrate reservoir host, such as mice, deer, lizards, and others (Lane et al., 2005). Seasonal activity for immature ticks (larvae and nymphs) in California are from January through October, and historical data suggest that Lyme disease cases in the Western US mimic this similar pattern. Merging the ecology, epidemiology, and

geographic distribution of this species is important for identifying better strategies for tick control by understanding tick movement, behavior, and pathogen presence in certain populations.

Pathogens and hosts associated with *Ixodes pacificus*

Ixodes pacificus, as noted above, is the main vector of *B. burgdorferi*, in the North American West and also transmits other pathogens of medical importance such as *A. phagocytophilum*, *Bartonella*, and *Ehrlichia* species (Holden et al., 2006). Grey squirrels (*Sciurus griseus*) and migratory birds, specifically order Passeriformes, maintain *Borrelia* in the environment (Castro & Wright, 2007). Larvae and nymphs of *I. pacificus* typically feed on Western fence lizards (*Sceloporus occidentalis*); however, these lizards have a highly lytic component in their blood which prevents sustained infection with *Borrelia* species (Lane et al., 2006). In addition to *B. burgdorferi*, other species such as *B. miyamotoi* and other lesser-known species such as *B. americana* and *B. bissettiae* are carried by *I. pacificus* (Salkeld et al., 2021). *Ixodes pacificus* can also carry *Anaplasma phagocytophilum*, the causative agent of anaplasmosis, and is maintained in the environment by several hosts which include the dusky-footed woodrat (*Neotoma fuscipes*) and tree squirrels (*Sciurus* sp.) (Nieto et al., 2009).

Bartonella is a genus of bacteria present in many arthropod vectors, including *I. pacificus*. Molecular analysis of *I. pacificus* showed a variety of *Bartonella* strains closely related to cattle *Bartonella* and several known human pathogenic *Bartonella* species and subspecies: *B. henselae*, *B. quintana*, *B. washoensis*, and *B. vinsonii* subsp. *berkhoffii*. This suggests that *I. pacificus* adults could be a source for *Bartonella* infections in humans (Chang et al., 2001). *Bartonella* is maintained in the environment by several rodent species:

ground squirrels (*Otospermophilus beecheyi*), deer mice (*Peromyscus maniculatus*), brush mice (*P. boylii*), Pinyon mice (*P. truei*) and dusky-footed woodrats (*N. fuscipes*). Additionally, *B. vinsonii* subsp. *arupensis* and *B. washoensis* have been described as zoonotic pathogens (Kosoy et al., 2003; Welch et al., 1999).

Pathogens from the genus *Ehrlichia* are the causative agents of two main human infections, human monocytic ehrlichiosis (*E. chaffeensis*) and human granulocytic ehrlichiosis (HGE) (unnamed species). In addition to human pathogens, *I. pacificus* also carries the species *E. equi*, the causative agent of equine granulocytic ehrlichiosis (EGE) (Rubel et al., 1998). *Ixodes pacificus* is capable of transmitting *E. equi* to previously uninfected horses experimentally; however, natural infection in the field has not been seen (Kramer et al., 1999; Richter et al., 1996). In addition to the above species, *Rickettsia* of the *E. phagocytophila* genogroup is carried by Ixodid ticks. It infects equines, ruminants, dogs, and humans.

Both ticks and their hosts have been shown to be infected with multiple pathogens at once. Specifically, coinfection has been recorded in deer mice, one of the main hosts for *I. pacificus* ticks may be coinfecting with and transmit Lyme borrelia along with other pathogens such as *A. phagocytophilum*; *B. microti*, the primary cause of babesiosis; or a tick-borne encephalitis virus. (Larson et al., 2018).

Life cycle and distribution

A county classification guideline was developed to accurately record *Ixodes pacificus* distributions in the US. In 1998, populations of *I. pacificus* ticks were established in 95 counties in six states with 56 of those counties being within California (Eisen et al., 2016); ticks are now reported in an additional 16 counties (total of 111 counties).

Distribution occurs mainly on the coast, where the humidity and temperature are higher than further inland its but the range is slowly expanding (Dennis et al., 1998). *Ixodes pacificus* distribution models suggest a potential expansion along the Oregon-Washington border with suitable habitats in 11 additional counties, representing a 12% possible increase in established counties (Hahn et al., 2016).

1.7 *Ixodes* population structure

Population genetic studies give insight into the dispersal patterns of ticks, including their direction of movement, distance, and possible factors (geographical, host availability, environmental) which may influence movement. This can be achieved with the estimation of gene flow between tick populations (McCoy, 2008, Araya-Anchetta et al., 2015). Movement of ticks is limited due to lack of wings, and as such, dispersal of ticks is reliant on the movement of its hosts (Falco and Fish., 1991, McCoy et al., 2001). The information for population structure of *I. pacificus* is currently minimal. Kain et al. (1997) conducted the first (of only two) genetic population structure. Allozyme data provided an equivocal picture of the population structure and biogeography of *I. pacificus* (Kain et al., 1997, Kain et al., 1999). Eleven of 12 loci sampled across the range of *I. pacificus*, including Utah, exhibited little genetic differentiation. In contrast, the twelfth locus, glucose-6-phosphate isomerase (GPI), showed a pattern of genetic differentiation that suggested either an adaptive cline or secondary contact along a broad zone in central and southern California. This zone demarcated a northern group of populations (Vancouver, Oregon, and Alameda co, CA) from a southern group (Mendocino, Santa Clara, San Diego, and Utah). Kain et al. (1997) recognized that either a rapid range expansion or high rates of gene flow could explain the pattern of allozyme variation. However, they proposed that the most plausible

explanation for the pattern of allozyme variation in *I. pacificus* was that it had a high rate of gene flow. However, another study by the same group (Kain et al., 1999) using mitochondrial DNA suggested little genetic variation except between a geographically isolated Utah locality and all other localities (same as in Kain et al., 1997). A study for *I. scapularis* suggested that reliance on minimal numbers of genetic markers such as allozymes may lead to incorrect identification of population structure (Sakamoto et al., 2014), therefore, a population structure using genome-level markers is needed to understand the expansion of this species.

1.8 Objectives – Part 2

We intended to develop the population structure of *I. pacificus* using a genome-wide single nucleotide polymorphism (SNP) approach using a Restriction site associated DNA (RAD) sequencing (RADseq). Our objective was to determine structure of *I. pacificus* in California and to test whether there are northern and southern group of populations as predicted by allozyme data (Kain et al., 1997). We will also use this dataset to identify which pathogens are associated with different populations throughout the state of California.

1.9 Gaps in the literature

Minimal data is available on the genetic structure of *I. pacificus*. From the technology available at the time, the mtDNA data was not able to differentiate haplotypes within the continuous portion of the range (Kain et al., 1999). The authors suggested that this probably is due to failures in current gene flow models to reflect adequately the complexity of evolutionary processes underlying the geographical differentiation of mtDNA haplotypes in *I. pacificus*. The authors recommended that it will be crucial to continue to develop

realistic gene flow models that can be applied in empirical studies because many systematic analyses are begun with little or no ecological or genetic data available on the organism being investigated (Kain et al., 1999). No other studies have attempted to resolve the population structure of *I. pacificus* in past two decades.

1.10 References

Almazán, Consuelo, Katherine M. Kocan, Edmour F. Blouin, and José de la Fuente. Vaccination with recombinant tick antigens for the control of *Ixodes scapularis* adult infestations. *Vaccine* 23, no. 46-47 (2005): 5294-5298.

Andreotti, Renato, Marisela S. Pedroso, Alexandre R. Caetano, and Natália F. Martins. Comparison of predicted binders in *Rhipicephalus* (*Boophilus*) *microplus* intestine protein variants Bm86 Campo Grande strain, Bm86 and Bm95. *Revista Brasileira de Parasitologia Veterinária* 17, no. 2 (2008): 93-98.

Ayllon, Nieves, Margarita Villar, Ruth C. Galindo, Katherine M. Kocan, Radek Šíma, Juan A. Lopez, Jesus Vazquez et al. Systems biology of tissue-specific response to *Anaplasma phagocytophilum* reveals differentiated apoptosis in the tick vector *Ixodes scapularis*. *PLoS genetics* 11, no. 3 (2015): e1005120.

Barker, S.C. and Murrell, A. (2004) Systematic and Evaluation of Ticks with a List of Valid Genus and Species Names. *Journal of Parasitology*, 129, 15-36.

<http://dx.doi.org/10.1017/S0031182004005207>

Bastos, Reginaldo G., Massaro W. Ueti, Donald P. Knowles, and Glen A. Scoles. The *Rhipicephalus* (*Boophilus*) *microplus* Bm86 gene plays a critical role in the fitness of ticks fed on cattle during acute *Babesia bovis* infection. *Parasites & vectors* 3, no. 1 (2010): 1-11.

Belongia EA. Epidemiology and impact of coinfections acquired from *Ixodes* ticks. *Vector Borne Zoonotic Dis.* 2002;2(4):265-73.

Burger TD, Shao R, Beati L, Miller H, Barker SC. Phylogenetic analysis of ticks (Acari: Ixodida) using mitochondrial genomes and nuclear rRNA genes indicates that the genus *Amblyomma* is polyphyletic. *Mol Phylogenet Evol.* 2012 Jul;64(1):45-55. doi: 10.1016/j.ympev.2012.03.004. Epub 2012 Mar 17. PMID: 22465402.

Brossard, M. "The use of vaccines and genetically resistant animals in tick control. *Revue scientifique et technique-Office international des épizooties* 17 (1998): 188-193.

Büchel, Kerstin, Juliane Bendin, Amina Gharbi, Sibylle Rahlenbeck, and Hans Dautel. Repellent efficacy of DEET, Icaridin, and EBAAP against *Ixodes ricinus* and *Ixodes scapularis* nymphs (Acari, Ixodidae). *Ticks and tick-borne diseases* 6, no. 4 (2015): 494-498.

Carr AL, Mitchell RD III, Dhammi A, Bissinger BW, Sonenshine DE, Roe RM. Tick Haller's Organ, a New Paradigm for Arthropod Olfaction: How Ticks Differ from

Insects. *Int J Mol Sci.* 2017 Jul 18;18(7):1563. doi: 10.3390/ijms18071563. PMID: 28718821; PMCID: PMC5536051.

Castro, Martin B., and Stan A. Wright. "Vertebrate hosts of *Ixodes pacificus* (Acari: ixodidae) in California. *Journal of Vector Ecology* 32, no. 1 (2007): 140-149.

Cerutti, Heriberto. "RNA interference: traveling in the cell and gaining functions?". *TRENDS in Genetics* 19, no. 1 (2003): 39-46.

Chang, C. C., Bruno B. Chomel, R. W. Kasten, V. Romano, and N. Tietze. Molecular evidence of Bartonella spp. in questing adult *Ixodes pacificus* ticks in California. *Journal of Clinical Microbiology* 39, no. 4 (2001): 1221-1226.

Chaverra-Rodriguez, Duverney, Elena Dalla Benetta, Chan C. Heu, Jason L. Rasgon, Patrick M. Ferree, and Omar S. Akbari. Germline mutagenesis of *Nasonia vitripennis* through ovarian delivery of CRISPR-Cas9 ribonucleoprotein. *Insect Molecular Biology* 29, no. 6 (2020): 569-577.

Chaverra-Rodriguez, Duverney, Vanessa M. Macias, Grant L. Hughes, Sujit Pujhari, Yasutsugu Suzuki, David R. Peterson, Donghun Kim, Sage McKeand, and Jason L. Rasgon. Targeted delivery of CRISPR-Cas9 ribonucleoprotein into arthropod ovaries for heritable germline gene editing. *Nature communications* 9, no. 1 (2018): 1-11.

Cramaro, Wibke J., Dominique Revets, Oliver E. Hunewald, Regina Sinner, Anna L. Reye, and Claude P. Muller. Integration of *Ixodes ricinus* genome sequencing with

transcriptome and proteome annotation of the naïve midgut. *BMC genomics* 16, no. 1 (2015): 1-15.

De la Fuente J, Estrada-Pena A, Venzal JM, Kocan KM, Sonenshine DE. Overview: ticks as vectors of pathogens that cause disease in humans and animals. *Front Biosci.* 2008;13:6938-46.

de la Fuente, J., Contreras, M., Estrada-Peña, A., and Cabezas-Cruz, A. (2017b). Targeting a global health problem: vaccine design and challenges for the control of tick-borne diseases. *Vaccine* 35, 5089–5094. doi: 10.1016/j.vaccine.2017.07.097

de la Fuente, J., Estrada-Peña, A., Venzal, J. M., Kocan, K. M., and Sonenshine, D. E. (2008). Overview: ticks as vectors of pathogens that cause disease in humans and animals. *Front. Biosci.* 13, 6938–6946. doi: 10.2741/3200

De la Fuente, J., K. M. Kocan, and E. F. Blouin. Tick vaccines and the transmission of tick-borne pathogens. *Veterinary research communications* 31, no. 1 (2007): 85-90.

de la Fuente, José, Consuelo Almazán, Mario Canales, José Manuel Pérez de la Lastra, Katherine M. Kocan, and Peter Willadsen. A ten-year review of commercial vaccine performance for control of tick infestations on cattle. *Animal Health Research Reviews* 8, no. 1 (2007): 23-28.

De La Fuente, Jose, Manuel Rodríguez, Miguel Redondo, Carlos Montero, JoséCarlos García-García, Luis Méndez, Emerio Serrano et al. Field studies and cost-

effectiveness analysis of vaccination with Gavac™ against the cattle tick *Boophilus microplus*. *Vaccine* 16, no. 4 (1998): 366-373.

Dennis, David T., Trudi S. Nekomoto, John C. Victor, William S. Paul, and Joseph Piesman. Reported distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the United States. *Journal of medical entomology* 35, no. 5 (1998): 629-638.

Dykstra, Elizabeth A., Hanna N. Oltean, David Kangiser, Nicola Marsden-Haug, Stephen M. Rich, Guang Xu, Min-Kuang Lee, Muhammad G. Morshed, Christine B. Graham, and Rebecca J. Eisen. Ecology and epidemiology of tickborne pathogens, Washington, USA, 2011–2016. *Emerging infectious diseases* 26, no. 4 (2020): 648.

Edlow Infectious Disease Clinics of North America 2008, 22 (3): xiii-xv

Egkwu, Noble, Daniel E. Sonenshine, Brooke W. Bissinger, and R. Michael Roe. Transcriptome of the female synganglion of the black-legged tick *Ixodes scapularis* (Acari: Ixodidae) with comparison between Illumina and 454 systems. *PLoS One* 9, no. 7 (2014): e102667.

Eisen, Lars, and Kirby C. Stafford III. Barriers to effective tick management and tick-bite prevention in the United States (Acari: Ixodidae). *Journal of medical entomology* 58, no. 4 (2021): 1588-1600.

Eisen, Rebecca J., Lars Eisen, and Charles B. Beard. County-scale distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the continental United States. *Journal of medical entomology* 53, no. 2 (2016): 349-386.

Falco, Richard C., and Durland Fish. Horizontal movement of adult *Ixodes dammini* (Acari: Ixodidae) attracted to CO₂-baited traps. *Journal of Medical Entomology* 28, no. 5 (1991): 726-729.

Francischetti, Ivo MB, Van My Pham, Ben J. Mans, John F. Andersen, Thomas N. Mather, Robert S. Lane, and José MC Ribeiro. The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). *Insect biochemistry and molecular biology* 35, no. 10 (2005): 1142-1161.

Francischetti, Ivo MB, Van My Pham, Ben J. Mans, John F. Andersen, Thomas N. Mather, Robert S. Lane, and José MC Ribeiro. The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). *Insect biochemistry and molecular biology* 35, no. 10 (2005): 1142-1161.

Franta Z, Frantová H, Konvičková J, et al. Dynamics of digestive proteolytic system during blood feeding of the hard tick *Ixodes ricinus*. *Parasit Vectors*. 2010;3:119. Published 2010 Dec 14. doi:10.1186/1756-3305-3-119

Gao, J., Luo, J., Fan, R., Schulte-Spechtel, U.C., Fingerle, V., Guan, G., Zhao, H., Li, Y., Ren, Q., Ma, M. and Liu, Z., 2009. Characterization of a concealed antigen Hq05

from the hard tick *Haemaphysalis qinghaiensis* and its effect as a vaccine against tick infestation in sheep. *Vaccine*, 27(3), pp.483-490.

Graf, J-F., R. Gogolewski, N. Leach-Bing, G. A. Sabatini, M. B. Molento, E. L. Bordin, and G. J. Arantes. Tick control: an industry point of view. *Parasitology* 129, no. S1 (2004): S427-S442.

Gulia-Nuss, Monika, Andrew B. Nuss, Jason M. Meyer, Daniel E. Sonenshine, R. Michael Roe, Robert M. Waterhouse, David B. Sattelle et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nature communications* 7, no. 1 (2016): 1-13.

Hahn, Micah B., Catherine S. Jarnevich, Andrew J. Monaghan, and Rebecca J. Eisen. Modeling the geographic distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the contiguous United States. *Journal of medical entomology* 53, no. 5 (2016): 1176-1191.

Herrmann, John A et al. Temporal and Spatial Distribution of Tick-Borne Disease Cases among Humans and Canines in Illinois (2000-2009). *Environmental health insights* vol. 8, Suppl 2 15-27. 9 Nov. 2014, doi:10.4137/EHI.S16017

Heu, Chan C., Francine M. McCullough, Junbo Luan, and Jason L. Rasgon. CRISPR-Cas9-based genome editing in the silverleaf whitefly (*Bemisia tabaci*). *The CRISPR journal* 3, no. 2 (2020): 89-96.

Holden, Kevin, John T. Boothby, Rickie W. Kasten, and Bruno B. Chomel. Co-detection of *Bartonella henselae*, *Borrelia burgdorferi*, and *Anaplasma phagocytophilum* in *Ixodes pacificus* ticks from California, USA. *Vector-Borne & Zoonotic Diseases* 6, no. 1 (2006): 99-102.

Holden, Kevin, John T. Boothby, Sulekha Anand, and Robert F. Massung. Detection of *Borrelia burgdorferi*, *Ehrlichia chaffeensis*, and *Anaplasma phagocytophilum* in ticks (Acari: Ixodidae) from a coastal region of California. *Journal of Medical Entomology* 40, no. 4 (2003): 534-539.

Horak IG, Camicas JL, Keirans JE. The Argasidae, Ixodidae and Nuttalliellidae (Acari: Ixodida): a world list of valid tick names. *Exp Appl Acarol.* 2002;28(1-4):27-54. doi: 10.1023/a:1025381712339. PMID: 14570115.

Horn, Martin, Martina Nussbaumerová, Miloslav Šanda, Zuzana Kovářová, Jindřich Srba, Zdeněk Franta, Daniel Sojka et al. Hemoglobin digestion in blood-feeding ticks: mapping a multi-peptidase pathway by functional proteomics. *Chemistry & biology* 16, no. 10 (2009): 1053-1063.

Hornbostel, V. L., Richard S. Ostfeld, and Michael A. Benjamin. Effectiveness of *Metarhizium anisopliae* (Deuteromycetes) against *Ixodes scapularis* (Acari: Ixodidae) engorging on *Peromyscus leucopus*. *Journal of Vector Ecology* 30, no. 1 (2005): 91.

Jia, Na, Jinfeng Wang, Wenqiang Shi, Lifeng Du, Yi Sun, Wei Zhan, Jia-Fu Jiang et al. Large-scale comparative analyses of tick genomes elucidate their genetic diversity and vector capacities. *Cell* 182, no. 5 (2020): 1328-1340.

Josek, Tanya, Kimberly KO Walden, Brian F. Allan, Marianne Alleyne, and Hugh M. Robertson. A foreleg transcriptome for *Ixodes scapularis* ticks: candidates for chemoreceptors and binding proteins that might be expressed in the sensory Haller's organ. *Ticks and tick-borne diseases* 9, no. 5 (2018): 1317-1327.

Kaaya, Godwin P., Esther N. Mwangi, and Elizabeth A. Ouna. Prospects for Biological Control of Livestock Ticks, *Rhipicephalus appendiculatus* and *Amblyomma variegatum*, Using the Entomogenous Fungi *Beauveria bassiana* and *Metarhizium Anisopliae*. *Journal of invertebrate pathology* 67, no. 1 (1996): 15-20.

Kain, Douglas E., Felix AH Sperling, and Robert S. Lane. Population genetic structure of *Ixodes pacificus* (Acari: Ixodidae) using allozymes. *Journal of medical entomology* 34, no. 4 (1997): 441-450.

Kain, Douglas Eugene. Population structures of the western black-legged tick, *Ixodes pacificus* (Acari: Ixodidae) and a study of the antiquity of the Lyme disease spirochete in California. University of California, Berkeley, 1996.

Kain, Douglas E., Felix AH Sperling, Howell V. Daly, and Robert S. Lane. Mitochondrial DNA sequence variation in *Ixodes pacificus* (Acari: Ixodidae). *Heredity* 83, no. 4 (1999): 378-386.

Kavi, Harsh H., Harvey R. Fernandez, Weiwu Xie, and James A. Birchler. RNA silencing in *Drosophila*. *Febs Letters* 579, no. 26 (2005): 5940-5949.

Kilpatrick AM, Dobson ADM, Levi T, Salkeld DJ, Swei A, Ginsberg HS, Kjemtrup A, Padgett KA, Jensen PM, Fish D, Ogden NH, Diuk-Wasser MA, 2017. Lyme disease ecology in a changing world: consensus, uncertainty and critical gaps for improving control. *Philos. Trans. R. Soc. Lond., B, Biol. Sci* 372.

Kosoy, Michael, Mike Murray, Robert D. Gilmore Jr, Ying Bai, and Kenneth L. Gage. *Bartonella* strains from ground squirrels are identical to *Bartonella washoensis* isolated from a human patient. *Journal of Clinical Microbiology* 41, no. 2 (2003): 645-650.

Kramer, Vicki L., Mark P. Randolph, Lucia T. Hui, William E. Irwin, Anthony G. Gutierrez, and Duc J. Vugia. Detection of the agents of human ehrlichioses in ixodid ticks from California. *The American journal of tropical medicine and hygiene* 60, no. 1 (1999): 62-65.

Krause PJ, Telford SR 3rd, Spielman A, Sikand V, Ryan R, Christianson D, et al. Concurrent Lyme disease and babesiosis. Evidence for increased severity and duration of illness. *JAMA*. 1996;275(21):1657-60.

Krause, Peter J., Durland Fish, Sukanya Narasimhan, and Alan G. Barbour. *Borrelia miyamotoi* infection in nature and in humans. *Clinical Microbiology and Infection* 21, no. 7 (2015): 631-639.

Lane, R. S., J. Mun, L. Eisen, and R. J. Eisen. Refractoriness of the western fence lizard (*Sceloporus occidentalis*) to the Lyme disease group spirochete *Borrelia bissettii*. *Journal of Parasitology* 92, no. 4 (2006): 691-696.

Lane, Robert S., Richard N. Brown, Joseph Piesman, and Chindi A. Peavey. Vector competence of *Ixodes pacificus* and *Dermacentor occidentalis* (Acari: Ixodidae) for various isolates of Lyme disease spirochetes. *Journal of medical entomology* 31, no. 3 (1994): 417-424.

Larson, Scott R., Xia Lee, and Susan M. Paskewitz. Prevalence of tick-borne pathogens in two species of *Peromyscus* mice common in northern Wisconsin. *Journal of Medical Entomology* 55, no. 4 (2018): 1002-1010.

Levine JF, Wilson ML, Spielman A. Mice as reservoirs of the Lyme disease spirochete. *Am J Trop Med Hyg.* 1985 Mar;34(2):355-60. doi: 10.4269/ajtmh.1985.34.355. PMID: 3985277.

Luttrell MP, Nakagaki K, Howerth EW, Stallknecht DE, Lee KA, 1994. Experimental infection of *Borrelia burgdorferi* in white-tailed deer. *J. Wildl. Dis* 30, 146–154.

Macias, Vanessa M., Sage McKeand, Duverney Chaverra-Rodriguez, Grant L. Hughes, Aniko Fazekas, Sujit Pujhari, Nijole Jasinskiene, Anthony A. James, and Jason L. Rasgon. Cas9-mediated gene-editing in the malaria mosquito *Anopheles stephensi* by ReMOT Control. *G3: Genes, Genomes, Genetics* 10, no. 4 (2020): 1353-1360.

Martínez-Arzate, SAUL GABRIEL, Juan Carlos Sánchez-Bermúdez, Samuel Sotelo-Gómez, H. M. Diaz-Albiter, W. Hegazy-Hassan, E. Tenorio-Borroto, A. Barbabosa-Pliego, and J. C. Vázquez-Chagoyán. Genetic diversity of Bm86 sequences in *Rhipicephalus* (*Boophilus*) *microplus* ticks from Mexico: analysis of haplotype distribution patterns. *BMC genetics* 20, no. 1 (2019): 1-12.

McCoy, K. D., T. Boulinier, C. Tirard, and Y. Michalakis. Host specificity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick *Ixodes uriae*. *Journal of evolutionary biology* 14, no. 3 (2001): 395-405.

Mello, Craig C., and Darryl Conte. Revealing the world of RNA interference. *Nature* 431, no. 7006 (2004): 338-342.

Mencke, N. Acaricidal and repellent properties of permethrin, its role in reducing transmission of vector-borne pathogens. *Parassitologia* 48, no. 1-2 (2006): 139-140.

Merino, Octavio, Sandra Antunes, Juan Mosqueda, Juan A. Moreno-Cid, José M. Pérez de la Lastra, Rodrigo Rosario-Cruz, Sergio Rodríguez, Ana Domingos, and José de la Fuente. Vaccination with proteins involved in tick–pathogen interactions reduces vector infestations and pathogen infection. *Vaccine* 31, no. 49 (2013): 5889-5896

Mulenga, Albert, and Kelly Erikson. A snapshot of the *Ixodes scapularis* degradome. *Gene* 482, no. 1-2 (2011): 78-93.

Narasimhan, Sukanya, Felix Santiago, Raymond A. Koski, Brandon Brei, John F. Anderson, Durland Fish, and Erol Fikrig. Examination of the *Borrelia burgdorferi* transcriptome in *Ixodes scapularis* during feeding. *Journal of bacteriology* 184, no. 11 (2002): 3122-3125.

Nava, *et al.* An overview of systematics and evolution of ticks *Front. Biosci.*, 14 (2009), pp. 2857-2877

Nieto, Nathan C., Janet E. Foley, Jamie Bettaso, and Robert S. Lane. Reptile infection with *Anaplasma phagocytophilum*, the causative agent of granulocytic anaplasmosis. *Journal of Parasitology* 95, no. 5 (2009): 1165-1170.

O'Donoghue, A., Eroy-Reveles, A., Knudsen, G. et al. Global identification of peptidase specificity by multiplex substrate profiling. *Nat Methods* 9, 1095–1100 (2012). <https://doi.org/10.1038/nmeth.2182>

Odongo, David, Lucy Kamau, Robert Skilton, Stephen Mwaura, Cordula Nitsch, Anthony Musoke, Evans Taracha, Claudia Daubenberger, and Richard Bishop. Vaccination of cattle with TickGARD induces cross-reactive antibodies binding to conserved linear peptides of Bm86 homologues in *Boophilus decoloratus*. *Vaccine* 25, no. 7 (2007): 1287-1296.

Perner, Jan, Jan Provazník, Jana Schrenková, Veronika Urbanová, José Ribeiro, and Petr Kopáček. RNA-seq analyses of the midgut from blood-and serum-fed *Ixodes ricinus* ticks. *Scientific reports* 6, no. 1 (2016): 1-18.

- Piesman J, 1979. Host-associations and seasonal abundance of immature *Ixodes dammini* in southeastern Massachusetts. *Ann. Entomol. Soc. Am* 72, 829–832.
- Pirali-Kheirabadi, Khodadad, Hamidreza Haddadzadeh, Mehdi Razzaghi-Abyaneh, Saeed Bokaie, Rasoul Zare, Mehran Ghazavi, and Masoomeh Shams-Ghahfarokhi. Biological control of *Rhipicephalus* (Boophilus) *annulatus* by different strains of *Metarhizium anisopliae*, *Beauveria bassiana* and *Lecanicillium psalliotae* fungi. *Parasitology Research* 100, no. 6 (2007): 1297-1302.
- Porter, W.T., Motyka, P.J., Wachara, J. *et al.* Citizen science informs human-tick exposure in the Northeastern United States. *Int J Health Geogr* 18, 9 (2019).
<https://doi.org/10.1186/s12942-019-0173-0>
- Rego, Ryan OM, Jos JA Trentelman, Juan Anguita, Ard M. Nijhof, Hein Sprong, Boris Klempa, Ondrej Hajdusek *et al.* Counterattacking the tick bite: towards a rational design of anti-tick vaccines targeting pathogen transmission. *Parasites & vectors* 12, no. 1 (2019): 1-20.
- Reubel, Gerhard H., Robert B. Kimsey, Jeffrey E. Barlough, and John E. Madigan. Experimental transmission of *Ehrlichia equi* to horses through naturally infected ticks (*Ixodes pacificus*) from northern California. *Journal of Clinical Microbiology* 36, no. 7 (1998): 2131-2134.

Richter Jr, P. J., R. B. Kimsey, John E. Madigan, J. E. Barlough, J. S. Dumler, and D. L. Brooks. *Ixodes pacificus* (Acari: Ixodidae) as a vector of *Ehrlichia equi* (Rickettsiales: Ehrlichieae). *Journal of Medical Entomology* 33, no. 1 (1996): 1-5.

Rodriguez-Valle, Manuel, Paula Moolhuijzen, Roberto A. Barrero, Chian Teng Ong, Greta Busch, Thomas Karbanowicz, Mitchell Booth et al. Transcriptome and toxin family analysis of the paralysis tick, *Ixodes holocyclus*. *International journal for parasitology* 48, no. 1 (2018): 71-82.

Sajid, Andaleeb, Jaqueline Matias, Gunjan Arora, Cheyne Kurokawa, Kathleen DePonte, Xiaotian Tang, Geoffrey Lynn et al. mRNA vaccination induces tick resistance and prevents transmission of the Lyme disease agent. *Science translational medicine* 13, no. 620 (2021): eabj9827.

Sakamoto, Joyce M., Jerome Goddard, and Jason L. Rasgon. Population and demographic structure of *Ixodes scapularis* Say in the eastern United States. *PLoS One* 9, no. 7 (2014): e101389.

Salkeld, Daniel J., Danielle M. Lagana, Julie Wachara, W. Tanner Porter, and Nathan C. Nieto. Examining prevalence and diversity of tick-borne pathogens in questing *Ixodes pacificus* ticks in California. *Applied and Environmental Microbiology* 87, no. 13 (2021): e00319-21.

Schwarz, Alexandra, Alejandro Cabezas-Cruz, Jan Kopecký, and James J. Valdés. Understanding the evolutionary structural variability and target specificity of tick

salivary Kunitz peptides using next generation transcriptome data. *BMC evolutionary biology* 14, no. 1 (2014): 1-16.

Schwarz, Alexandra, Björn M. von Reumont, Jan Erhart, Andrezza C. Chagas, José MC Ribeiro, and Michalis Kotsyfakis. De novo *Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *The FASEB Journal* 27, no. 12 (2013): 4745-4756.

Sharma, Arvind, Michael N. Pham, Jeremiah B. Reyes, Randeep Chana, Won C. Yim, Chan C. Heu, Donghun Kim et al. Cas9-mediated gene editing in the black-legged tick, *Ixodes scapularis*, by embryo injection and ReMOT Control. *iScience* 25, no. 3 (2022): 103781.

Shirai, Yu, and Takaaki Daimon. Mutations in cardinal are responsible for the red-1 and peach eye color mutants of the red flour beetle *Tribolium castaneum*. *Biochemical and Biophysical Research Communications* 529, no. 2 (2020): 372-378.

Sojka, D., Franta, Z., Horn, M., Caffrey, C. R., Mareš, M., & Kopáček, P. (2013). New insights into the machinery of blood digestion by ticks. *Trends in parasitology*, 29(6), 276-28

Sonenshine D. E. 1991. *Biology of ticks*. Oxford University Press, New York

Spielman A, Wilson ML, Levine JE, Piesman J. 1985. Ecology of *Ixodes dammini*-borne human babesiosis and Lyme disease. *Annual Review of Entomology* 30: 439-460

Teglas, Mike B., and Janet Foley. Differences in the transmissibility of two *Anaplasma phagocytophilum* strains by the North American tick vector species, *Ixodes pacificus* and *Ixodes scapularis* (Acari: Ixodidae). *Experimental & applied acarology* 38, no. 1 (2006): 47-58.

Telford SR 3rd, Mather TN, Moore SI, Wilson ML, Spielman A. Incompetence of deer as reservoirs of the Lyme disease spirochete. *Am J Trop Med Hyg.* 1988 Jul;39(1):105-9. doi: 10.4269/ajtmh.1988.39.105. PMID: 3400797.

Tokarz R, Jain K, Bennett A, Briese T, Lipkin WI. Assessment of polymicrobial infections in ticks in New York State. *Vector Borne Zoonotic Dis.* 2010;10(3):217-21. doi: 10.1089/vbz.2009.0036.

Trager, William. Further observations on acquired immunity to the tick *Dermacentor variabilis* Say. *The Journal of Parasitology* 25, no. 2 (1939): 137-139.

Vechtova, Pavlina, Zoltan Fussy, Radim Cegan, Jan Sterba, Jan Erhart, Vladimir Benes, and Libor Grubhoffer. Catalogue of stage-specific transcripts in *Ixodes ricinus* and their potential functions during the tick life-cycle. *Parasites & vectors* 13, no. 1 (2020): 1-19.

Welch, David F., Karen C. Carroll, Erik K. Hofmeister, David H. Persing, Denise A. Robison, Arnold G. Steigerwalt, and Don J. Brenner. Isolation of a new subspecies, *Bartonella vinsonii* subsp. *arupensis*, from a cattle rancher: identity with isolates

found in conjunction with *Borrelia burgdorferi* and *Babesia microti* among naturally infected mice. *Journal of Clinical Microbiology* 37, no. 8 (1999): 2598-2601.

Zhou, Xuguo, Faith M. Oi, and Michael E. Scharf. Social exploitation of hexamerin: RNAi reveals a major caste-regulatory factor in termites. *Proceedings of the National Academy of sciences* 103, no. 12 (2006): 4499-4504.

Chapter 2- A multi-omics approach towards the identification of novel midgut targets for control of the *Ixodes scapularis*

2.1 Abstract

Ixodes scapularis is a major vector that transmits causative agents of Lyme disease and several other pathogens of human significance. The tick midgut is the main tissue involved in blood-digestion and the first organ to have contact with pathogens ingested through the blood meal. However, knowledge of the process of blood digestion by the ticks and protein expression in the digestive tract is limited. A systems biology approach based on RNA sequencing and TMT protein quantification was therefore used to gain insight into tick midgut during blood ingestion (unfed and partially fed) and digestion (1-, 7-, and 14-days post detachment from the host) A total of 2,580 differentially expressed transcripts and 524 proteins were identified as involved with blood intake, digestion, and nutrient absorption. Detoxification genes, proteases, protease inhibitors, metabolism, and immune-related proteins were most abundantly over-expressed in response to blood-feeding. 12 representative genes from these major gene categories were chosen as potential vaccine candidates and, using RNA interference, the effect of these gene knockdowns on tick biology was investigated. Knockdown of these genes had variable negative impacts on tick physiology such as disruption of the gut membrane, inability to fully engorge, inability to produce eggs, and increased mortality. These and additional gene targets in the midgut provide opportunities to explore novel tick control strategies.

2.2 Introduction

Ticks are hematophagous ectoparasites of medical and veterinary importance worldwide (Edlow 2008). They cause injury through feeding and also transmit a range of pathogens that affect humans and animals, and even cause significant financial loss throughout the world (Eisen and Eisen, 2018). Of about 50,000 vector-borne disease cases reported annually in the United States alone, 95% are due to tick-borne pathogens and >70% are attributed to Lyme disease (Adams et al., 2016). *Ixodes scapularis*, commonly known as the black-legged tick, is responsible for transmitting *Borrelia burgdorferi*, the causative agent of Lyme disease, and six other pathogens of human significance in the Eastern and Central United States (De la Fuente 2008). A single tick species can transmit multiple pathogens. For instance, *I. scapularis* alone may carry up to seven pathogens and can be co-infected with multiple pathogens, and may cause multiple diseases with a single bite (Hersh et al., 2014).

Like all hard ticks, *I. scapularis* requires multiple blood meals during its lifecycle; a blood meal is required to molt to each successive developmental stage, and adult females require blood for egg development (Sonenshine 1992). Adult females remain attached and feed on the host for up to 10 days. The feeding occurs in two phases: the slow feeding phase which begins with tick attachment and may last up to 9 days, followed by the rapid feeding phase which lasts 12-24 hrs (Gulia-Nuss et al., 2016). The tick midgut is the first point of contact for pathogens ingested with the blood meal and therefore several studies have explored midgut proteins that interact with pathogens to develop a transmission-blocking vaccine. Therefore, an anti-tick vaccine has been proposed as a viable strategy to reduce tick burden and tick-borne disease transmission.

The anti-tick vaccine strategy has been successful for cattle tick, *Rhipicephalus (Boophilus) microplus* (Bastos et al., 2010). This vaccine is based on a midgut transmembrane protein (BM86) of *B. microplus*. Tick feeding on immunized cattle led to a reduction in the number of engorged ticks and tick fecundity by 90% (Willadsen et al., 1986, Willadsen, 2006). This vaccine is sold as TickGARD and Gavac (Odongo et al., 2007). However, BM86 is not antigenic against *Ixodes spp.* ticks, and the vaccine against *I. scapularis* orthologs of this protein did not produce any effect on attachment, feeding, and fecundity (Koci et al., 2021). Thus, it is important to identify novel tick antigen candidates for developing an anti-tick vaccine for *Ixodes spp.* ticks (de la Fuente and Kocan, 2003, de la Fuente et al., 2018). Increasing developments in omics studies of ticks and tick-borne pathogens, as well as the availability of reference genomes for *Ixodes spp.* (Gulia-Nuss et al., 2016; Cramaro et al., 2017, Jia et al., 2020) have now made it possible to begin to untangle our understanding of the genetic factors and molecular pathways involved in different aspects of tick biology, including tick-host-pathogen interactions (de la Fuente, 2012, de la Fuente 2017a, Contreras et al., 2019).

In the present study, we carried out a timecourse transcriptome and proteome of the *I. scapularis* midgut from unfed, partially fed, and replete ticks (soon after drop off to actively egg-laying) to understand the dynamic process of blood ingestion and digestion and to identify proteins that might be important for tick feeding and reproduction. Out of thousands of differentially expressed genes, the major categories were detoxification enzymes, proteases, protease inhibitors, lipid carriers, and immune-related genes. To further understand the function of these gene families, we chose to knock down 12 genes using RNA interference. The knockdown produced variable effects ranging from early

detachment to no egg deposition, suggesting that these could be pursued as potential anti-tick vaccine candidates.

2.3 Materials and Methods

Tick samples and feeding

Pathogen-free, unfed adult *I. scapularis* ticks were acquired from the tick rearing facility at the Oklahoma State University, Stillwater, Oklahoma, and kept in an incubator at 95% relative humidity (RH) and 20° C in an incubator in our laboratory until blood-fed on New Zealand white rabbits. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Nevada, Reno (IACUC # 21-01-1118).

Midgut RNA extraction

For transcriptomes, adult females were collected at different time points: unfed (UF), removed from the host five days after attachment (partially engorged, PE), and at 1, 7, and 14 days (D) after voluntary host drop off (post blood meal, PBM) (Figure 1). Ticks were surface cleaned with 70% ethanol. Whole midguts were dissected in cold phosphate buffered saline (PBS) from three ticks and pooled (six for unfed samples) at each time point. Intact Midguts were washed in cold PBS to remove blood. Once cleaned of blood, guts were immediately transferred to a cold 1.7 ml tube containing 200 µl of Trizol and stored at -80°C until processed.

Total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA) and a Zymo Directzol kit (Zymo Research, Irvine, California).

Illumina Sequencing

Total RNA samples were submitted to Genewiz Inc. (New Jersey, USA) for Illumina RNA library construction and sequencing. The mRNA enriched and amplified library fragments were purified and checked for quality and final concentration using an Agilent 2100 Bioanalyzer with a High Sensitivity DNA Chip (Agilent Technologies, USA). The final quantified libraries were pooled in equimolar amounts for sequencing on four lanes of an Illumina HiSeq 2500 DNA sequencer, utilizing a 150 bp paired-end sequencing flow cell with a HiSeq Reagent Kit (Illumina, USA).

RNASeq Data Analyses

Illumina Forward and reverse reads were trimmed for quality with fastp and checked with the fastqc tool. Trimmed reads were aligned to the IscaW1 assembly (VectorBase) and the IscaW1.6 Gene set (VectorBase) (Giraldo-Calderón et al., 2015). The genome was indexed using the HISAT2 index tool (Kim et al., 2019) and filtered reads were aligned to the genome using hierarchical indexing for spliced alignment of transcripts (HISAT2) next-generation sequencing alignment program. The resulting Sequence Alignment/Map (SAM) file was converted to Binary Alignment/Map (BAM) format using SAMTools (Li et al., 2009). BAM files for all samples were then merged into a single count matrix using the tool featureCounts from the Subread package (Liao et al., 2014). Reads in BAM files were compared to a gene transfer format (GTF) annotation file available at VectorBase. Unique reads which fell within an exon region were counted.

Normalization and differential gene expression analysis was carried out using a count matrix from featureCounts and the DESeq2 program (Love et al., 2014). Reads were normalized across all samples using UF samples as the control compared to PE, 1,

7, and 14 D PBM. P-values and log₂ fold changes (Log₂FC) were calculated using the Wald test. The Benjamini-Hochberg procedure was used to calculate the p-adjusted value/false discovery rate (padj/FDR). Transcripts with an FDR less than 0.05 and Log₂FC greater than 2 were considered as differentially expressed (DE) for each comparison. VectorBase analysis tool was used for gene ontology (Giraldo-Calderón et al., 2015).

For clustering, all factors were analyzed at once using the Likelihood Ratio Test (LRT). Briefly, DESeq2 runs LRT that was used to identify any transcripts that show a change in expression across all different levels starting with UF samples to 14 D PBM. The hclust program was used for hierarchical clustering of transcripts using the Pearson method to merge rows (transcripts) into larger clusters or clades in a bottom-up iterative fashion.

Transcriptome data validation by qRT-PCR

Total RNA was isolated in independent triplicates from tick midguts collected at the same time points as mentioned above for the RNASeq. cDNA was synthesized from DNase treated RNA samples (500ng- 1ug) using iScript cDNA synthesis kit and following the manufacturer's protocols (BioRad, Hercules, CA, USA). Quantitative RT-PCR was conducted following Sharma et al. (2019). Briefly, cDNA was diluted 10x with deionized H₂O before using it as a template in qRT-PCR experiments. One microliter cDNA was used in each 10 µl qRT-PCR reaction. Each sample was run in triplicate wells of a 96-well plate. qRT-PCR was performed on CFX Touch Real-Time PCR Detection System using SYBR green master mix (BioRad, Hercules, CA, USA). All reactions were performed with an initial 5 min at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at

60°C, and 30 s at 72°C, and a melt curve was analyzed at 58–95°C in 5 s increments.

Relative expression was calculated using the $2^{-\Delta\Delta C_t}$ method.

Proteomics

For proteome samples, similar to the transcriptome samples, adult females were collected prior to blood feeding (UF) (n = 3), or after voluntary detachment (1 D, and 7 D PBM) (n = 4). They were dissected, and midguts were washed to remove blood in cold 1X PBS (Santa Cruz Biotechnology, Dallas, TX, USA) containing one Pierce™ Protease Inhibitor tablet (Thermo Fisher, Rockford, IL, USA) (PBS+PI). Midguts were transferred to 1.7 mL microtubes containing 200 µl PBS+PI and homogenized using microtube pestles. An additional 400 µL of PBS+PI was added to samples after homogenization; followed by centrifugation for 10 min at 12,000 x g. The supernatant was then transferred to a new microtube and centrifuged again at 12000 x g. The supernatant was collected and sent to the UNR Proteomics Core (University of Nevada, Reno) for Tandem Mass Tagging using Orbitrap (Thermo Scientific, San Jose, CA, USA).

Isolated protein samples were digested and desalted according to Lundby et al. (2012). Briefly, protein concentrations were determined using EZQ Protein Quantification Kit (Invitrogen #R33200). Proteins were then reduced, alkylated, and subjected to methanol-chloroform precipitation. Extracted proteins were digested with endoproteinase Lys-C (Wako, Richmond, VA), followed by digestion with trypsin (Promega, Madison, WI) and desalting and concentration using C₁₈ Sep-Pak cartridges (Waters, Milford, MA, USA).

Peptides were mass tagged using TMT 11-plex isobaric label kit (Thermo-Fisher Cat # 90110) and TMT11-131C (Cat #A34807) following the manufacturer's protocol.

Pooled TMT-labeled peptides were fractionated by basic pH reversed-phase (BPRP) fractionation on an Ultimate 3000 HPLC (Thermo Scientific) using an integrated fraction collector. Elution was performed using 10 min. gradient of 0-20% solvent B followed by a 50 min. gradient of solvent B from 20-45% (Solvent A 5.0% Acetonitrile, 10 mM ammonium bicarbonate pH 8.0, Solvent B 90.0% Acetonitrile, 10 mM ammonium bicarbonate pH 8.0) on a Zorbax 300Extend-C18 column (Agilent) at a flow rate of 0.4 ml/min. A total of 24 fractions were collected at 37 s intervals in a looping fashion for 60 min. Peptide elution was monitored at a wavelength of 220nm using a Dionex Ultimate 3000 variable wavelength detector (Thermo Scientific). Each fraction was then centrifuged to near dryness and desalted using C₁₈ Sep-Pak cartridges followed again by centrifugation to near dryness and reconstitution with 20 ul of 5% acetonitrile and 0.1% formic acid.

Basic pH reversed-phase fractions were separated using an UltiMate 3000 RSLCnano system (Thermo Scientific, San Jose, CA) on a self-packed UChrom C18 column (100 um x 35 cm). Separation was performed using a 180-min. gradient of solvent B from 2-27% (Solvent A: 0.1% formic acid; Solvent B: acetonitrile, 0.1% formic acid) at 50 °C using a digital PicoView nanospray source (New Objectives, Woburn, MA) that was modified with a custom-built column heater and an ABIRD background suppressor (ESI Source Solutions, Woburn, MA). The column used for separation was a self-packed Pico-Frit (New Objectives, Woburn, MA) column with a 15um tip was packed with 40 cm of 1.9 um ReproSil-Pur 120 C18-AQ (Dr. Maisch GmbH, Germany) at 9000 psi using a nano-LC column packing kit (nanoLCMS, Gold River, CA).

Mass spectral analysis was performed using an Orbitrap Fusion mass spectrometer (Thermo Scientific, San Jose, CA). TMT analysis was performed using an MS3 multi-notch approach (Gulia-Nuss et al., 2016). The MS1 precursor selection range is from 400-1400 m/z at a resolution of 120K and automatic gain control (AGC) target of 2.0×10^5 with a maximum injection time of 100 ms. Quadrupole isolation at 0.7 Th for MS² analysis using CID fragmentation in the linear ion trap with a collision energy of 35%. The AGC was set to 4.0×10^3 with a maximum injection time of 150ms. The instrument was operated in a top speed data-dependent mode with a most intense precursor priority with dynamic exclusion set to an exclusion duration of 60 s with a 10ppm tolerance. MS2 fragment ions were captured in the MS3 precursor population. These MS3 precursors were then isolated within a 2.5 Da window and subjected to high energy collision-induced dissociation (HCD) with a collision energy of 55%. The ions were then detected in the Orbitrap at a resolution of 60,000 with an AGC of 5.0×10^4 and a maximum injection time of 150 ms. The data was then analyzed using Sequest (Thermo Fisher Scientific, San Jose, CA, version v.27, rev. 11.) and Proteome Discoverer (Thermo Scientific, San Jose, CA.) using a custom protein fasta (Ixodes-scapularis-Wikel-Peptides_201811) containing 20,473 sequences, assuming the digestion enzyme Trypsin and max number of missed cleavages set to 2. Reporter Ion quantification settings were normalized to total peptide amount, scale on all average, co-isolation threshold 50, and quant value correction factors applied. Peptides tables were exported to Excel for further analysis.

Differentially expressed proteins (DEP) in 1 D and 7 D PBM were identified that met the criteria: an FDR less than 0.05 and $\text{Log}_2\text{FC} > 2$ for each comparison. VectorBase tool was used for GO and pathway enrichment analysis.

Co-analysis of transcriptome and proteome data:

For comparison of transcripts and proteins identified from transcriptomics and proteomics data, DEG and DEP were grouped to identify common and unique products at each time point. 1 D PBM DEP and DEG were compared, and 7 D PBM DEP and DEG were compared. Pearson correlation was conducted between transcriptome and proteome using excel.

RNA interference:

Double-stranded (ds) RNA were synthesized for nine transcripts: Cystatin (ISCW000447), Gastric Triacylglycerol Lipase (ISCW012188), Heme Lipoprotein Precursor (ISCW21710), three Paramyosin paralogs (ISCW016446, ISCW003712, ISCW006917), Troponin I (ISCW023443), Chitinase (ISCW003950), and Peritrophic Membrane Binding Chitin Protein (ISCW000955) (Table 6). Total RNA was extracted from unfed or 1 D PBM tick midguts and cDNA was synthesized as described above. Primers were designed with a T7 promoter sequence (5'-TAATACGACTCACTATAGGG-3') on the 5' end of both forward and reverse primers. RT-PCR conditions were the same as mentioned above in the RT-PCR section. Gel bands were extracted using the QIAquick gel extraction kit (Qiagen, Hilden, Germany), and 100 ng of cleaned PCR product was used as a template for dsRNA synthesis using the T7 MegaScript kit (Invitrogen, Carlsbad, CA, USA). Newly synthesized dsRNA was purified using phenol-chloroform and ethanol precipitated.

Unfed adult female ticks (N=16 per gene per replicate) were injected with 1 μ L of dsRNA (2 μ g/ μ L). Injections were performed with a u-200 insulin syringe on the ventral right side between the 3rd and 4th leg of the tick as described in detail in Reyes et al (2020). Control ticks (N=15) were injected with 1 μ L of RNase/DNase/Protease-free water. Injected ticks were immediately placed in a holding container at 95% RH and observed for 2 h for recovery before going back into incubator at 20⁰ C. Ticks were allowed to recover for 7 D before feeding on a New Zealand white rabbit. These experiments were replicated twice.

Detached ticks were collected daily, weighed, photographed immediately after dropping off, and stored in individual containers in the incubator at 20°C and 95% RH. A batch of females from each treatment was kept for fecundity assessment. Females were observed daily for mortality and egg-laying. Once a female stopped laying eggs, total egg mass was weighed. To measure knockdown efficiency, three UF ticks were collected, and the rest were placed on a rabbit and collected post repletion at 1 D PBM (N=3). These UF and PBM samples were used for qRT-PCR. Relative expression was calculated using the $2^{-\Delta\Delta C_t}$ method.

2.4 RESULTS

Midgut transcriptome and read mapping:

The average number of reads from transcriptomes ranged from 42-54 million (Table 1). A total of 40-53 million reads were recovered after filtering and adapter removal. These clean reads were aligned to the *I. scapularis* Wikel genome sequence (IscaW) and annotation set IscaW1.6. The alignment rate ranged from 47-58% (Table 1).

Midgut transcriptome composition

Differentially expressed genes (DEG) in all samples compared to unfed ranged from 600 to 1,389. Total DEG were 600, 892, 1,389, and 711 in PE, 1D, 7D, and 14 D PBM samples, respectively. Out of these, 368, 487, 722, and 383 transcripts were upregulated whereas, 232, 405, 667, and 328 transcripts were downregulated in PE, 1D, 7D, and 14D samples, respectively (Figure 2).

To assess the enrichment of molecular function of our transcripts, we examined Gene Ontology (GO) term annotations associated with the DE transcripts in PE samples and after blood meal ingestion (1, 7, and 14 D PBM). Transcripts common among all timepoints were also assessed.

A total of 142 DEG transcripts were common among all blood-fed time points, 116 transcripts were upregulated and 26 were downregulated (Figure 2). All transcripts except one had the same expression pattern throughout the timecourse. ISCW010590, an uncharacterized protein with unknown protein domains was downregulated at PE and upregulated at all replete timepoints (1 – 14 D PBM). Among characterized and annotated transcripts, the most abundant upregulated groups were cytochrome P450, glutathione S-transferases, AMP-dependent CoA ligase, and Major facilitator superfamily and Immunoglobulin I-set was downregulated. Additional important upregulated DEG families were: Kunitz domain trypsin inhibitors, trypsin Inhibitor-like containing cysteine-rich domain, sulfotransferases), , glutathione peroxidases, Von Willebrand factor/Lipid transport proteins (a monobactams), MD-2-related lipid-recognition domain-containing cuticle protein containing an R&R consensus and serine carboxypeptidasesparamyosin, and methyltransferase genes were upregulated. Out of the

116 upregulated transcripts, 50 were uncharacterized or unannotated. Eight downregulated transcripts were uncharacterized. ([Datasheet-1](#))

Differentially expressed transcripts in PE:

Out of a total of 218 DEG transcripts in PE compared to UF, 111 were upregulated and 107 were downregulated (Figure 2). A serine carboxypeptidase gene was most upregulated (ISCW024713, 12.6 Log₂FC) and a fatty acyl-CoA elongase protein was most downregulated (ISCW002735, -25.6 Log₂FC) (Table 2). 92 transcripts (~42%) were uncharacterized, out of which 45 were upregulated and 47 were downregulated. The most abundant upregulated protein families were: Cytochrome P450, major facilitator superfamily, and glycoside hydrolase deacetylase, while protein kinase family was downregulated. Other important upregulated DEG belonged to gene families: cystatin, phospholipase B-like, FAD-dependent oxidoreductase, and chitin synthases while ABC-Transporters and pink-eyed dilution protein were downregulated ([Data sheet-1](#))

Differentially expressed transcripts at 1 D PBM:

Out of a total of 361 differentially expressed transcripts identified at 1 D PBM compared to UF (Figure 2), 118 were upregulated and 243 downregulated. The most upregulated gene was a secreted protein, which contains a MD-2 lipid recognition protein domain (ISCW013195, Log₂FC 10.2) and the most downregulated gene was carnitine acyltransferase (ISCW007276, Log₂FC -23.45), an enzyme that catalyzes the exchange of acyl groups between carnitine and coenzyme A (CoA) for fatty acid metabolism (Table 2). 139 transcripts (~38%) were uncharacterized (41 were upregulated and 98 were downregulated). In the characterized genes, the most abundant upregulated gene families were: serpins, papain, trypsin-like serine protease, Glutathione S-transferase,

Insulin-like growth factor-binding protein, lipid transport protein and trypsin inhibitors and DEG downregulated families were: cytochrome P450, CRAL-TRIO lipid-binding domain, cholinesterases, chitinases, protein kinases, ABC transporter-like transcripts, Nuclear hormone receptors, major facilitator sugar transporter-like, methyltransferases, , and insect cuticle protein ,. ([Datasheet-1](#))

Differentially expressed transcripts at 7 D PBM:

A total of 564 transcripts were uniquely identified at 7 D PBM compared to UF (Figure 2). 243 transcripts were upregulated, and 321 transcripts were downregulated. The most upregulated gene was a serine carboxypeptidase (ISCW024713, Log2FC 13.5) and the most downregulated gene was also an uncharacterized protein (ISCW000528, Log2FC -22.5) part of the Ribonuclease Inhibitor family (Table 2). Of all the unique transcripts at this timepoint, 167 (29%) were uncharacterized, 102 were upregulated and 65 were downregulated. The most abundant DEG upregulated gene families were: Cytochrome P450, insect cuticle proteins, major facilitator superfamily, Thioredoxin domain, protein kinases, serine proteases, ABC transporter-like, Fibronectin type II, granulin, trypsin inhibitor kunitz domain containing transcripts, and tick histamine-binding proteins. The most abundant DEG downregulated families were: RNA recognition motif domain, Ankyrin repeat containing, , LDL receptor proteins and leucine-rich repeat proteins. Members of leucine-rich repeat domains, protein kinases, serine proteases, and major facilitator family were both up and down regulated. , ([Data sheet-1](#))

Differentially expressed transcripts at 14D PBM:

A total of 128 transcripts were uniquely identified at 14 D PBM compared to UF (Figure 2). There were 83 downregulated and 45 upregulated transcripts. The most upregulated transcript was an Uncharacterized protein (ISCW013728, Log2FC 13.1) and the most downregulated was a fatty acyl-CoA elongase (ISCW002735, Log2FC -21.5) (Table 2). 43 transcripts were uncharacterized: 32 upregulated and 11 downregulated. The most abundant gene families identified were Zinc finger C2H2-type, RNA recognition motif, Protein kinase, and Leucine-rich repeat. ([Data sheet-1](#))

Proteases and Protease inhibitors

To further understand blood digestion in *I. scapularis*, transcripts for proteolytic enzymes were examined. A total of 80 proteases (Figure 3A) and 27 protease inhibitors (PI) (Figure 3B) were differentially expressed in our data. Eight proteases and PI were common and upregulated in all timepoints (Figure 3A, B). 35 proteases and PIs (31 upregulated and 4 downregulated) were identified at PE, 61 (38 upregulated and 23 downregulated) at 1 D PBM, 69 (48 upregulated and 21 downregulated) at 7 D PBM and 37 (27 upregulated and 10 downregulated) were identified at 14 D PBM (Table 2).

Eleven proteases and PI (9 upregulated and 2 downregulated) were unique to PE, 20 (9 upregulated and 11 downregulated) to 1 D PBM, 19 (11 upregulated and 8 downregulated) to 7 D PBM, and 3 (1 upregulated and 2 downregulated) were unique to 14 D PBM. The most abundant family of proteolytic enzymes were serine proteases (Figure 3C, D) and the most abundant family of protease inhibitors were chymotrypsin and elastase inhibitors (Figure 3 C). Among serine proteases, the most abundant subfamily of proteases was S1A (17 transcripts) (Figure 3D). The second most abundant

group was the S10 family (8 transcripts) (Figure 3D). The most abundant protease inhibitor family was I8 (14 transcripts) (Figure 3D).

Immunity-related genes in response to blood-feeding and digestion

To understand immune response due to blood acquisition in *I. scapularis*, immune-related transcripts were examined. A total of 48 immunity-related transcripts were differentially expressed in our data. Two immunity-related transcripts (both were phospholipid-hydroperoxide glutathione peroxidase) were common among all time points and are suggested to be putatively involved in the gut-microbe homeostasis pathway (Figure 4, Table 3). 16 immunity-related transcripts were identified in PE, 11 were upregulated, and 5 were downregulated. Of these 16 transcripts, 6 were unique to this timepoint (Figure 4 A, B, Table 3). These six transcripts (all upregulated) included: lysozyme, a caspase, part of the Toll pathway, an antimicrobial peptide (AMP), and a thioester-containing protein (Figure 4 C, D, Table 3).

12 immunity-related transcripts were differentially expressed only at 1D PBM. Three transcripts were AMPs, one putatively related to agglutination, a defensin, a lysozyme, a protease inhibitor, argonaute, a thioester containing protein, and a duox (Figure 4 C, D, Table 3). Eight transcripts were expressed only at 7D PBM. These transcripts comprised protease inhibitors, non-self-recognition, JAK/STAT pathway protein, and a putative agglutination protein (Figure 4 C, D, Table 3). Two immunity-related transcripts were identified only in 14 D PBM samples. One transcript was involved in free-radical defense, and another was a putative defensin (Figure 4 C, D; Table 3).

Temporal expression changes during blood-meal ingestion and digestion.

With the combination of the Likelihood Ratio Test from DESeq2 and Hierarchical clustering of transcripts with the Pearson correlation, 2,180 DE transcripts were identified throughout the time points. These transcripts fell into six distinct clusters based on their expression levels (Figure 5). Cluster 1 (oxidoreductase activity, chitin synthase, iron, and heme-binding) was upregulated at PE, downregulated at 1 D PBM and the expression remained unchanged from 1D PBM until 14 D PBM (Figure 5). Cluster 2 transcripts (protein binding, kinases, and Acyl-CoA hydrolases) were upregulated at UF and downregulated at all other time points (Figure 5). Cluster 3 transcripts (zinc ion binding, transcription coregulators, heat shock proteins, structural proteins, and drug binding) were upregulated at UF and 1 D PBM and downregulated at other time points (Figure 5). Cluster 4 transcripts were upregulated at PE, 7 D PBM and 14 D PBM, and downregulated at UF and 1 D (Figure 5). Transcripts in this group were associated with indoleacetamide hydrolase activity, enzyme inhibitors, monooxygenase activity, iron ion binding, heme binding, transporter activity, and lipid binding. Cluster 5 transcripts were upregulated from 1 D PBM until 14 D PBM (Figure 5). Transcripts in this group were associated with rhodopsin kinase activity, triglyceride lipase activity, lipid transporters, carbohydrate-binding, steroid binding, and calcium ion binding. Cluster 6 transcripts were upregulated in all time points from PE to 14 D; however, expression was higher in later timepoints (1-14D) (Figure 5). Transcripts in this group were associated with aminoacyl transferase activity, gluconolactonase activity, vitamin binding, chitin-binding, steroid binding, and extracellular matrix binding.

Validation of RNA-seq data using qRT-PCR

Nine transcripts, Protein slowmo, general transcription factor, stearoyl-CoA-desaturase, gamma-glutamyl transpeptidase, cystatin, gastric triacylglycerol lipase, heme lipoprotein precursor, paramyosin, and troponin I, were selected for validation of the dataset. Of these transcripts, all except protein slowmo were upregulated when compared to UF. All transcripts agreed with the transcriptional expression of RNA-seq data (Figure 6).

Analysis of differentially expressed proteins

A total of 524 proteins were identified in the midgut proteome at two time points (UF vs 1 D PBM and UF vs 7 D PBM) (Figure 7 A, B). 439 DEPs were identified in 1 D PBM (225 upregulated, 214 downregulated) and 423 DEPs were identified in 7 D PBM (225 upregulated, 198 downregulated). 101 and 85 were unique to 1 D and 7 D PBM, respectively (Figure 8 C).

The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis indicated six metabolic pathways were enriched among upregulated proteins in UF vs 1 D and 7D PBM: Purine metabolism, Aminoacyl-tRNA biosynthesis, Monobactam biosynthesis, Inositol phosphate metabolism, Selenocompound metabolism, and Sulfur metabolism (Figure 8). A total of 41 and 42 metabolic pathways were enriched among downregulated proteins in 1 (Figure 9) and 7 D PBM (Figure 10), respectively. Top 5 enriched pathways were: biosynthesis of antibiotics, propanoate metabolism, fructose and mannose metabolism, tyrosine metabolism, and butanoate metabolism.

At the 1 D PBM timepoint, the top five upregulated significant GO functions were: protein binding, translation factor activity, RNA binding, oxidoreductase activity,

and cofactors (Figure 11). At 7 D PBM timepoint, the top five upregulated significant GO functions were: Organic cyclic compound binding, mRNA binding, phosphantetheine binding, aminoacyl-tRNA ligase activity, and structural constituent of ribosome (Figure 12).

Co-expression analysis of Transcriptome and Proteome

Co-expression of genes and proteins at the same timepoint was analyzed. For 1 D PBM transcriptome and proteome, 45 were common in both datasets with low Pearson correlation ($r = 0.38$ P-value < 0.01) (Figure 13). Of these 45, 17 transcripts and proteins were upregulated in both datasets and 10 were downregulated. Out of 17 upregulated, eight were uncharacterized and seven characterized proteins were: selenophosphate synthase, adenylylsulfate kinase, AMP-dependent CoA ligase, fatty acid synthase; cytochrome P450, laminin beta 1 chain, serine proteinase, secreted protein part of the endoplasmic reticulum resident protein 29 family, and tumor rejection antigen (Gp96, a heat shock protein 70) (Table 4). Of the 10 downregulated in both datasets, five were uncharacterized, and other proteins were: medium-chain acyl-CoA dehydrogenase, malate dehydrogenase, cytochrome P450, cysteine proteinase cathepsin L, and argininosuccinate synthase (Table 3).

40 transcripts were common in both 7 D PBM transcriptome and proteome with moderate Pearson correlation ($r = 0.6$ P-value < 0.0001) (Figure 14). Out of 40, 15 were upregulated and 14 were downregulated (Figure 14). Nine of 15 proteins were uncharacterized, other proteins identified were: Cytochrome P450, NAD(P) transhydrogenase, IMP dehydrogenase, adenylylsulfate kinase, AMP dependent CoA ligase, and tumor rejection antigen (Gp96) (Table 5). Out of 14 downregulated proteins in

both datasets, five were uncharacterized, and other proteins were cysteine proteinase cathepsin L, glycogen phosphorylase, argininosuccinate synthase, Thimet oligopeptidase, Hsp70, DNA replication factor/protein phosphatase inhibitor SET/SPR-2, ADP/ATP translocase, Cytochrome P450, and Acetyl-CoA hydrolase (Table 5).

Functional characterization of proteins for putative vaccine candidates

From transcriptome and proteome data, we chose 12 transcripts for functional analysis using RNAi (Figure 15). Three serine proteases (SP1, SP2, and SP4, discussed in detail in Chapter 3), chitinase (carbohydrate metabolism), cystatin (enzyme regulator), gastric triacylglycerol lipase (lipid metabolism), heme lipoprotein precursor (transport protein), troponin I (Ca⁺⁺ induced muscle contraction complex, also known to play a role as an angiogenesis inhibitor), paramyosin (muscle contraction), and peritrophic membrane chitin binding protein (peritrophic matrix formation). Knockdown levels were confirmed by qRT-PCR (Figure 16). 11 out of 12 genes tested were successfully knockdown and the knockdown efficiency varied from 51% to 88% (Figure 16). We were unable to PCR amplify gene peritrophic membrane chitin binding protein with several combinations of primers and could not confirm its knockdown efficiency. Paramyosin, Troponin I, and SP2 knockdown females did not produce any eggs until 6 weeks post-feeding and eventually died (Table 5). Sequences were tagged with T7 promoter sequence (5' TAATACGACTCACTATAG 3') (Table 6). Chitinase knockdown females produced 62% fewer eggs than control. Cystatin knockdown females produced 40% fewer eggs when compared to control. Gastric triacylglycerol lipase knockdown resulted in a 45% reduction in total eggs produced and heme lipoprotein precursor knockdown resulted in a 28% reduction in the total number of eggs laid. Peritrophic membrane

binding chitin protein, SP1, and SP4 knockdown resulted in 46%, 35%, and 63% reduction in eggs, respectively. Paramyosin 2 and 3 knockdown had the least effect on egg production and resulted in 11 and 21% reduction, respectively (Table 7, Figure 17).

Troponin I knockdown resulted in an early drop-off from the host. On average, troponin-1 knockdown ticks dropped off within 3 days of attachment compared to an average of 8 days of controls. Troponin 1 knockdown females had visible red streaks. We dissected 3 troponin knockdown females at 14 days post blood meal and kept the rest for egg-laying. Midgut in these dissected females was degraded resulting in the blood throughout the body cavity.

2.5 Discussion

This work represents the first timecourse transcriptomic and proteomic data of *I. scapularis* female gut during blood acquisition (partially engorged, PE), soon after repletion (1 D PBM), and during egg development (7D PBM) and egg-laying (14D PBM). Ticks start digesting a blood meal while actively feeding on the host; however, the majority of blood is taken during the last 12-24h of feeding (rapid phase) (Starck et al., 2018), therefore, blood digestion continues in replete females. Prior studies on blood digestion in *Ixodes spp.* have primarily focused on PE females (five days on the host) (Anderson et al., 2008, Kotsyfakis et al., 2015, Cramaro et al., 2015, Perner et al., 2016, Xu et al., 2016, Chmelar et al., 2016, Moreira et al., 2017, Oleaga et al., 2017, Charrier et al., 2018, Araujo et al., 2019, Franta et al., 2010, Schwarz et al., 2014, Cramaro et al., 2015) due to the relevance of this timepoint to pathogen acquisition and transmission (des Vignes et al., 2001). However, feeding to repletion and post-repletion blood digestion for egg production are important considerations for tick control.

Over 2,500 transcripts were differentially regulated throughout the timepoints. The highest number of differentially expressed transcripts were at 7 D PBM, and the least was at PE. This could be due to changes in physiology, at an earlier timepoint, ticks have started blood acquisition (slow phase), and digestion and metabolic pathways for storage are likely not yet initiated, whereas, at 7D PBM, females are actively digesting blood and making yolk proteins to provision eggs. The decrease in transcript levels at 14 D PBM could be due to the reason that ticks have a single gonadotrophic cycle and the female will soon die. We used Gene ontology (GO) analysis to examine a representative set of transcripts and proteins in particular metabolic networks. We examined GO across tick blood digestion timelines to obtain a general overview of physiological processes predominating at each time point. The GO classification considers three categories: molecular function, component (cellular localization), and biological process.

The proteolytic enzymes made up one of the largest groups of upregulated transcripts and serine proteases were among the major groups. Previous studies have suggested that the most abundant tick proteolytic enzymes are cathepsins, asparaginyl endopeptidase, and metallopeptidases, which break down protein-rich bloodmeal within acidic digestive vesicles (Lara et al., 2005, Horn et al., 2009 Sojka et al., 2008, Sojka et al., 2013). In our data, we identified a total of 80 protease transcripts. Out of these 80, 35 were serine proteases, followed by metallo- (22 transcripts), cysteine- (17 transcripts), and threonine- (6 transcripts) proteases. Cysteine proteinases and cathepsin L enzymes that catalyze reactions at acidic pH were downregulated in our transcriptome and proteome data suggesting that these differences are likely due to the different time points used in these analyses and a shift towards digestion at alkaline pH in the gut in replete

females. Ticks may switch from using proteases that are active in acidic environments of digestive vesicles while actively feeding on the host to using both acidic and alkaline pH enzymes and digest a large amount of blood in both digestive vesicles and gut lumen (alkaline pH suitable for serine proteases) (Reyes et al., 2020). Ribiero et al. (1988) and another study on replete ticks (Heekin et al., 2013) also suggested upregulation of serine proteases, supporting the current study. The S10 family domain of serine proteases were one of the most abundant proteases (after S1A) in our data. In *Haemaphysalis longicornis* ticks one of the S10 proteases was upregulated during blood-feeding. This protease was isolated from the midgut and was identified as a 53 kDa protein with a pI of 7.5. Enzymatic functional assays revealed that it was active at a broad range of pH and temperature values and cleaved hemoglobin, a major component of the blood meal (Motubu et al., 2007).

I2 (Kunitz-like serine protease inhibitors) and I8 (TIL domain elastase inhibitors) were the most prevalent protease inhibitors in our dataset. A comparative transcriptome study in ticks also suggested these two families are predominant in *I. scapularis* (Porter et al., 2017). The Kunitz-like serpins have been well studied in *R. microplus* and exhibit activity against bovine trypsin and human neutrophils (Lima et al. 2010). There are currently few functional studies for family I8 members in ticks (Porter et al., 2017), however, in *R. microplus*, inhibitory properties have been verified for some I8 members. These include ixodidin, a TIL domain chymotrypsin/elastase inhibitor with antimicrobial properties (Fogaça et al., 2006), and BmSI-7, another TIL domain protein shown to inhibit the bacterial protease subtilisin A and the fungal protease Pr1 (sasaki et al., 2008). An I8 family member from *I. ricinus*, was upregulated after feeding, shows similarity to

von Willebrand Factor (Rudenko et al., 2005), and thus may play a role in platelet aggregation inhibition. Elastin is an insoluble protein responsible for the elastic properties of mammalian connective tissues. Elastin is an extremely hydrophobic protein that is not a substrate for most other proteinases. In mammalian physiology, elastase is necessary to provide homeostasis for tissue repair and inhibition of these elastases might be important for successful feeding in ticks.

We also identified and characterized cystatin in the gut. A review of tick cystatins by Schwarz et al. (2012) listed only one cystatin in *R. microplus*, *Bmcystatin*, which regulates embryogenesis. However, in other tick species several cystatins, have been identified and their functions encompass regulation of digestion, detoxification, innate immunity, pathogen transmission, and immunosuppression (Schwarz et al. 2012). In this study, knockdown of the most upregulated cystatin resulted in decreased blood meal intake and fewer eggs, suggesting its role in blood digestion and embryogenesis (Table 7).

A total of 48 transcripts encoding immunity-related genes were identified as differentially expressed (Table 3). Glutathione peroxidase genes were upregulated throughout blood feeding. In *Drosophila melanogaster* has been shown this gene family has been shown to be involved in gut-microbe homeostasis and play a very important role in differentiation between commensal and pathogenic microbes (MacPherson and Harris, 2004, Buchon et al., 2013). Glutathione peroxidase genes ensure the preservation of bacteria which is beneficial to arthropods, while also trigger an immune response against pathogens (Lazzaro and Rolff, 2011).

A dual oxidase (DUOX) gene was highly upregulated in 1 D PBM guts. The membrane-bound enzyme DUOX is known to produce bactericidal reactive oxygen species (ROS) which thought to antagonize bacterial growth and play a major role in gut mucosal immunity in insects (Jang et al., 2021). In addition to the direct immune function, DUOX-mediated ROS also contributes to gut homeostasis in insects by stabilizing the peritrophic membrane of the midgut through the formation of covalent dityrosine bonds between matrix proteins (Kim et al., 2014; Kumar et al., 2010). Thus, DUOX can interact in the gut with ingested microbes either by producing bactericidal ROS or by generating a physical barrier. The high expression of DUOX at 1 D PBM suggests that it might be involved with peritrophic matrix formation soon after repletion. .

The most abundant group of proteolytic enzymes identified in the time course belonged to the clan PA class S1A serine peptidases which contain the catalytic triad His, Asp, and Ser (Page et al., 2008, Rawlings et. al., 2010). Serine proteases also serve as key regulators for coagulation, AMP synthesis, and melanization of several pathogens (Gorman and Paskewitz, 2001; Janeway and Medzhitov, 2002; Jiravanichpaisal et al., 2006). Members of the S1A protease family are involved in many physiological processes, including the regulation of invertebrate immune responses (Cao and Jiang, 2018). Many S1A family members have been linked to the antimicrobial immune response (Buchon et al., 2009; Patnogie and Leclerc, 2017). Protease inhibitors, for example, serine protease inhibitors (serpins) are imperative for arthropod immunity (Kanost, 1999, Gulley et al., 2013, Reichart et al., 2011). In *I. scapularis*, many were shown previously to be expressed in the gut and salivary glands of UF and PE ticks and it is speculated that these serpins may be used for tick feeding and disease transmission

(Mulenga et al., 2009). Our data suggest many serine proteases and protease inhibitors are upregulated throughout time points (PE-14 D PBM) suggesting that while some of these might be involved in blood digestion, many might be transcribed in response to blood-feeding and immune response due to ingested host factors (Smith and Pal, 2014).

Our data showed the upregulation of proteins associated with Monobactam biosynthesis. These N-thiolated beta-lactams were given the name “monobactams” because they have a flexible monocyclic ring and lack a carboxylic acid moiety, with no second ring fused to the beta-lactam ring. Monobactams are effective only against aerobic Gram-negative bacteria (Turos et al., 2000). Interestingly, the tick microbiome is predominantly composed of Gram-negative bacteria of the phylum Proteobacteria (Benson MJ, et al. 2004; Narasimhan and Fikrig, 2015). Therefore, the increased protein levels of these monobactams in 1D and 7D PBM ticks might be for keeping the bacteria levels in check.

Hierarchical clustering resulted in a total of six separate clusters (Figure 5). One of the clusters (cluster 1) was strictly upregulated at PE, downregulated at 1 D PBM, and remained unchanged after that (Figure 5). This distinct group of genes expressed are predicted oxidoreductases, chitin synthase, and iron and heme-binding. These groups were also upregulated in our proteomics data suggesting that not only transcripts are expressed but also translated into proteins. Many genes involved in oxidoreductase activity in this group are part of the cytochrome P450 family. Cytochrome P450s are essential for the metabolism of foreign compounds, and regulate hormones and ecophysiology (Darmauw et al., 2020). Ticks ingest a large blood meal and can feed on a variety of hosts which has led to an expansion of the CYP450 repertoire (over 200 CYP

genes in the *I. scapularis* genome (Gulia-Nuss et al., 2016). Therefore, upregulation of CYP450s in response to a blood meal is not surprising. However, the increased expression of tick histamine binding protein at 7 D PBM and secreted salivary gland peptides at 14 D PBM suggest that these genes might be wrongly annotated. Tick salivary glands begin to degrade after repletion and drop off and by 7D salivary glands are almost disintegrated (Lejal et al., 2019).

We chose 12 genes for transcriptional silencing using dsRNA RNAi showed variable phenotypes. The chitin synthases family of genes includes chitin-binding proteins, chitinases, and chitin deacetylases. Chitin is the major component of the peritrophic matrix in arthropod gut tissues (Kramer and Koga, 1986.). Peritrophic matrix formation begins with the initiation of blood intake to keep the midgut epidermis safe from the blood bolus (Zhu et al., 2016). Chitinase plays an important role in cuticle remodeling during a blood meal to accommodate the increased gut size (You et al., 2003). Chitinase knockout using CRISPR-Cas9 was well tolerated and was not embryo lethal (Sharma et al., 2022), however, the resulting larvae were not fed and the effect of chitinase knockout during blood-feeding (peritrophic matrix formation and cuticular expansion) is yet to be understood. Chitinase knockdown by RNAi in adult females resulted in early detachment from the host (4.7 days vs. 6.7 days in control) and one of the smallest egg masses suggesting that this gene is important for blood-feeding. (Table 7). Cystatin, a protein inhibitor of papain-like cysteine proteases, has been shown to influence feeding in ticks, hemoglobin digestion, and immunity (Zhou et al., 2006, Yamaji et al., 2009). In our data, cystatin knockdown resulted in higher mortality of injected ticks (Table 7) and ~50% reduction in egg mass. Out of three paramyosin

paralogues, only 1, paramyosin 1, knockdown had a significant effect on tick survival and fecundity. Paramyosin constitutes the thick myosin filaments in striated, obliquely striated, and smooth muscles (Szent-gyorgi et al., 1971, Kantha et al., 1990). Knockdown of this protein might have an effect on muscle contraction important for insertion of hypostome (mouthparts) in the host skin and feeding. 75% of paramyosin 1 knockdown females died before feeding and none of the fed females deposited any eggs. Troponin 1-like protein is a mediator of Ca^{2+} regulation of muscle contraction in vertebrates and contraction of the indirect flight muscles in *D. melanogaster* (Peckman et al., 1990, Agianian et al., 2004). Additionally, troponin has been hypothesized as an angiogenesis inhibitor (Fukumoto et al., 2006). Ticks are pool feeders i.e., they break the blood vessel of the host and feed on the blood pooled at the site (Fukumoto et al., 2006). The host immune system repairs the broken blood vessels (angiogenesis); however, secreted molecules in tick saliva such as troponin 1 inhibit this angiogenesis allowing ticks to feed (Fukumoto et al., 2006). Knockdown of troponin 1 in female ticks before feeding resulted in an early drop-off (3 days as opposed to 6.7 days in control) and no eggs. The ticks that acquired some blood during attachment had red streaks visible through the cuticle and females that remained alive after 14 D post drop off had disintegrated guts when dissected. Although more work is needed to confirm that troponin 1 acts as an angiogenesis inhibitor, our data suggest this possibility and additional role in maintaining gut structure. Out of three serine proteases selected for RNAi, knockdown of one, SP2, resulted in no eggs and SP4 produced the smallest egg clutch among all. These data are discussed in detail in Chapter 3.

Several other genes of interest include the myeloid differentiation factor 2 (MD-2)-related lipid-recognition protein and pink-eyed dilution protein. The MD-2 genes are thought to be involved in immune responses through recognizing bacteria lipopolysaccharide in mammals, arthropods and plants (Inohara and Nunez, 2002). However, the physiological roles of MD-2 in other biological processes are largely unknown. In the brown planthopper, *Nilaparvata lugens*, knockdown of an MD-2 gene led to molting failure and mortality at the nymph–adult transition phase, and impaired egg laying and hatching. (Wang et al., 2021). MD-2 was upregulated in all time points in our data from PE-14D. RNAi of this gene might help us understand its function in tick physiology. Another gene, pink-eyed dilution protein (P-protein), plays a critical role in melanin synthesis in mammals. Mutation in this protein may cause complete or partial albinism (Hirobe et al., 2011). Melanins are complex pigments; typically consist of co-polymers of black or brown eumelanins and red or yellow pheomelanins. The mouse pink-eyed dilution (p) locus is known to control eumelanin synthesis (Chen et al., 2002). Although this protein has not been yet studied in ticks, our data showed downregulation of this gene during blood feeding meaning that expression was higher in unfed females. The phenotypic markers such as coat color and eye color has shown to be important for transgenic animal development. This gene might provide a target for phenotypic screening of transgenic ticks.

In summary, our data suggest that the tick gut transcriptome is dynamic and changes throughout the feeding and blood digestion phases. The transcriptome shift from actively feeding (PE) when the genes related to homeostasis and housekeeping are upregulated to upregulation of genes involved in carbohydrate, lipid, and other

biosynthesis pathways post repletion and finally to lipid and carbohydrate storage and vitellogenesis at 14 D PBM. Detoxification enzymes, proteases, and protease inhibitors are upregulated at all time points. Serine proteases made up the majority of digestive enzymes; however, many of these might be related to immune response. We also identified proteins in monobactam biosynthesis pathway which act on Gram-negative bacteria, a major constituent of the tick microbiome. Deletion of these genes might provide insight into tick-microbe interactions. We selected 12 candidates based on different functions for further characterization. Three out of 12 gene knockdowns led to a complete shutdown of egg development and deposition. One of the targets (troponin) knockdown resulted in an early drop-off. These and additional candidate genes need to be further evaluated as potential vaccine targets.

2.6 References

Adams, Philip P., Carlos Flores Avile, and Mollie W. Jewett. A dual luciferase reporter system for *B. burgdorferi* measures transcriptional activity during tick-pathogen interactions. *Frontiers in cellular and infection microbiology* 7 (2017): 225.

Agianian, Bogos, Uroš Kržič, Feng Qiu, Wolfgang A. Linke, Kevin Leonard, and Belinda Bullard. A troponin switch that regulates muscle contraction by stretch instead of calcium. *The EMBO Journal* 23, no. 4 (2004): 772-779.

Anderson, Jennifer M., Daniel E. Sonenshine, and Jesus G. Valenzuela. Exploring the mialome of ticks: an annotated catalogue of midgut transcripts from the hard tick, *Dermacentor variabilis* (Acari: Ixodidae). *BMC genomics* 9.1 (2008): 1-37.

Araujo, Ricardo N., Naylene Silva, Antonio Mendes-Sousa, Rafaela Paim, Gabriel CA Costa, Luciana R. Dias, Karla Oliveira et al. RNA-seq analysis of the salivary glands and midgut of the Argasid tick *Ornithodoros rostratus*. *Scientific reports* 9, no. 1 (2019): 1-13.

Bastos, Reginaldo G., Massaro W. Ueti, Donald P. Knowles, and Glen A. Scoles. The *Rhipicephalus (Boophilus) microplus* Bm86 gene plays a critical role in the fitness of ticks fed on cattle during acute *Babesia bovis* infection. *Parasites & vectors* 3, no. 1 (2010): 1-11.

Benson, Micah J., Jeffrey D. Gawronski, Douglas E. Eveleigh, and David R. Benson. Intracellular symbionts and other bacteria associated with deer ticks (*Ixodes scapularis*) from Nantucket and Wellfleet, Cape Cod, Massachusetts. *Applied and environmental microbiology* 70, no. 1 (2004): 616-620.

Buchon, Nicolas, Nichole A. Broderick, and Bruno Lemaitre. Gut homeostasis in a microbial world: insights from *Drosophila melanogaster*. *Nature Reviews Microbiology* 11, no. 9 (2013): 615-626.

Charrier, N. Pierre, Marjorie Couton, Maarten J. Voordouw, Olivier Rais, Axelle Durand-Hermouet, Caroline Herve, Olivier Plantard, and Claude Rispe. Whole body transcriptomes and new insights into the biology of the tick *Ixodes ricinus*. *Parasites & vectors* 11, no. 1 (2018): 1-15.

Chen, Kun, Prashiela Manga, and Seth J. Orlow. Pink-eyed dilution protein controls the processing of tyrosinase. *Molecular biology of the cell* 13, no. 6 (2002): 1953-1964.

Chen, Shifu, Yanqing Zhou, Yaru Chen, and Jia Gu. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, no. 17 (2018): i884-i890.

Chmelař, Jindřich, Jan Kotál, Shahid Karim, Petr Kopacek, Ivo MB Francischetti, Joao HF Pedra, and Michail Kotsyfakis. Sialomes and mialomes: a systems-biology view of tick tissues and tick–host interactions. *Trends in parasitology* 32, no. 3 (2016): 242-254.

Contreras, Marinela, Margarita Villar, and José De la Fuente. A vaccinomics approach for the identification of tick protective antigens for the control of *Ixodes ricinus* and *Dermacentor reticulatus* infestations in companion animals. *Frontiers in physiology* (2019): 977.

Cramaro, Wibke J., Dominique Revets, Oliver E. Hunewald, Regina Sinner, Anna L. Reye, and Claude P. Muller. Integration of *Ixodes ricinus* genome sequencing with transcriptome and proteome annotation of the naïve midgut. *BMC genomics* 16, no. 1 (2015): 1-15.

Cramaro, Wibke J., Oliver E. Hunewald, Lesley Bell-Sakyi, and Claude P. Muller. Genome scaffolding and annotation for the pathogen vector *Ixodes ricinus* by ultra-long single molecule sequencing. *Parasites & vectors* 10, no. 1 (2017): 1-9.

D.E. Sonenshine (1991) *Biology of Ticks*, Vol. 1 Oxford University Press New York.

de la Fuente, Jose, and Katherine M. Kocan. Advances in the identification and characterization of protective antigens for recombinant vaccines against tick infestations. *Expert Review of Vaccines* 2, no. 4 (2003): 583-594.

- de la Fuente, Jose, Katherine M. Kocan, Consuelo Almazan, and Edmour F. Blouin. Targeting the tick-pathogen interface for novel control strategies. *Front Biosci* 13 (2008): 6947-6956.
- de la Fuente, José. Controlling ticks and tick-borne diseases... looking forward. *Ticks and tick-borne diseases* 9, no. 5 (2018): 1354-1357.
- Dermauw, Wannes, Thomas Van Leeuwen, and René Feyereisen. Diversity and evolution of the P450 family in arthropods. *Insect Biochemistry and Molecular Biology* 127 (2020): 103490
- des Vignes, Franka, Joseph Piesman, Richard Heffernan, Terry L. Schulze, Kirby C. Stafford III, and Durland Fish. Effect of tick removal on transmission of *Borrelia burgdorferi* and *Ehrlichia phagocytophila* by *Ixodes scapularis* nymphs. *The Journal of infectious diseases* 183, no. 5 (2001): 773-778.
- Edlow, Jonathan A., and Daniel C. McGillicuddy. Tick paralysis. *Infectious disease clinics of North America* 22, no. 3 (2008): 397-413.
- Eisen, Rebecca J., and Lars Eisen. The blacklegged tick, *Ixodes scapularis*: an increasing public health concern. *Trends in parasitology* 34, no. 4 (2018): 295-309.
- Fogaça AC, Almeida IC, Eberlin MN, Tanaka AS, Bulet P, Daffre S. Ixodidin, a novel antimicrobial peptide from the hemocytes of the cattle tick *Boophilus microplus* with inhibitory activity against serine proteinases. *Peptides*. 2006;27:667-74.

Franta, Zdeněk, Helena Frantová, Jitka Konvičková, Martin Horn, Daniel Sojka, Michael Mareš, and Petr Kopáček. Dynamics of digestive proteolytic system during blood feeding of the hard tick *Ixodes ricinus*. *Parasites & vectors* 3, no. 1 (2010): 1-11.

Fukumoto, Shinya, Takeshi Sakaguchi, Myungio You, Xuenan Xuan, and Kozo Fujisaki. Tick troponin I-like molecule is a potent inhibitor for angiogenesis. *Microvascular research* 71, no. 3 (2006): 218-221.

Giraldo-Calderón, Gloria I., Scott J. Emrich, Robert M. MacCallum, Gareth Maslen, Emmanuel Dialynas, Pantelis Topalis, Nicholas Ho et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic acids research* 43, no. D1 (2015): D707-D713.

Gorman, Maureen J., and Susan M. Paskewitz. Serine proteases as mediators of mosquito immune responses. *Insect biochemistry and molecular biology* 31, no. 3 (2001): 257-262.

Gulia-Nuss, Monika, Andrew B. Nuss, Jason M. Meyer, Daniel E. Sonenshine, R. Michael Roe, Robert M. Waterhouse, David B. Sattelle et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nature communications* 7, no. 1 (2016): 1-13.

Gulley, Melissa M., Xin Zhang, and Kristin Michel. The roles of serpins in mosquito immunology and physiology. *Journal of insect physiology* 59, no. 2 (2013): 138-147.

Heekin, A.M., Guerrero, F.D., Bendele, K.G. et al. Gut transcriptome of replete adult female cattle ticks, *Rhipicephalus (Boophilus) microplus*, feeding upon a *Babesia bovis*-infected bovine host. *Parasitol Res* 112, 3075–3090 (2013).

<https://doi.org/10.1007/s00436-013-3482-4>

Heekin, Andrew M., Felix D. Guerrero, Kylie G. Bendele, Leo Saldivar, Glen A. Scoles, Scot E. Dowd, Cedric Gondro, Vishvanath Nene, Appolinaire Djikeng, and Kelly A. Brayton. Gut transcriptome of replete adult female cattle ticks, *Rhipicephalus (Boophilus) microplus*, feeding upon a *Babesia bovis*-infected bovine host. *Parasitology research* 112, no. 9 (2013): 3075-3090.

Hersh, Michelle H., Shannon L. LaDeau, M. Andrea Previtali, and Richard S. Ostfeld. When is a parasite not a parasite? Effects of larval tick burdens on white-footed mouse survival. *Ecology* 95, no. 5 (2014): 1360-1369.

Hirobe, Tomohisa, Shosuke Ito, and Kazumasa Wakamatsu. "The mouse pink-eyed dilution allele of the P gene greatly inhibits eumelanin but not pheomelanin synthesis." *Pigment Cell & Melanoma Research* 24, no. 1 (2011): 241-246.

Horn, Martin, Martina Nussbaumerová, Miloslav Šanda, Zuzana Kovářová, Jindřich Srba, Zdeněk Franta, Daniel Sojka et al. Hemoglobin digestion in blood-feeding ticks: mapping a multi-peptidase pathway by functional proteomics. *Chemistry & biology* 16, no. 10 (2009): 1053-1063.

Inohara, Naohiro, and Gabriel Nuñez. ML—a conserved domain involved in innate immunity and lipid metabolism. *Trends in biochemical sciences* 27, no. 5 (2002): 219-221.

Janeway Jr, Charles A., and Ruslan Medzhitov. Innate immune recognition. *Annual review of immunology* 20, no. 1 (2002): 197-216.

Jang S, Mergaert P, Ohbayashi T, et al. Dual oxidase enables insect gut symbiosis by mediating respiratory network formation. *Proc Natl Acad Sci U S A*. 2021;118(10):e2020922118. doi:10.1073/pnas.2020922118

Jia, Na, Jinfeng Wang, Wenqiang Shi, Lifeng Du, Yi Sun, Wei Zhan, Jia-Fu Jiang et al. Large-scale comparative analyses of tick genomes elucidate their genetic diversity and vector capacities. *Cell* 182, no. 5 (2020): 1328-1340.

Jiravanichpaisal, Pikul, Bok Luel Lee, and Kenneth Söderhäll. Cell-mediated immunity in arthropods: hematopoiesis, coagulation, melanization and opsonization. *Immunobiology* 211, no. 4 (2006): 213-236.

Kanehisa, Minoru, and Susumu Goto. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28, no. 1 (2000): 27-30.

Kanost, Michael R. Serine proteinase inhibitors in arthropod immunity. *Developmental & Comparative Immunology* 23, no. 4-5 (1999): 291-301.

Kantha, Sachi Sri, Shugo Watabe, and Kanehisa Hashimoto. Comparative biochemistry of paramyosin—a review. *Journal of food biochemistry* 14, no. 1 (1990): 61-88.

Kim, Daehwan, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology* 37, no. 8 (2019): 907-915.

Kim, Sung-Hee, and Won-Jae Lee. Role of DUOX in gut inflammation: lessons from *Drosophila* model of gut-microbiota interactions. *Frontiers in cellular and infection microbiology* 3 (2014): 116.

Koči, Juraj, Sandhya Bista, Payal Chirania, Xiuli Yang, Chrysoula Kitsou, Vipin Singh Rana, Ozlem Buyuktanir Yas, Daniel E. Sonenshine, and Utpal Pal. Antibodies against EGF-like domains in *Ixodes scapularis* BM86 orthologs impact tick feeding and survival of *Borrelia burgdorferi*. *Scientific reports* 11, no. 1 (2021): 1-13.

Kongsuwan, Kritaya, Peter Josh, Ying Zhu, Roger Pearson, Joanne Gough, and Michelle L. Colgrave. Exploring the midgut proteome of partially fed female cattle tick (*Rhipicephalus (Boophilus) microplus*). *Journal of insect physiology* 56, no. 2 (2010): 212-226.

Kotsyfakis, Michalis, Alexandra Schwarz, Jan Erhart, and José Ribeiro. Tissue-and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Scientific reports* 5, no. 1 (2015): 1-10.

Kramer, Karl J., and Daizo Koga. Insect chitin: physical state, synthesis, degradation and metabolic regulation. *Insect biochemistry* 16, no. 6 (1986): 851-877.

Kumar, Sanjeev, Alvaro Molina-Cruz, Lalita Gupta, Janneth Rodrigues, and Carolina Barillas-Mury. A peroxidase/dual oxidase system modulates midgut epithelial immunity in *Anopheles gambiae*. *Science* 327, no. 5973 (2010): 1644-1648.

Lara, F. A., U. Lins, G. H. Bechara, and P. L. Oliveira. Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *Journal of Experimental Biology* 208, no. 16 (2005): 3093-3101.

Lazzaro, Brian P., and Jens Rolff. Danger, microbes, and homeostasis. *Science* 332, no. 6025 (2011): 43-44.

Lejal, Emilie, Sara Moutailler, Ladislav Šimo, Muriel Vayssier-Taussat, and Thomas Pollet. Tick-borne pathogen detection in midgut and salivary glands of adult *Ixodes Ricinus*. *Parasites & vectors* 12, no. 1 (2019): 1-8.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. The sequence alignment/map format and SAMtools. *Bioinformatics* 25, no. 16 (2009): 2078-2079.

Liao, Yang, Gordon K. Smyth, and Wei Shi. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, no. 7 (2014): 923-930.

Lima, Cassia A., Ricardo JS Torquato, Sergio D. Sasaki, Giselle Z. Justo, and Aparecida S. Tanaka. Biochemical characterization of a Kunitz type inhibitor similar to

dendrotoxins produced by *Rhipicephalus (Boophilus) microplus* (Acari: Ixodidae) hemocytes. *Veterinary parasitology* 167, no. 2-4 (2010): 279-287.

Love, Michael I., Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, no. 12 (2014): 1-21.

Lundby, A., Secher, A., Lage, K. *et al.* Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nat Commun* 3, 876 (2012).

<https://doi.org/10.1038/ncomms1871>

Macpherson, Andrew J., and Nicola L. Harris. Interactions between commensal intestinal bacteria and the immune system. *Nature Reviews Immunology* 4, no. 6 (2004): 478-485.

Moreira, Higo Nasser Santanna, Rafael Mazioli Barcelos, Pedro Marcus Pereira Vidigal, Raphael Contelli Klein, Carlos Emmanuel Montandon, Talles Eduardo Ferreira Maciel, Juliana Fernandes Areal Carrizo *et al.* A deep insight into the whole transcriptome of midguts, ovaries and salivary glands of the *Amblyomma sculptum* tick. *Parasitology international* 66, no. 2 (2017): 64-73.

Motobu M, Tsuji N, Miyoshi T, Huang X, Islam MK, Alim MA, Fujisaki K. Molecular characterization of a blood-induced serine carboxypeptidase from the ixodid tick *Haemaphysalis longicornis*. *FEBS J.* 2007 Jul;274(13):3299-312. doi: 10.1111/j.1742-4658.2007.05852.x. Epub 2007 Jun 3. PMID: 17542992

Mulenga, Albert, Rabuesak Khumthong, and Katelyn C. Chalaire. *Ixodes scapularis* tick serine proteinase inhibitor (serpin) gene family; annotation and transcriptional analysis.

BMC genomics 10, no. 1 (2009): 1-18.

Narasimhan, Sukanya, and Erol Fikrig. Tick microbiome: the force within. *Trends in parasitology* 31, no. 7 (2015): 315-323.

Odongo, David, Lucy Kamau, Robert Skilton, Stephen Mwaura, Cordula Nitsch, Anthony Musoke, Evans Taracha, Claudia Daubenberger, and Richard Bishop.

Vaccination of cattle with TickGARD induces cross-reactive antibodies binding to conserved linear peptides of Bm86 homologues in *Boophilus decoloratus*. *Vaccine* 25, no. 7 (2007): 1287-1296.

Oleaga, Ana, Prosper Obolo-Mvoulouga, Raúl Manzano-Román, and Ricardo Pérez-Sánchez. A proteomic insight into the midgut proteome of *Ornithodoros moubata* females reveals novel information on blood digestion in argasid ticks. *Parasites & Vectors* 10, no. 1 (2017): 1-21.

Page, Michael J., and Enrico Di Cera. Serine peptidases: classification, structure and function. *Cellular and Molecular Life Sciences* 65, no. 7 (2008): 1220-1236.

Peckham, M., J. E. Molloy, J. C. Sparrow, and D. C. S. White. Physiological properties of the dorsal longitudinal flight muscle and the tergal depressor of the trochanter muscle of *Drosophila melanogaster*. *Journal of Muscle Research & Cell Motility* 11, no. 3 (1990): 203-215.

Perner, Jan, Jan Provazník, Jana Schrenková, Veronika Urbanová, José Ribeiro, and Petr Kopáček. RNA-seq analyses of the midgut from blood-and serum-fed *Ixodes ricinus* ticks. *Scientific reports* 6, no. 1 (2016): 1-18.

Porter, L., Radulović, Ž.M. & Mulenga, A. A repertoire of protease inhibitor families in *Amblyomma americanum* and other tick species: inter-species comparative analyses. *Parasites Vectors* 10, 152 (2017). <https://doi.org/10.1186/s13071-017-2080-1>

Rawlings, Neil D. Peptidase inhibitors in the MEROPS database. *Biochimie* 92, no. 11 (2010): 1463-1483.

Reichhart, Jean Marc, David Gubb, and Vincent Leclerc. The *Drosophila* serpins: multiple functions in immunity and morphogenesis. *In Methods in enzymology*, vol. 499, pp. 205-225. Academic Press, 2011.

Reyes, Jeremiah, Cuauhtemoc Ayala-Chavez, Arvind Sharma, Michael Pham, Andrew B. Nuss, and Monika Gulia-Nuss. Blood digestion by trypsin-like serine proteases in the replete Lyme disease vector tick, *Ixodes scapularis*. *Insects* 11, no. 3 (2020): 201.

Ribeiro, José. Role of saliva in tick/host interactions. *Experimental & applied acarology* 7, no. 1 (1989): 15-20.

Rudenko N, Golovchenko M, Edwards MJ, Grubhoffer L. Differential expression of *Ixodes ricinus* tick genes induced by blood feeding or *Borrelia burgdorferi* infection. *J Med Entomol.* 2005;42:36–41.

Sasaki SD, de Lima CA, Lovato DV, Juliano MA, Torquato RJS, Tanaka AS. BmSI-7, a novel subtilisin inhibitor from *Boophilus microplus*, with activity toward Pr1 proteases from the fungus *Metarhizium anisopliae*. *Exp Parasitol*. 2008;118:214–20.

Schwarz, Alexandra, Alejandro Cabezas-Cruz, Jan Kopecký, and James J. Valdés. Understanding the evolutionary structural variability and target specificity of tick salivary Kunitz peptides using next generation transcriptome data. *BMC evolutionary biology* 14, no. 1 (2014): 1-16.

Schwarz, Alexandra, James J. Valdés, and Michalis Kotsyfakis. The role of cystatins in tick physiology and blood feeding. *Ticks and tick-borne diseases* 3, no. 3 (2012): 117-127.

Sharma, Arvind, et al. Dynamics of Insulin Signaling in the Black-Legged Tick, *Ixodes Scapularis*. *Frontiers in Endocrinology (Lausanne)*, vol. 10, Frontiers Research Foundation, 2019, pp. 292–292, <https://doi.org/10.3389/fendo.2019.00292>.

Sharma, Arvind, Michael N. Pham, Jeremiah B. Reyes, Randeep Chana, Won C. Yim, Chan C. Heu, Donghun Kim et al. Cas9-mediated gene editing in the black-legged tick, *Ixodes scapularis*, by embryo injection and ReMOT Control. *iScience* 25, no. 3 (2022): 103781.

Smith, Alexis A., and Utpal Pal. Immunity-related genes in *Ixodes scapularis*—perspectives from genome information. *Frontiers in cellular and infection microbiology* 4 (2014): 116.

Starck, J. Matthias, Lisa Mehnert, Anja Biging, Juliana Bjarsch, Sandra Franz-Guess, Daniel Kleeberger, and Marie Hörnig. Morphological responses to feeding in ticks (*Ixodes ricinus*). *Zoological Letters* 4, no. 1 (2018): 1-19.

Szent-Györgyi, Andrew G., Carolyn Cohen, and John Kendrick-Jones. Paramyosin and the filaments of molluscan “catch” muscles: II. Native filaments: Isolation and characterization. *Journal of molecular biology* 56, no. 2 (1971): 239-258.

Wang, Wei, Ya Ma, Rui-Rui Yang, Xu Cheng, Hai-Jian Huang, Chuan-Xi Zhang, and Yan-Yuan Bao. An MD-2-related lipid-recognition protein is required for insect reproduction and integument development. *Open biology* 11, no. 12 (2021): 210170.

Whitaker, Douglas, Mary Christman, and Maintainer Douglas Whitaker. Package ‘clustsig’. *R package* (2014).

Willadsen, P., and D. H. Kemp. Vaccination with ‘concealed antigens’ for tick control. *Parasitology Today* 4, no. 7 (1988): 196-198.

Willadsen, Peter. Tick control: thoughts on a research agenda. *Veterinary parasitology* 138, no. 1-2 (2006): 161-168.

Xu, Xing-Li, Tian-Yin Cheng, Hu Yang, and Zhi-Hui Liao. De novo assembly and analysis of midgut transcriptome of *Haemaphysalis flava* and identification of genes involved in blood digestion, feeding and defending from pathogens. *Infection, Genetics and Evolution* 38 (2016): 62-72.

Yamaji, Kayoko, Naotoshi Tsuji, Takeharu Miyoshi, M. Khyrul Islam, Takeshi Hatta, M. Abdul Alim, M. Anisuzzaman, Shiro Kushibiki, and Kozo Fujisaki. A salivary cystatin, HISC-1, from the ixodid tick *Haemaphysalis longicornis* play roles in the blood-feeding processes. *Parasitology research* 106, no. 1 (2009): 61-68.

You, Myungjo, Xuenan Xuan, Naotoshi Tsuji, Tsugihiko Kamio, DeMar Taylor, Naoyoshi Suzuki, and Kozo Fujisaki. Identification and molecular characterization of a chitinase from the hard tick *Haemaphysalis longicornis*. *Journal of Biological Chemistry* 278, no. 10 (2003): 8556-8563.

Zhou, Jinlin, Mami Ueda, Rika Umemiya, Badgar Battsetseg, Damdinsuren Boldbaatar, Xuenan Xuan, and Kozo Fujisaki. A secreted cystatin from the tick *Haemaphysalis longicornis* and its distinct expression patterns in relation to innate immunity. *Insect biochemistry and molecular biology* 36, no. 7 (2006): 527-535.

Zhu, Jiwei, Sayed M. Khalil, Robert D. Mitchell, Brooke W. Bissinger, Noble Egekwu, Daniel E. Sonenshine, and R. Michael Roe. Mevalonate-farnesal biosynthesis in ticks: comparative synganglion transcriptomics and a new perspective. *PLoS One* 11, no. 3 (2016): e0141084.

2.7 Figures and Tables

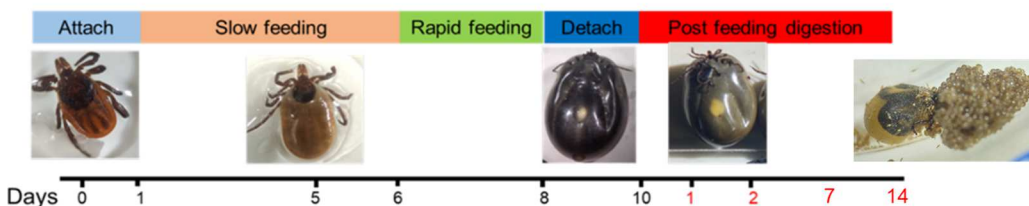


Figure 1. Blood feeding and digestion timecourse. Adult female ticks will attach and slowly ingest 1/3 of total meal over 7-9 days (slow feeding phase). Once mated, females will go in rapid feeding phase (12-24 hrs) where they take 2/3 of their blood meal. Ticks will then fall off host and continue digesting this meal and will begin laying eggs ~14 days post repletion.

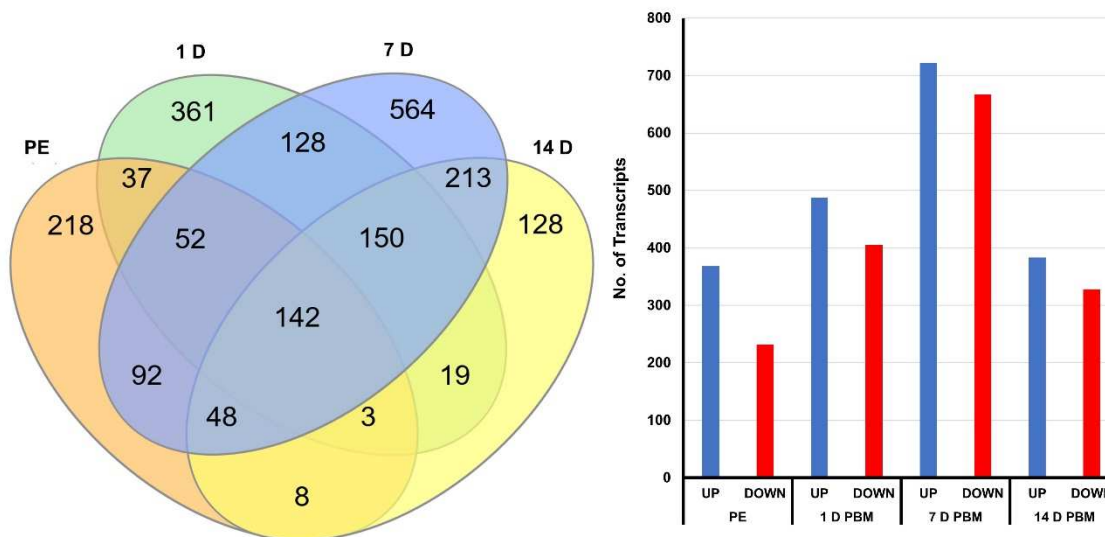


Figure 2. Venn diagram of differentially expressed transcripts identified at all timepoints and profile of differentially expressed transcripts identified. (A)

Transcripts identified from DESeq2 Wald test using a cutoff of -2 and 2 for fold change and adjusted P-value of less than 0.05. All Transcripts identified at partially engorged

(PE), 1 D PBM, 7 D PBM, and 14 D PBM. PBM= post blood meal. **(B)** Number of transcripts upregulated (blue) and downregulated (red) compared to unfed. PE=partially engorged (5 days on host), PBM= post blood meal. 1 D= one day after drop off from host. 7 D and 14D PBM= 7 and 14 days after drop off from the host, respectively.

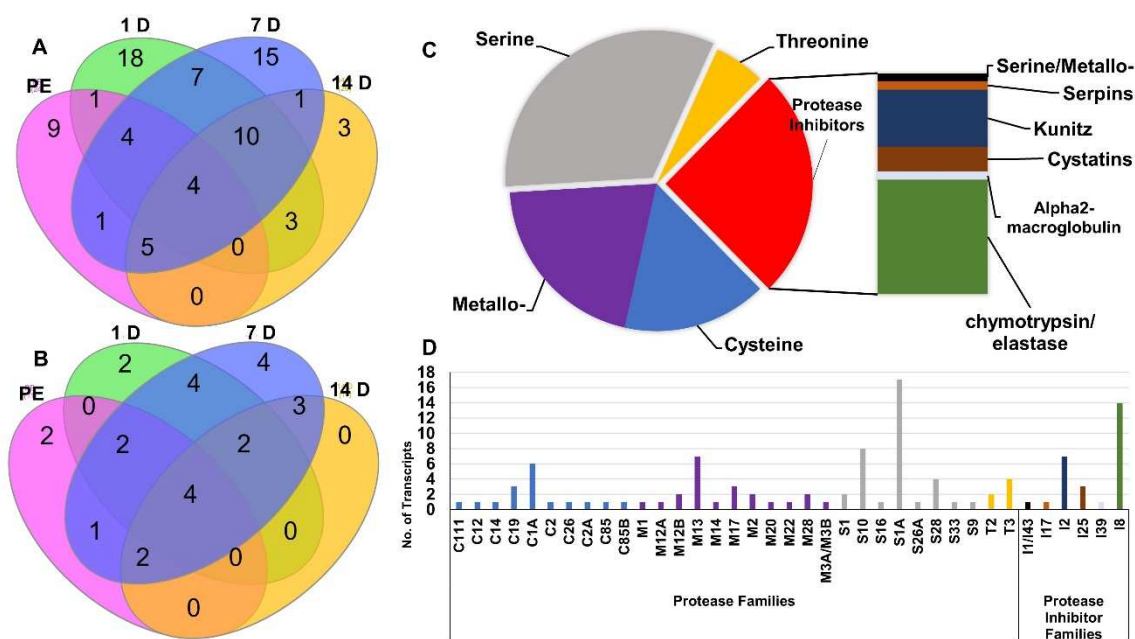


Figure 3. Identification of differentially expressed proteolytic enzyme transcripts as compared to UF. Venn diagram depicting the number of differentially expressed transcripts belonging to proteases (A) and protease Inhibitors (B). C) Classification of proteases by active site (pie chart), along with the classification of the type of protease inhibitors (bar-in-pie). D) MEROPS classification of all identified protease and protease inhibitor families identified.

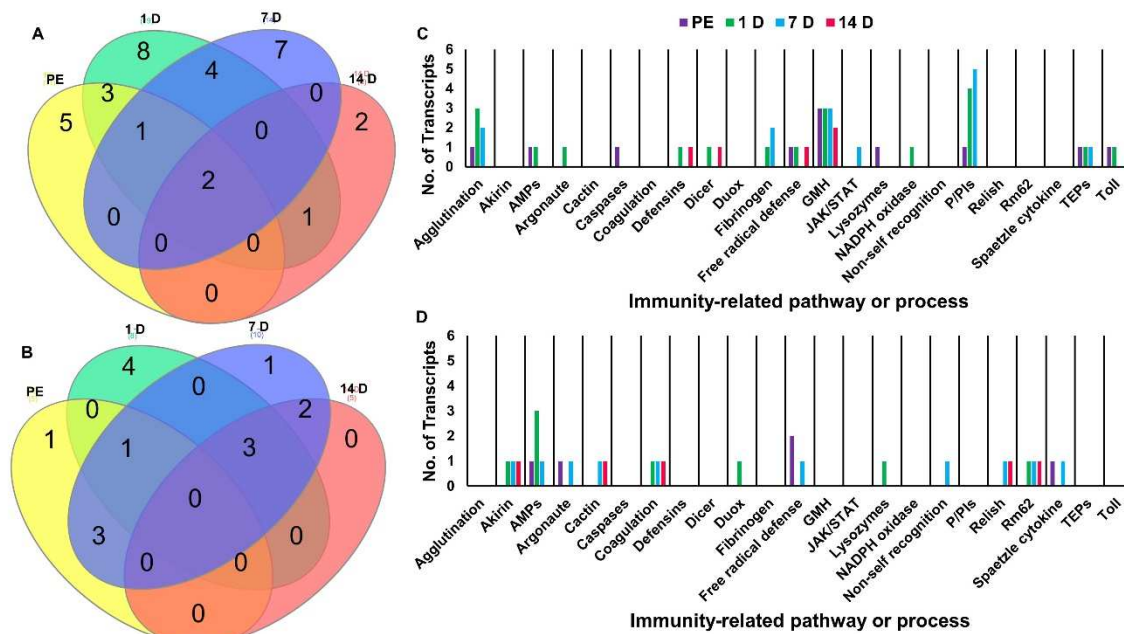


Figure 4. Identification of differentially expressed immunity-related transcripts as compared to UF transcripts. Timepoints are PE (Yellow), 1 D PBM (green), 7 D PBM (blue), and 14 D PBM (magenta). A) Venn diagram depicting the number of differentially expressed upregulated immunity-related transcripts throughout all timepoints. B) Venn diagram depicting the number of differentially expressed downregulated immunity-related transcripts throughout all timepoints. Abundance of all upregulated (C) and downregulated (D) immunity-related transcripts separated by pathway.

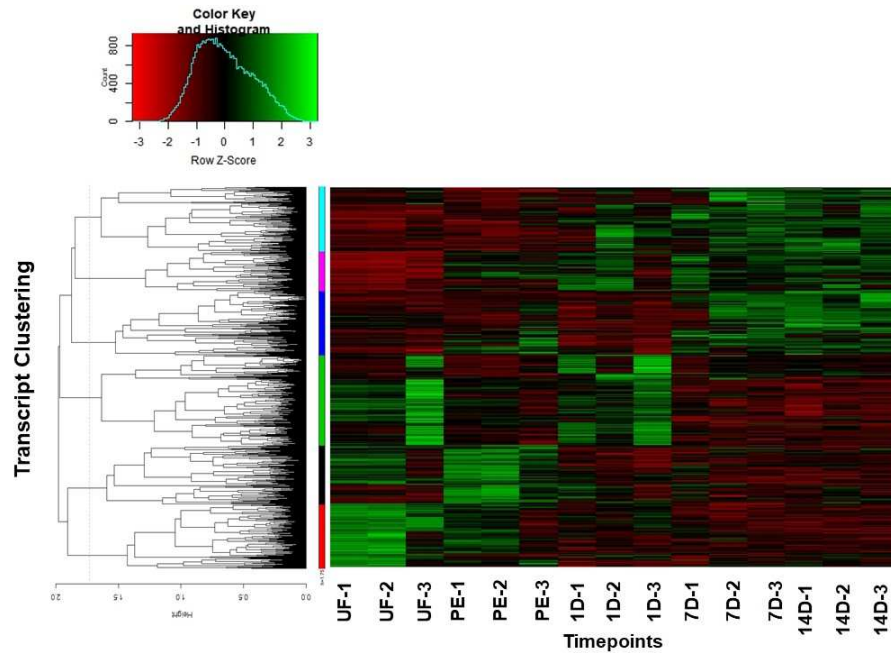


Figure 5. Time series hierarchical clustering from the midguts of UF and fed *I. scapularis* ticks collected at different blood-feeding timepoints. The Z-score represents the expression status of transcripts over time as calculated by the Likelihood Ratio Test. The color bar indicates a group of transcripts that followed a distinct pattern together throughout the blood digestion time course. Cluster bars are distinguished by: Red (cluster 1), Black (cluster 2), Green (cluster 3), Blue (cluster 4), pink (cluster 5), and turquoise (cluster 6).

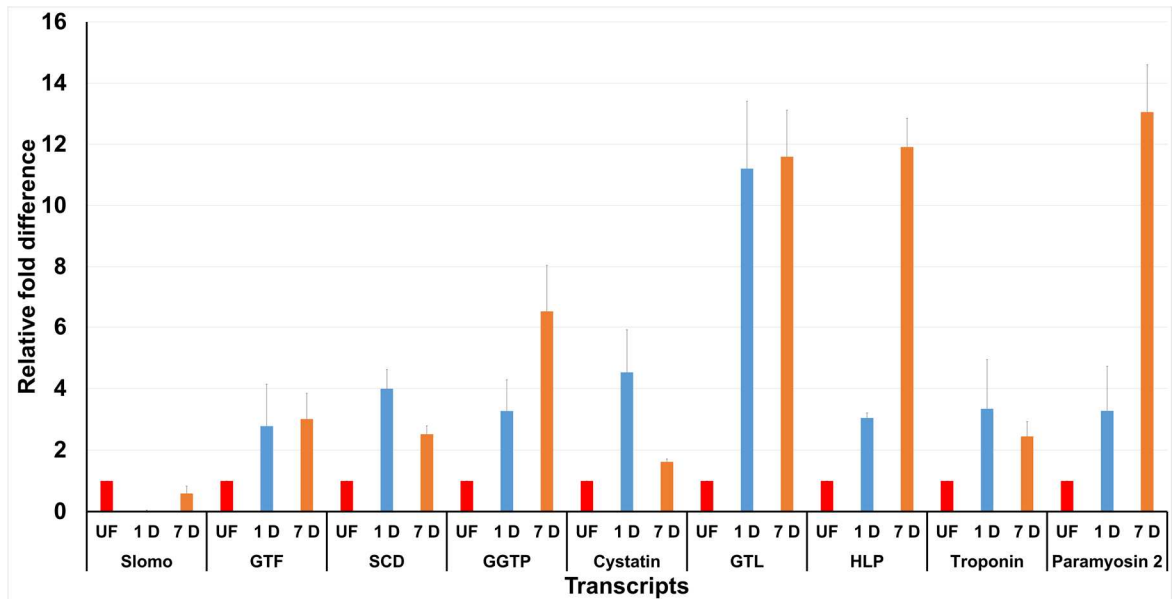


Figure 6. qRT-PCR relative transcript expression selected for validation of RNA-seq dataset. Nine transcripts were selected to validate omics data via qRT-PCR on RNA extracted from unfed (UF - Red), 1 day PBM (1 D - Blue) and 7 day PBM (7 D - Orange) ticks. Transcripts chosen are: Slomo, General transcription factor (GTF), Stearyl-CoA Desaturase, Gamma-glutamyltranspeptidase (GGTP), Cystatin, Gastric Triacylglycerol Lipase (GTL), Heme Lipoprotein precursor (HLP), Troponin I, and Paramyosin 2. and Data were analyzed relative to a housekeeping gene, tubulin, and $2^{-\Delta\Delta C_t}$ method was used. Three biological replicates were used. Data is presented as mean \pm SE.

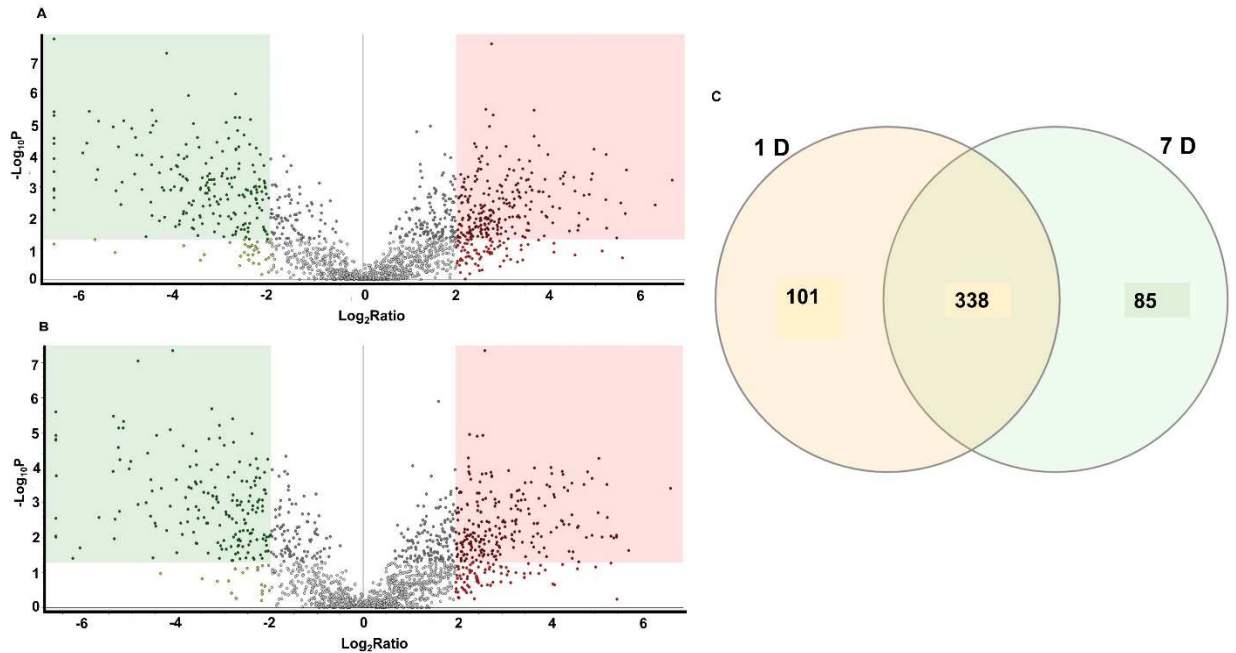


Figure 7. Profile of differentially expressed proteins identified. Volcano plots of log fold change vs P_{adj} differentially expressed proteins (DEPs) depicted by green (upregulated) or red (downregulated) dots in UF vs A) 1 D PBM and B) 7 D PBM. C) Venn diagram of differentially expressed proteins identified at 1 D PBM (yellow) and 7 D PBM (green).

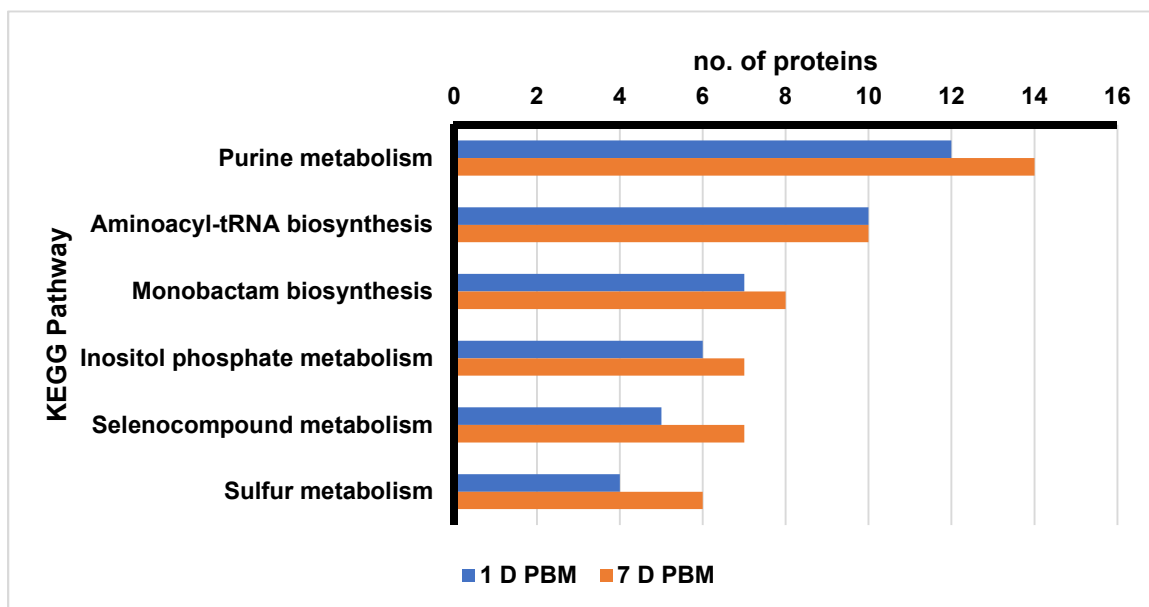


Figure 8. KEGG pathway analysis of upregulated proteins at 1 and 7 D PBM.

Number of differentially expressed proteins belonging to their respective metabolic pathways with 1 D PBM (blue) and 7 D PBM (orange).

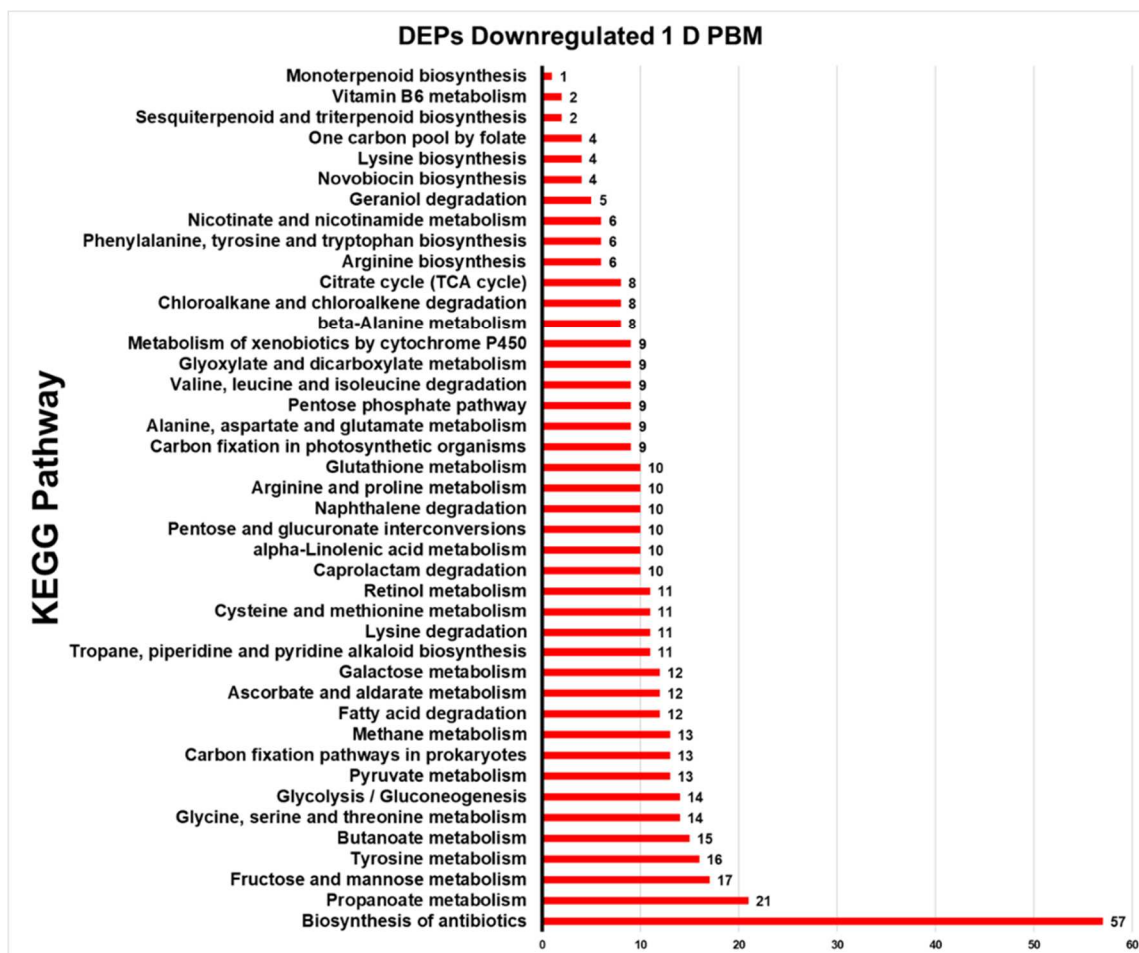


Figure 9. KEGG pathway analysis of downregulated proteins at 1 PBM. Number of Differentially expressed proteins belonging to each specific metabolic pathway for 1 D PBM.

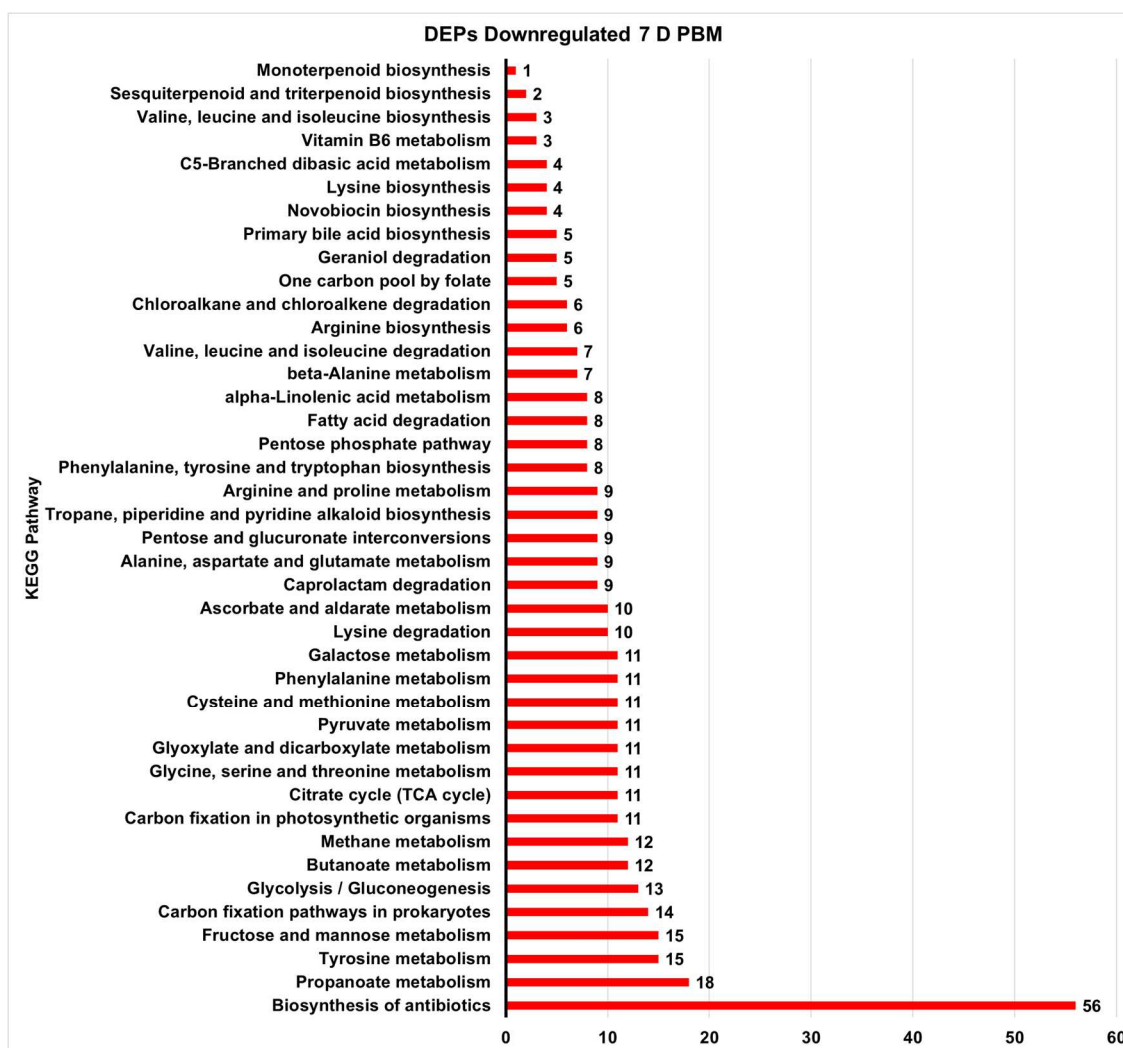


Figure 10. KEGG pathway analysis of downregulated proteins at 7 D PBM. Number of Differentially expressed proteins belonging to each specific metabolic pathway for 7 D PBM.

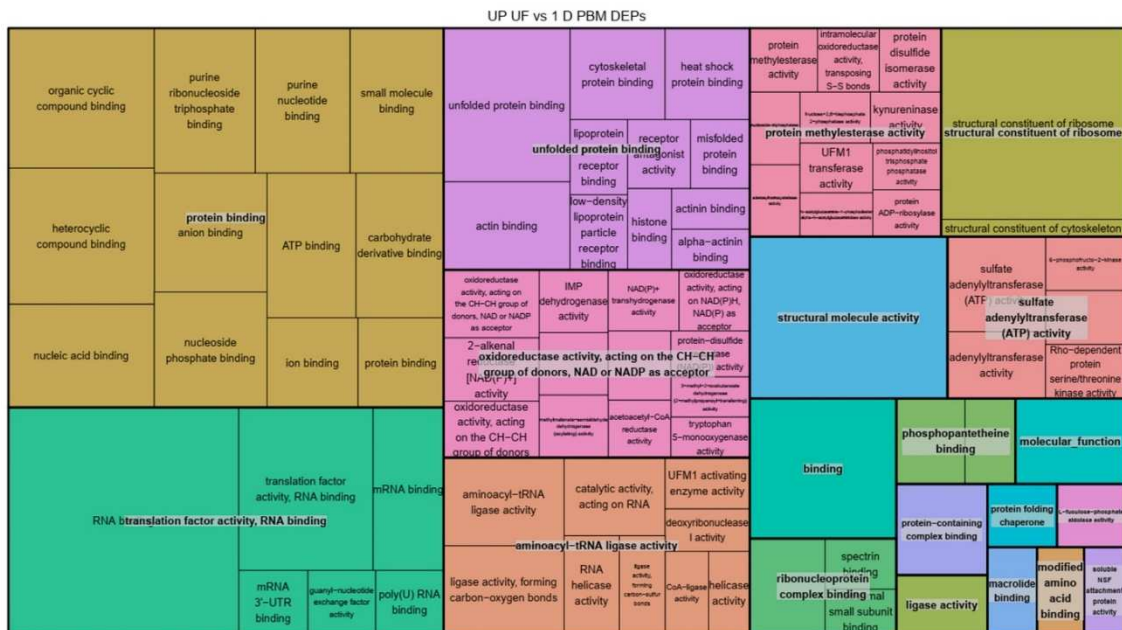


Figure 11. Molecular function gene ontology analysis of upregulated proteins at 1 D PBM. Molecular function ranking from top left to bottom right with proteins most abundant in a specific group starting at the top left (orange - protein binding) to least abundant group in bottom right (purple - soluble (*N*-ethylmaleimide sensitive fusion) NSF attachment protein activity).

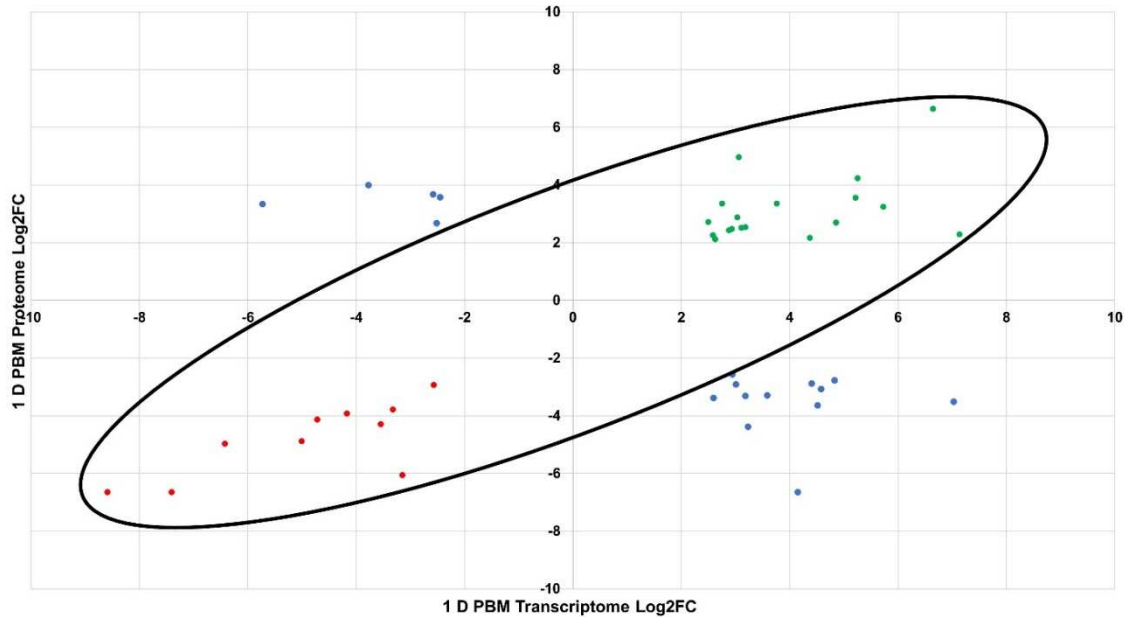


Figure 13. Co-expression analysis of transcriptome and proteome at 1 D PBM when compared to UF samples. Transcripts and proteins which had the same expression pattern are within the black oval, those which didn't agree are outside the oval and depicted as blue dots. Transcripts and proteins which were downregulated in both datasets are red points and those which are both upregulated are green points. Transcripts and proteins are described in Table 3. Pearson correlation value $r = 0.38$ and P -value < 0.01

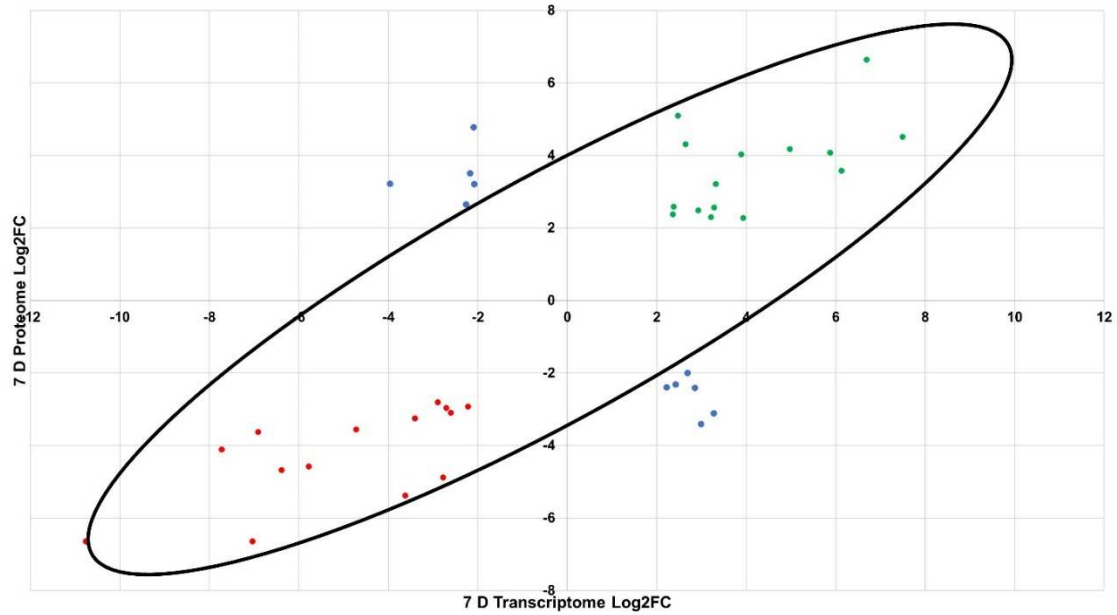


Figure 14. Co-expression analysis of transcriptome and proteome at 1 D PBM when compared to UF samples. Transcripts and proteins which had the same expression pattern are within the black oval, those which didn't agree are outside the oval and depicted as blue dots. Transcripts and proteins which were downregulated in both datasets are red points and those which are both upregulated are green points. Transcripts and proteins are described in Table 4. Pearson correlation of $r = 0.6$ and $P\text{-value} < 0.0001$

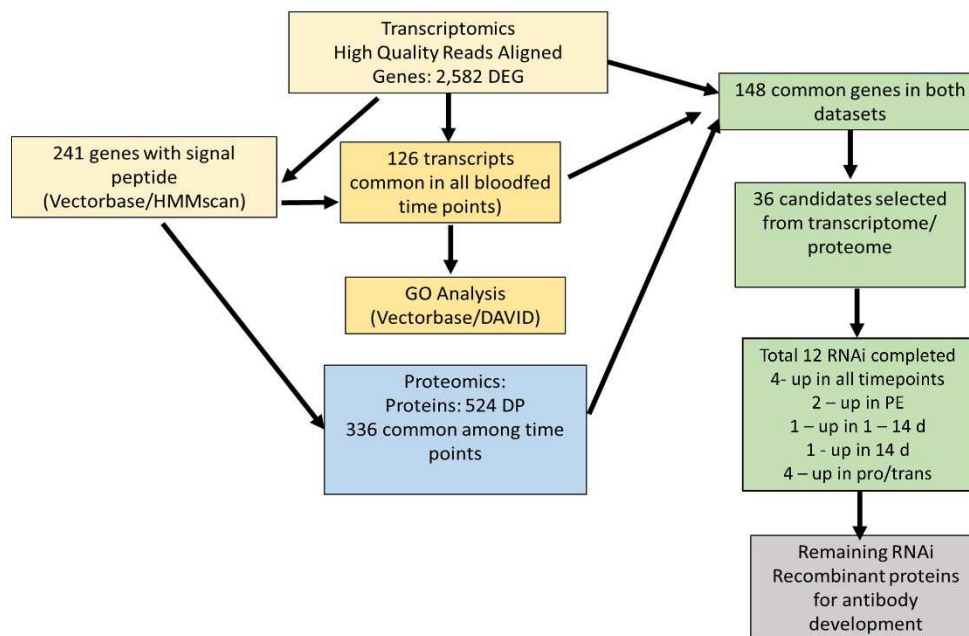


Figure 15. Selection of RNAi targets from mialome transcripts and proteins. Method used for identifying targets for transcriptional silencing.

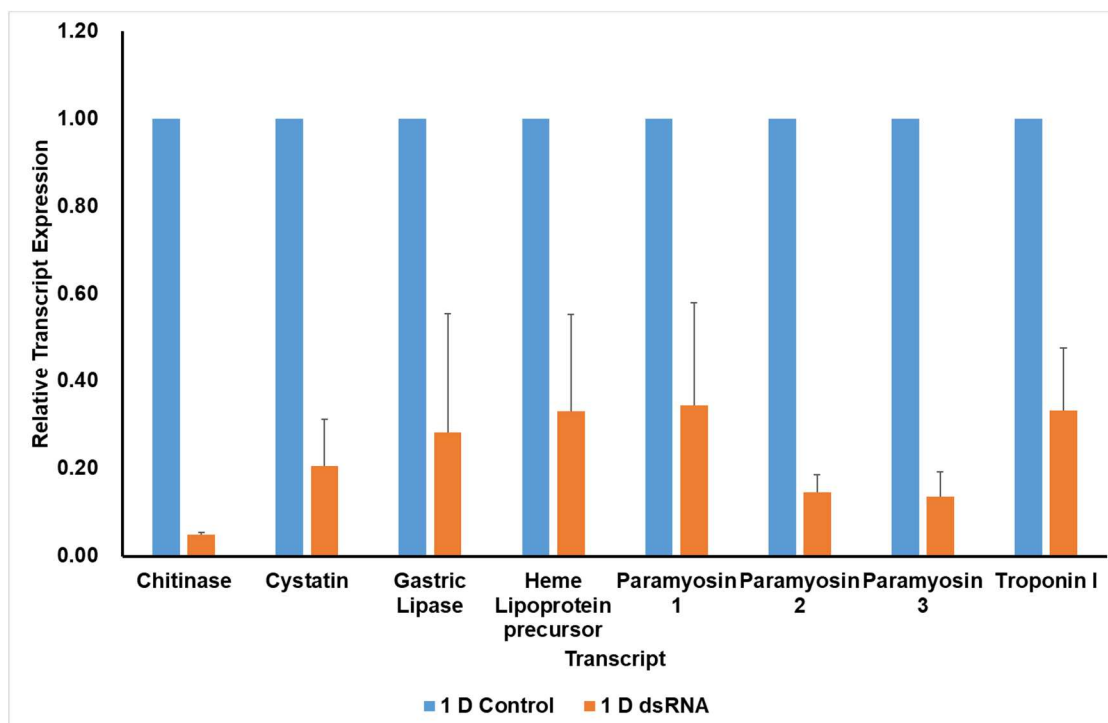


Figure 16. Knockdown efficiency of RNAi targets. Knockdown efficiency was tested by comparing control (water injected – blue) ticks and dsRNA (orange) treated ticks.

Serine protease transcript expression shown in Chapter 3.

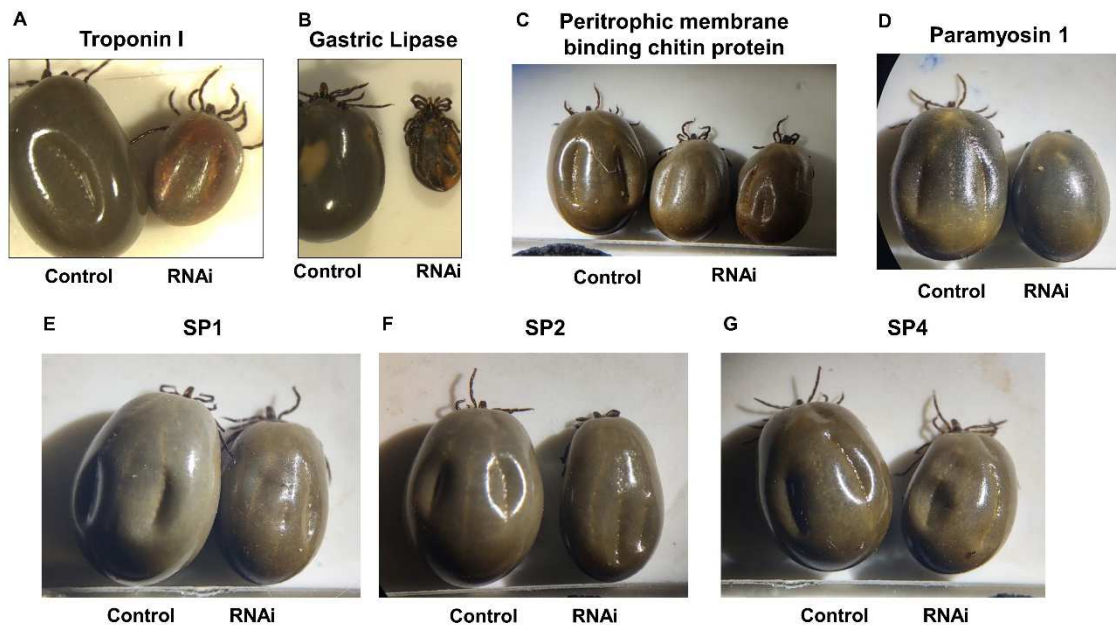


Figure 17. Visual phenotype of dsRNA injected ticks. Control (water-injected) to RNAi (dsRNA treated) ticks. Images from some of the transcripts silenced are A) Troponin I, B) Gastric triacylglycerol lipase, C) Peritrophic membrane binding chitin protein, D) Paramyosin, and several paralogs of serine proteases (E, F, G).

Table 1. Statistics of reads from Illumina sequencing data at all timepoints. Number of total reads from raw sequencing data, clean reads after filtering data using fastp, and average alignment of replicates to Wikel genome.

Timepoint	Average Total Reads (Counts per Million)	Average Clean Reads (Counts per Million)	Alignment Rate
UF	46.9	45.9	52%
PE	54.2	53	55%
1 D PBM	46.8	45.2	58%
7 D PBM	41.8	40.6	52%
14 D PBM	42.8	41.4	47%

Table 2. Top 5 up or downregulated differentially expressed transcripts per timepoint. DEGs from Transcript ID, product description, time point differentially expressed and protein family the transcript belongs to.

Top 5 up or downregulated DEG transcripts per timepoint			
GeneID	Gene Name	Log2FC	Protein Family
Partially Engorged			
ISCW024713	serine carboxypeptidase, putative	12.6	N/A
ISCW002216	cystatin, putative	12.0	Cystatin domain
ISCW000447	Cystatin, putative	12.0	Cystatin domain
ISCW008507	uncharacterized protein	11.7	MAM domain; Concanavalin A-like lectin
ISCW013357	uncharacterized protein	10.7	N/A
ISCW020763	Hsp70, putative (Fragment)	-9.3	Chaperone DnaK; Heat shock protein 70 family
ISCW024176	Secreted protein, putative (Fragment)	-9.4	TRAF-like
ISCW013179	uncharacterized protein	-9.6	3'-5' exonuclease domain
ISCW008943	uncharacterized protein	-22.3	Fibronectin type III
ISCW002735	fatty acyl-CoA elongase, putative	-25.6	ELO family
1 Day PBM			
ISCW013195	Secreted protein, putative	10.2	MD-2-related lipid-recognition domain; Immunoglobulin E-set
ISCW005502	ML domain-containing protein, putative	9.6	MD-2-related lipid-recognition domain; Immunoglobulin E-set
ISCW024071	Secreted salivary gland peptide, putative (Fragment)	9.6	N/A
ISCW007369	Acid methyltransferase, putative	9.4	Methyltransferase type 11
ISCW018910	uncharacterized protein	9.3	N/A
ISCW012645	uncharacterized protein	-10.5	N/A
ISCW015845	unspecified product	-11.6	Immunoglobulin subtype 2
ISCW015846	Neurotrimin, putative (Fragment)	-12.2	Immunoglobulin subtype 2
ISCW026011	unspecified product	-15.7	N/A
ISCW007276	carnitine acyltransferase, putative	-23.5	Acyltransferase ChoActase/COT/CPT
7 Day PBM			
ISCW024713	serine carboxypeptidase, putative	13.5	N/A
ISCW013728	Uncharacterized protein (Fragment) (Fragment)	12.3	N/A
ISCW021228	Vitellogenin, putative	12.1	von Willebrand factor, type D domain; Lipid transport protein
ISCW013727	Vitellogenin, putative	11.9	von Willebrand factor, type D domain; Lipid transport protein
ISCW013000	uncharacterized protein	11.9	N/A
ISCW022961	esterase, putative	-10.8	Cholinesteras
ISCW002736	uncharacterized protein	-20.7	N/A
ISCW009672	Immunoglobulin-binding protein, putative	-21.7	Ganglioside GM2 activator
ISCW016523	uncharacterized protein	-21.9	N/A
ISCW000528	uncharacterized protein	-22.5	N/A
14 Day PBM			
ISCW013728	Uncharacterized protein (Fragment) (Fragment)	13.1	N/A
ISCW004018	uncharacterized protein	12.9	N/A
ISCW024713	serine carboxypeptidase, putative	12.7	N/A
ISCW001698	Serine protease inhibitor, putative	12.5	Pancreatic trypsin inhibitor Kunitz domain
ISCW021228	Vitellogenin, putative	12.5	von Willebrand factor, type D domain; Lipid transport protein
ISCW009672	Immunoglobulin-binding protein, putative	-20.6	Ganglioside GM2 activator
ISCW021619	uncharacterized protein	-20.9	N/A
ISCW016523	uncharacterized protein	-21.1	N/A
ISCW006379	selenoprotein N, 1, putative	-21.1	N/A
ISCW002735	fatty acyl-CoA elongase, putative	-21.5	ELO family

Table 3. Differentially expressed immunity-related transcripts during feeding timecourse. Transcript ID, product description, time point differentially expressed, immunity-related pathway or process, and the protein family these transcripts belong to.

elements	Product Description	PE	1 D	7 D	14 D	Immune Pathway or process	Protein Family
ISCW022517	phospholipid-hydroperoxide glutathione peroxidase, putative	2.60	3.08	2.92	3.42	Gut-microbe homeostasis	Glutathione peroxidase
ISCW015098	Phospholipid-hydroperoxide glutathione peroxidase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QJX7]	8.44	5.84	6.37	5.85	Gut-microbe homeostasis	Glutathione peroxidase
ISCW009616	serpin 7 precursor, putative	2.69	2.34	2.31		Proteases/Protease inhibitors	Serpin domain
ISCW002695	Secreted salivary gland peptide, putative [Source:UniProtKB/TrEMBL;Acc:B7PC21]	-2.97	-2.68	-2.92		Antimicrobial	Serine proteases, trypsin domain
ISCW002701	ATP-dependent RNA helicase, putative [Source:UniProtKB/TrEMBL;Acc:B7PA01]	-3.78	-3.96	-4.48		Rm62	DEAD/DEAH box helicase domain;Helicase, C-terminal
ISCW023283	Protective antigen 4D8 [Source:UniProtKB/TrEMBL;Acc:Q4VRW2]	-3.09	-2.99	-3.79		Akirin	Akirin
ISCW001322	proclotting enzyme precursor, putative	-6.68	-6.41	-10.15		Coagulation	Serine proteases, trypsin domain
ISCW019584	phospholipid-hydroperoxide glutathione peroxidase, putative	6.65	4.38			Gut-microbe homeostasis	Thioredoxin-like superfamily
ISCW021704	hemelipoglyco-carrier protein	5.50	5.26			Agglutination	Lipovitelin-phosvitin complex, superhelical domain;Lipid transport protein, beta-sheet shell
ISCW012382	superoxide dismutase Cu-Zn, putative	3.39	3.45			Free radical defense	Superoxide dismutase-like, copper/zinc binding domain superfamily
ISCW022696	Eukaryotic translation initiation factor 2C, putative [Source:UniProtKB/TrEMBL;Acc:B7QBT5]	-3.37	-3.33			Argonaute	PAZ domain superfamily;Ribonuclease H-like superfamily
ISCW024422	Superoxide dismutase, putative [Source:UniProtKB/TrEMBL;Acc:B7PVY8]	-3.03	-2.26			Free radical defense	Superoxide dismutase-like, copper/zinc binding domain superfamily
ISCW018887	uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7PQ56]	-5.00	-4.30			Spatzle cytokine	Cystine-knot cytokine
ISCW013746	ixoderin precursor, putative	4.56	4.91			Fibrinogen-related	Fibrinogen, alpha/beta/gamma chain, C-terminal globular domain
ISCW023621	Secreted salivary gland peptide, putative [Source:UniProtKB/TrEMBL;Acc:B7QL35]	4.11	3.52			Proteases/Protease inhibitors	Serpin domain
ISCW021709	hemelipoglyco-carrier protein	6.45	4.78			Agglutination	von Willebrand factor, type D domain;Lipid transport protein, N-terminal;Vitellogen, open beta-sheet
ISCW023622	serpin-4 precursor, putative	5.33	5.02			Proteases/Protease inhibitors	Serpin domain
ISCW000889	Dicer-1, putative [Source:UniProtKB/TrEMBL;Acc:B7P5Y7]	2.23	2.70			Dicer	Helicase, C-terminal;Ribonuclease III domain;Dicer dimerisation domain;Helicase/UvrB, N-terminal
ISCW017768	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PK31]	-2.83	-2.50			Cactin	Cactin, C-terminal;Cactin, central domain

elements	Product Description	PE	1 D	7 D	14 D	Immune Pathway or process	Protein Family
ISC W018935	Nuclear factor nf-kappa-B P105 subunit, putative [Source:UniProtKB/TrEMBL;Acc:B7PL56]			-2.43	-2.94	Relish	Rel homology domain (RHD), DNA-binding domain;Rel homology dimerisation domain
ISC W001645	lysozyme, putative	4.05				Lysozymes	Lysozyme-like domain superfamily
ISC W015027	superoxide-dismutase, putative	-4.23				Free radical defense	Superoxide dismutase-like, copper/zinc binding domain superfamily
ISC W003039	Caspase, apoptotic cysteine protease, putative [Source:UniProtKB/TrEMBL;Acc:B7PD12]	3.69				Caspases	Peptidase C14A, caspase catalytic domain
ISC W007727	Toll, putative (Fragment)	4.10				Toll	Toll/interleukin-1 receptor homology (TIR) domain superfamily
ISC W005926	Scapularisin preproprotein, putative [Source:UniProtKB/TrEMBL;Acc:B7PUU6]	3.55				Antimicrobial	Knottin, scorpion toxin-like superfamily
ISC W019887	Alpha-2-macroglobulin, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PU24]	2.88				Thio-ester containing proteins (TEPs)	Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid;Alpha-macroglobulin, receptor-binding domain superfamily
ISC W024299	hemelipoglyco-carrier protein	7.03				Agglutination	von Willebrand factor type D domain
ISC W016747	Preprodefensin, putative [Source:UniProtKB/TrEMBL;Acc:B7P9L7]	5.74				Defensins	Knottin, scorpion toxin-like superfamily
ISC W001646	Lysozyme, putative [Source:UniProtKB/TrEMBL;Acc:B7P0W5]	-4.17				Lysozymes	Lysozyme-like domain superfamily
ISC W023623	serpin-4 precursor, putative	6.17				Proteases/ Protease inhibitors	Serpin superfamily
ISC W002630	NADPH oxidase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PD Q2]	4.69				NADPH oxidase	EF-hand domain pair;Ferredoxin-NADP reductase (FNR), nucleotide-binding domain
ISC W021130	Translation initiation factor 2C, putative [Source:UniProtKB/TrEMBL;Acc:B7Q990]	4.03				Argonaute	PAZ domain superfamily;Ribonuclease H-like superfamily
ISC W002331	Secreted salivary gland peptide, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PC03]	-4.99				Antimicrobial	Peptidase S1, PA clan
ISC W018425	Secreted protein, putative [Source:UniProtKB/TrEMBL;Acc:B7PPI3]	2.58				Antimicrobial	Endoplasmic reticulum resident protein 29, C-terminal domain superfamily;Thioredoxin-like superfamily
ISC W023777	Alpha-macroglobulin, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QMC8]	3.39				Thio-ester containing proteins (TEPs)	Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid;immunoglobulin E-set
ISC W004495	Toll, putative [Source:UniProtKB/TrEMBL;Acc:B7PED2]	4.43				Toll	Leucine-rich repeat, typical subtype
ISC W007865	Dual oxidase 1, putative [Source:UniProtKB/TrEMBL;Acc:B7PVC0]	-3.25				Duox	EF-hand domain pair;Haem peroxidase superfamily;Ferredoxin-NADP reductase (FNR), nucleotide-binding domain;Riboflavin synthase-like beta-barrel
ISC W017494	Map-kinase activating death domain protein, putative [Source:UniProtKB/TrEMBL;Acc:B7P802]	-2.46				Antimicrobial	cDENN domain;uDENN domain;dDENN domain

elements	Product Description	PE	1 D	7 D	14 D	Immune Pathway or process	Protein Family
ISCW021005	Secreted protein, putative [Source:UniProtKB/TrEMBL;Acc:B7Q956]			-2.50		Non-self recognition	Cysteine-rich flanking region, C-terminal;Leucine-rich repeat, typical subtype
ISCW02080	Protein disulfide isomerase 1, putative [Source:UniProtKB/TrEMBL;Acc:B7PBW3]			2.71		Proteases/Protease inhibitors	Thioredoxin-like superfamily
ISCW03089	uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7PC26]			2.59		Thio-ester containing proteins (TEPs)	Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid;Alpha-macroglobulin, receptor-binding domain superfamily
ISCW09628	Receptor tyrosine phosphatase type r2b, putative [Source:UniProtKB/TrEMBL;Acc:B7Q018]			3.58		JAK/STAT pathway(Domeless)	Protein-tyrosine phosphatase-like
ISCW024686	secreted salivary gland peptide, putative Thioredoxin peroxidase, putative [Source:UniProtKB/TrEMBL;Acc:B7QJH0]			6.38		Fibrinogen-related protein	Fibrinogen-like, C-terminal
ISCW013767	alkaline phosphatase, putative			3.52		Gut-microbe homeostasis	Thioredoxin-like superfamily
ISCW023785	Mannose-binding endoplasmic reticulum-golgi intermediate compartment lectin [Source:UniProtKB/TrEMBL;Acc:B7P364]			4.46		Proteases/Protease inhibitors	Alkaline-phosphatase-like, core domain superfamily
ISCW016179	Cu ²⁺ /Zn ²⁺ superoxide dismutase SOD1, putative [Source:UniProtKB/TrEMBL;Acc:B7Q898]			2.96		Agglutination	Concanavalin A-like lectin/glucanase domain superfamily
ISCW011852	Defensin, putative [Source:UniProtKB/TrEMBL;Acc:B7QGE4]			4.62		Free radical defense	Superoxide dismutase, copper/zinc binding domain
ISCW022102				5.64		Defensins	Defensin, invertebrate/fungal

Table 4. Co-expressed transcripts and proteins at 1 D PBM. List of gene identifiers co-expressed and differentially expressed at 1 D PBM when compared to unfed timepoint, their transcript or protein expression, protein family, and their respective gene ontology.

Gene ID	Product Description	Transcriptome Log2FC 1D PBM	Proteome Log2FC 1D PBM	P FamDescription
ISCW007318	Cytochrome P450, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PS65]	7.14	2.29	Cytochrome P450
ISCW024299	hemilipoglyco-carrier protein	7.03	-3.51	N/A
ISCW020176	AMP dependent CoA ligase, putative [Source:UniProtKB/TrEMBL;Acc:B7Q394]	6.64	6.64	AMP-dependent synthetase/ligase
ISCW012383	Secreted protein, putative	5.72	3.25	N/A
ISCW011065	Mucin [Source:UniProtKB/TrEMBL;Acc:B7QEW9]	5.25	4.24	von Willebrand factor, type D domain
ISCW023551	Salivary lipid interacting protein, putative [Source:UniProtKB/TrEMBL;Acc:B7QH22]	5.22	3.55	MID-2-related lipid-recognition domain
ISCW015172	Serin proteinase, putative [Source:UniProtKB/TrEMBL;Acc:B7QNU3]	4.85	2.70	Serine proteases, trypsin domain
ISCW005304	Short chain type dehydrogenase, putative [Source:UniProtKB/TrEMBL;Acc:B7PMQ6]	4.83	-2.78	N/A
ISCW003178	histone H3, putative	4.58	-3.07	Histone H2AH2B/H3
ISCW003840	Diacylglycerol O-acyltransferase, putative [Source:UniProtKB/TrEMBL;Acc:B7PJ5]	4.51	-3.64	Diacylglycerol acyltransferase
ISCW011856	glucosylceramidase, putative	4.40	-2.88	Glycosyl hydrolase family 30, TIM-barrel domain;Glycosyl hydrolase family 30, beta sandwich domain
ISCW009053	Fatty acid synthase, putative [Source:UniProtKB/TrEMBL;Acc:B7PYN9]	4.37	2.17	Alcohol dehydrogenase, C-terminal;Beta-ketoacyl synthase, N-terminal;Phosphoantethine binding ACP domain;Acyl transferase;Thioesterase;Beta-ketoacyl synthase, C-terminal;Polyketide synthase, ketoreductase domain;Polyketide synthase, dehydratase domain;Ketoacyl-synthetase, C-terminal extension
ISCW007768	Serin proteinase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PRN3]	4.15	-6.64	Serine proteases, trypsin domain
ISCW018902	uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7PQW9]	3.76	3.36	N/A
ISCW017166	Lysosomal acid phosphatase, putative [Source:UniProtKB/TrEMBL;Acc:B7PCV9]	3.59	-3.29	Histidine phosphatase superfamily, clade-2
ISCW006582	uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7PQJ1]	3.23	-4.38	N/A
ISCW022358	Calcium-binding protein, putative [Source:UniProtKB/TrEMBL;Acc:B7QF77]	3.18	-3.31	SMP-30/Gluconolactonase/LRE-like region
ISCW009705	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7Q0Q1]	3.18	2.54	N/A
ISCW001657	IMP dehydrogenase, putative	3.11	2.52	IMP dehydrogenase/GMP reductase;CBS domain
ISCW000413	Adenylylsulfate kinase, putative [Source:UniProtKB/TrEMBL;Acc:B7P3J6]	3.06	4.97	Sulphate adenylyltransferase catalytic domain;ATP-sulfurylase PUA-like domain
ISCW022787	uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7QCR1]	3.03	2.88	PDZ domain
ISCW003975	NADP-dependent isocitrate dehydrogenase, putative	3.01	-2.91	isopropylmalate dehydrogenase-like domain
ISCW000849	Sele nophosphate synthase, putative [Source:UniProtKB/TrEMBL;Acc:B7P2I9]	2.94	-2.56	PurM-like, N-terminal domain;PurM-like, C-terminal domain

Gene ID	Product Description	Transcriptome Proteome		PFam Description
		Log2FC 1 D PBM	Log2FC 1 D PBM	
ISCW004612	Long chain fatty acid CoA ligase, putative [Source:UniProtKB/TrEMBL;Acc:B7PJA8]	2.94	2.48	AMP-dependent synthetase/ligase
ISCW016514	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PD64]	2.88	2.43	DnaJ domain;Tetratricopeptide repeat
ISCW023966	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QNM9]	2.75	3.36	Pseudouridine-5'-phosphate glycosidase
ISCW014125	Laminin beta 1 chain, putative [Source:UniProtKB/TrEMBL;Acc:B7QHE0]	2.63	2.12	Laminin EGF domain;Laminin, N-terminal
ISCW015316	Fatty acid-binding protein FABP, putative [Source:UniProtKB/TrEMBL;Acc:B7QMW0]	2.60	-3.38	Lipocalin/cytosolic fatty-acid binding domain
ISCW018425	Secreted protein, putative [Source:UniProtKB/TrEMBL;Acc:B7PPI3]	2.58	2.26	Endoplasmic reticulum resident protein 29, C-terminal;ERp29, N-terminal
ISCW022766	Tumor rejection antigen (Gp96), putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QC85]	2.50	2.72	Heat shock protein Hsp90 family;Histidine kinase/HSP90-like ATPase
ISCW005690	Hf62 protein, putative [Source:UniProtKB/TrEMBL;Acc:B7PKW5]	-2.45	3.58	W2 domain
ISCW022612	Ubiquitin carboxy-terminal hydrolase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QDB1]	-2.52	2.68	Peptidase C19, ubiquitin carboxy-terminal hydrolase;Peptidase C19, ubiquitin-specific peptidase, DUSP domain;Ubiquitin-like domain, USP-type
ISCW018085	Malate dehydrogenase, putative [Source:UniProtKB/TrEMBL;Acc:B7PEG0]	-2.57	-2.93	Malate/L-lactate dehydrogenase-like
ISCW013297	Uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7QA66]	-2.58	3.67	Caprin-1 dimerization domain
ISCW001274	Medium-chain acyl-CoA dehydrogenase, putative [Source:UniProtKB/TrEMBL;Acc:B7P3G6]	-3.15	-6.06	Acyl-CoA dehydrogenase/oxidase C-terminal;Acyl-CoA oxidase/dehydrogenase, central domain;Acyl-CoA dehydrogenase/oxidase, N-terminal
ISCW017094	Argininosuccinate synthase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PA12]	-3.32	-3.78	Argininosuccinate synthase
ISCW009338	EGF family domain-containing protein ATP-dependent RNA helicase, putative [Source:UniProtKB/TrEMBL;Acc:B7PA01]	-3.54	-4.29	EGF-like domain
ISCW002701	ABC transporter, putative Dtdp-glucose 4-6-dehydratase, putative [Source:UniProtKB/TrEMBL;Acc:B7QCT0]	-3.78	4.00	DEAD/DEAH box helicase domain;Helicase, C-terminal
ISCW012637	Sulfotransferase, putative [Source:UniProtKB/TrEMBL;Acc:B7PD99]	-4.17	-3.92	N/A
ISCW003098	Chitinase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QD76]	-4.72	-4.13	NAD-dependent epimerase/dehydratase
ISCW012984	Cytochrome P450, putative [Source:UniProtKB/TrEMBL;Acc:B7QA82]	-5.01	-4.88	Sulfotransferase domain
ISCW022071	Cysteine proteinase cathepsin L, putative [Source:UniProtKB/TrEMBL;Acc:B7P3N8]	-5.73	3.33	Glycoside hydrolase family 18, catalytic domain
ISCW000076	esterase, putative	-6.42	-4.97	Cytochrome P450
ISCW022961		-7.40	-6.64	Peptidase C1A, papain C-terminal
		-8.59	-6.64	Carboxylesterase_type B

Table 5. Co-expressed transcripts and proteins at 7 D PBM. List of gene identifiers co-expressed and differentially expressed at 7 D PBM when compared to unfed timepoint, their transcript or protein expression, protein family, and their respective gene ontology.

GeneID	Product Description	Transcriptome		Proteome	PFam Description
		Log2FC 7 D PBM	Log2FC 7 D PBM		
ISCW012383	Secreted protein, putative [Source:UniProtKB/TrEMBL;Acc:B7QEW8]	7.50	4.51		N/A
ISCW020176	AMP dependent CoA ligase, putative [Source:UniProtKB/TrEMBL;Acc:B7Q394]	6.69	6.64		AMP-dependent synthetase/ligase
ISCW017659	eIF2-interacting protein ABC50, putative [Source:UniProtKB/TrEMBL;Acc:B7PJP5]	-2.09	4.77		ABC transporter-like;ABC-transporter extension domain
ISCW007318	Cytochrome P450, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7P565]	6.13	3.58		Cytochrome P450
ISCW023551	Salivary lipid interacting protein, putative [Source:UniProtKB/TrEMBL;Acc:B7QH2]	5.88	4.08		MD-2-related lipid-recognition domain
ISCW001396	salivary protein 14 kDa precursor, putative Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7P6D4]	4.97	4.18		N/A
ISCW016514	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QNM9]	3.93	2.27		DnaJ domain;Tetratricopeptide repeat
ISCW023966	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QNM9]	3.89	4.03		Pseudouridine-5'-phosphate glycosidase
ISCW022787	uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:B7QCR1]	3.32	3.21		PDZ domain
ISCW001590	U5 snRNP-associated RNA splicing factor, putative [Source:UniProtKB/TrEMBL;Acc:B7P571]	-2.18	3.50		Prp18;Pre-mRNA processing factor 4 (PRP4)-like
ISCW002701	ATP-dependent RNA helicase, putative [Source:UniProtKB/TrEMBL;Acc:B7PA01]	-3.96	3.22		DEAD/DEAH box helicase domain;Helicase, C-terminal
ISCW015946	Zinc finger protein, putative [Source:UniProtKB/TrEMBL;Acc:B7P4Y3]	-2.08	3.21		ZC3H15/TMA46 family, C-terminal
ISCW009705	Uncharacterized protein (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QQQ1]	3.28	2.56		N/A
ISCW005690	Hfb2 protein, putative [Source:UniProtKB/TrEMBL;Acc:B7PKW5]	-2.26	2.65		W2 domain
ISCW001976	NAD(P) transhydrogenase, putative [Source:UniProtKB/TrEMBL;Acc:B7PCD1]	3.21	2.30		Alanine dehydrogenase/pyridine nucleotide transhydrogenase, NAD(H)-binding domain;NADP transhydrogenase beta-like domain;Alanine dehydrogenase/pyridine nucleotide transhydrogenase, N-terminal
ISCW014118	secreted protein, putative	2.93	2.48		N/A
ISCW008601	High-density lipoprotein-binding protein, putative [Source:UniProtKB/TrEMBL;Acc:B7Q368]	2.64	4.31		K Homology domain, type 1
ISCW000413	Adenylylsulfate kinase, putative [Source:UniProtKB/TrEMBL;Acc:B7P3J8]	2.47	5.10		Sulphate adenylyltransferase catalytic domain;ATP-sulfurylase PUA- like domain
ISCW022766	Tumor rejection antigen (Gp96), putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7QC85]	2.38	2.59		Heat shock protein Hsp90 family;Histidine kinase/HSP90-like ATPase
ISCW001657	IMP dehydrogenase, putative	2.36	2.38		IMP dehydrogenase/GMP reductase;CBS domain
ISCW003975	NADP-dependent isocitrate dehydrogenase, putative	2.69	-2.00		isopropylmalate dehydrogenase-like domain
ISCW015316	Fatty acid-binding protein FABP, putative [Source:UniProtKB/TrEMBL;Acc:B7QMW0]	2.43	-2.31		Lipocalin/cytosolic fatty-acid binding domain
ISCW001016	Reductase, putative [Source:UniProtKB/TrEMBL;Acc:B7P526]	2.22	-2.40		N/A

GeneID	Product Description	Transcriptome		Proteome	PFam Description
		Log2FC 7 D PBM	Log2FC 7 D PBM		
ISCW010163	Calcium-binding protein, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PYU7]	2.85	-2.41	-2.41	SMP-30/Gluconolactonase/LRE-like region
ISCW007743	flagelliform silk protein, putative Thimet oligopeptidase, putative [Source:UniProtKB/TrEMBL;Acc:B7PC82]	-2.90	-2.81	-2.81	LSM domain, eukaryotic/archaea-type
ISCW017437	ADP/ATP translocase, putative [Source:UniProtKB/TrEMBL;Acc:B7Q4Q3]	-2.22	-2.92	-2.92	Peptidase M3A/M3B catalytic domain
ISCW021753	DNA replication factor/protein phosphatase inhibitor SET/SPR-2, putative [Source:UniProtKB/TrEMBL;Acc:B7Q573]	-2.70	-2.97	-2.97	Mitochondrial substrate/solute carrier
ISCW021350	Short chain type dehydrogenase, putative [Source:UniProtKB/TrEMBL;Acc:B7PMQ6]	-2.60	-3.10	-3.10	Nucleosome assembly protein (NAP)
ISCW005304	Argininosuccinate synthase, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7PA12]	3.28	-3.11	-3.11	N/A
ISCW017094	Beta-N-acetylhexosaminidase, putative [Source:UniProtKB/TrEMBL;Acc:B7PE22]	-3.40	-3.25	-3.25	Argininosuccinate synthase
ISCW003575	ABC transporter, putative Sulfotransferase, putative [Source:UniProtKB/TrEMBL;Acc:B7PD99]	2.99	-3.41	-3.41	Glycoside hydrolase family 20, catalytic domain
ISCW024700	Hsp70, putative (Fragment) [Source:UniProtKB/TrEMBL;Acc:B7XP4]	-4.72	-3.56	-3.56	N/A
ISCW003098	EGF family domaincontaining protein Cytochrome P450, putative [Source:UniProtKB/TrEMBL;Acc:B7QA82]	-6.91	-3.63	-3.63	Sulfotransferase domain
ISCW020763	glycogen phosphorylase, putative Acetyl-CoA hydrolase, putative [Source:UniProtKB/TrEMBL;Acc:B7QJ4]	-7.73	-4.11	-4.11	Heat shock protein 70 family
ISCW009338	Cysteine protease cathepsin L, putative [Source:UniProtKB/TrEMBL;Acc:B7P3N8]	-6.78	-4.57	-4.57	EGF-like domain
ISCW022071	esterase, putative	-6.38	-4.68	-4.68	Cytochrome P450
ISCW000885		-2.78	-4.88	-4.88	Glycosyl transferase, family 35
ISCW023591		-3.62	-5.38	-5.38	Acetyl-CoA hydrolase/transferase;Acetyl-CoA hydrolase/transferase C-terminal domain
ISCW000076		-7.03	-6.64	-6.64	Peptidase C1A, papain C-terminal
ISCW022961		-10.76	-6.64	-6.64	Carboxylesterase, type B

Table 6. Primer sequences used with T7 promoter for RNAi. Transcript ID, sequence, and product size of transcripts.

Primer sequences for RNAi		
Ixodes scapularis gastric triacylglycerol lipase, putative, mRNA		
NCBI Reference Sequence: XM_002413665.1		
GTL1_F481	TGTGCGTTAGGTCCAGTGTC	613bp
GTL1_R1093	TGTAGCTCGCCATCAGCTTC	
Ixodes scapularis heme lipoprotein precursor, putative, mRNA		
NCBI Reference Sequence: XM_002411391.1		
HLP1_F890	TGGAGGAACTGAAGCACGTC	730bp
HLP1_R1619	TACTCCGGCTGGTTTGTGTC	
Ixodes scapularis paramyosin-1, putative, mRNA		
NCBI Reference Sequence: XM_002407245.1		
Para1_1F982	ACTGAAGTCCAGGCTCATGC	749bp
Para1_1R1730	AGTTCGGTGATCTGCACCTG	
Ixodes scapularis paramyosin-2, putative, mRNA		
NCBI Reference Sequence: XM_002406950.1		
Para2_1F1262	ACGTTCAAGTGCTCCAGACC	649bp
Para2_1R1910	TCAGTCCTCAGTTTCGCCAG	
Ixodes scapularis paramyosin-3, putative, mRNA		
NCBI Reference Sequence: XM_002407356.1		
Para3_2F298	GTGCAGGAACTTCTTTACGGC	604bp
Para3_2R901	CGTCAGACACCCTCTCCTTC	
Ixodes scapularis troponin I protein, putative, mRNA		
NCBI Reference Sequence: XM_002415740.1		
Trop1_1F143	CTGAGGAGCTCAAGAAGGAGC	406bp
Trop1_1R548	GCCCATTCGGGTTTGTCCAC	
Ixodes scapularis chitinase, putative, mRNA		
NCBI Reference Sequence: XM_002399485.1		
Chiti_1F397	AAAATTGCGTCGAGTGAGCG	697bp
Chiti_1R1093	CCTTATTCGCGAGGCTTTCG	
Peritrophic Membrane Chitin Binding Protein		
NCBI Reference Sequence: XM_002402097.1		
IscapPerMCB1F49	GGTCGTCAAACATTTCGGCTC	624bp
IscapPerMCB1R672	ACGGTTGTGCGGGATAGATG	

Table 7. Knockdown results of transcript and protein candidates via RNA

interference. The percentage of ticks that survived injection, fed to repletion, the average number of days on the host, and average egg mass (in milligrams) are shown. Rows highlighted in light blue did not produce any eggs. Ticks injected with water were used as control.

Transcripts and/or Proteins used in RNAi				
Gene	%Survived	% Fed to Repletion	Avg. days on host	Avg. Egg mass (mg)
Control	85	66.6	6.7	81.83
Chitinase	71.4	80	4.7	31.83
Cystatin	37.5	66.6	6.7	49.47
Gastric Lipase	75	66.6	6.1	45.06
Heme Lipoprotein Precursor	62.5	40	5.2	59.15
Paramyosin 1	25	50	5.0	0
Paramyosin 2	50	75	5.0	73.5
Paramyosin 3	56.3	77.8	5.0	64.95
Troponin I	87.5	14.2	3.0	0
MBCP	75	65	5.2	44.37
SP1	90	55	5.0	53.96
SP2	88	11	6.0	0
SP4	90	11.4	5.0	22.3

Chapter 3

**Blood Digestion by Trypsin-Like Serine Proteases in the Replete Lyme Disease
Vector Tick, *Ixodes scapularis***

This chapter is based on work published in:

Reyes, J.; Ayala-Chavez, C.; Sharma, A.; Pham, M.; Nuss, A.B.; Gulia-Nuss, M. Blood Digestion by Trypsin-Like Serine Proteases in the Replete Lyme Disease Vector Tick, *Ixodes scapularis*. *Insects* 2020, 11, 201. <https://doi.org/10.3390/insects11030201>

3.1 Abstract

Ixodes scapularis is the major vector of Lyme disease in the Eastern United States. Each active life stage (larva, nymph, and adult) takes a blood meal either for developing and molting to the next stage (larvae and nymphs) or for oviposition (adult females). This protein-rich blood meal is the only food taken by *Ixodes* ticks and therefore efficient blood digestion is critical for survival. Studies in partially engorged ticks have shown that the initial stages of digestion are carried out by cathepsin proteases within acidic digestive cells. In this study, we investigated the potential role of serine proteases in blood digestion in replete ticks. RNA interference was used for functional analysis and a trypsin-benzoyl-D, L-arginine 4-nitroanilide assay was used to measure active trypsin levels. Hemoglobinolytic activity was determined *in vitro*, with or without a serine protease inhibitor. Our data suggest that trypsin levels increase significantly after repletion. Knockdown of serine proteases negatively impacted blood-feeding, survival,

fecundity, levels of active trypsin in the midgut, and resulted in lower hemoglobin degradation. Incubation of midgut extract with a trypsin inhibitor resulted in 65% lower hemoglobin degradation. We provide evidence of the serine proteases as digestive enzymes in fully engorged, replete females. Understanding the digestive profile of trypsin during blood meal digestion in *I. scapularis* improves our understanding of the basic biology of ticks and may lead to new methods for tick control.

3.2 Introduction

Ixodes scapularis is a three-host tick that requires a blood meal to complete each developmental stage, in addition to adult females needing blood for egg development (Sonenshine 1992). The larvae and nymphs feed for 3–7 days whereas adult female feeding lasts for up to 10 days and consists of: (i) a slow feeding period up to 5–9 days post-attachment followed by (ii) rapid engorgement for 12–24 h before detachment from the host (Gulia-Nuss et al., 2016). Rapid engorgement accounts for about two-thirds of the total blood meal. The tick midgut comprises a major portion of the body and consists of a ventriculus (stomach) and several pairs of highly branched ceca extending into all regions of the body. Blood digestion putatively starts in the midgut soon after ingestion and continues for several days to weeks after dropping off the host. Proteins represent about 95% of the non-water content of vertebrate blood. Consequently, hematophagous arthropods require proteases as the main enzymes in the midgut to process a blood meal (Mulenga et al., 2011).

A typical animal genome contains 2–4% of genes encoding for proteolytic enzymes (Puente et al., 2005). Among these, serine proteases are the most abundant and functionally diverse group (Page et al., 2008). Over one-third of all known proteolytic

enzymes are serine proteases (Page et al., 2008). Out of a total of 233 putatively active *I. scapularis* proteases thus far identified, 63 (27%) are serine proteases (Mulenga et al., 2011). Hematophagous insects such as tsetse flies, mosquitoes, and many others digest the protein-rich blood meal mainly by using trypsin-like serine proteases that have a pH optimum in the alkaline range (~8.0 pH) (Santiago et al., 2017). In insects, protein digestion proceeds rapidly and takes place in the midgut lumen. The processing of host blood components in the tick midgut, however, appears to differ greatly from that in other hematophagous arthropods. In ticks, blood digestion is a slow process that occurs in the acidic environment of midgut intracellular vesicles (endosomes), mainly by the cathepsin-like proteases (Mulenga et al., 2011).

A multi-enzyme model for hemoglobin degradation was proposed for the European vector of Lyme disease, the castor bean tick, *I. ricinus* (Sojka et al., 2008). According to this model, the hemoglobin degradation pathway is initiated inside acidic digestive vesicles by cysteine and aspartic endopeptidases (cathepsin L, legumain, and cathepsin D), generating large peptide fragments (8–11 kDa), followed by the action of cathepsin B and C exopeptidases, generating smaller peptides (2–7 kDa). Finally, serine carboxypeptidase and leucine aminopeptidase may participate in the liberation of dipeptides and free amino acids. Other studies have suggested that the final stages of hemoglobin degradation take place both in and outside of the digestive vesicles (Horn et al., 2009). Trypsin-like serine proteases are active at high pH (Calvo et al., 2008), in contrast to the acidic-active cathepsins. Midgut homogenates of the hard tick, *I. scapularis* (formerly *I. dammini*), were shown to lyse erythrocytes from vertebrate blood at an alkaline pH, suggesting the involvement of trypsin enzymes. Ten major blood

digestive proteases (cathepsin, aminopeptidase, and serine proteases) were proposed to be involved in blood digestion in *I. scapularis* (Gulia-Nuss et al., 2016). The presence of four serine proteases on this list suggests previously unexplored roles of trypsin during *I. scapularis* blood ingestion.

Most studies in tick blood digestion have focused on partially engorged females (up to 5 days on host, slow feeding phase), resulting in little information on the digestive profile beyond this stage. Therefore, to better understand the digestive enzyme profile of replete *I. scapularis*, we first characterized the expression of ten proteases identified previously (Gulia-Nuss et al., 2016) and then measured trypsin activity in unfed, partially-fed, and post host detachment (also post-blood meal; hereafter PBM) ticks for 7 days (adults) and 28 days (larvae and nymphs) using benzoyl-D, L-arginine 4-nitroanilide (BApNA), a trypsin-specific substrate (Rascon et al., 2011, Gulia-Nuss et al., 2011)]. RNA interference was used for the functional analysis of three serine protease enzymes. Our data indicate that trypsin levels increase significantly after repletion and individual knockdown of serine proteases affect the proteolytic activity of midgut extracts, suggesting that these enzymes play a major, previously unrecognized, role in blood digestion in ticks PBM.

3.3 Materials and Methods

Tick Feeding and Sample Collection

Unfed adult pathogen-free *I. scapularis* were purchased from the tick rearing facility at the Oklahoma State University, Stillwater, OK. Ticks were maintained until use in an incubator at 95% relative humidity (RH) and 20 °C.

Larvae and nymphs were fed on mice at the University of Nevada, Reno (Nuss et al., 2017). Both stages were allowed to detach naturally PBM. Detached ticks were collected daily, returned to the incubator, and harvested at the appropriate PBM intervals. Larvae and nymphs were processed at 1, 2, 3, 7, 14, 21, and 28 days PBM. Two larvae or nymphs per sample were collected in triplicate for each time point. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Nevada, Reno (IACUC # 21-01-1118).

Adult ticks were fed on New Zealand white rabbits. Male and female adults were confined in capsules attached to the rabbit's back using Osto-Bond skin bonding latex adhesive (Montreal Ostomy, Canada). A published protocol (Almazán et al., 2018) was used with a few modifications. Briefly, feeding capsules were made out of 1.5" PVC tube with a Styrofoam lip glued for firm attachment to the rabbit's skin. Five to six capsules were attached to each rabbit one day prior to releasing ticks. A fine nylon mesh screen was glued to the top with a small opening secured with Velcro for ease of tick removal PBM. Rabbits were fitted with a collar for the duration of tick feeding. *Ixodes scapularis* adult females were collected 5 days post-host attachment (partially-fed), and at 1, 2, 3, 7, and 14 days PBM. Whole midguts were dissected in cold PBS buffer, washed (to remove blood), and two midguts were pooled per sample. For unfed samples, four midguts were pooled. Experiments were replicated with three biological cohorts. Once washed, midguts were immediately transferred to either a cold 1.7 mL tube containing 200 μ L of Trizol or Tris-HCl-CaCl₂ and stored at -80°C until processed.

Protease Expression

Total RNA was extracted from midguts or whole ticks (larvae and nymphs) using Trizol reagent and a Zymo Direct-Zol RNA purification kit (Zymo Research, Irvine, CA, USA) with DNase treatment step. A quantity of 1 µg DNase-treated RNA was used for cDNA synthesis, according to the kit protocol (iScript, BioRad, Hercules, CA, USA). For RT-PCR, 1 µL of 1:10 diluted cDNA was used as a template in a 20 µL reaction. RT-PCR conditions for all digestive enzymes were: initial denaturation at 95 °C for 5 min, 95 °C for 30 s, 55–58 °C for 30 s (Table 1), 72 °C for 30 s, repeated for 35 cycles, and a final extension at 72 °C for 10 min. 10 µL PCR product was separated by electrophoresis on a 1.2% agarose gel along with a DNA ladder (Apex DNA Ladder II; Genesee, San Diego, CA, USA) and visualized by using ethidium bromide-free dye (Amresco, Solon, OH, USA). Primer sequences for all proteases are listed in Table 1. Tubulin was used as a housekeeping control (Sharma et al., 2019).

Gel band intensity was analyzed by densitometry using Image Lab 5.2.1 (Gel Doc EZ-Imager, BioRad). The DNA ladder was used as the standard for generating a linear regression model to determine PCR product abundance. The ratio of the tubulin control band intensity was used to standardize the values of the band from each gene at different time points. After calculating the ratios, the band with the highest expression in each gene was set at 1 and the values were adjusted proportionally and plotted to depict the change in expression of each gene over time.

Sample Preparation for BApNA Assay

Six larvae or nymphs from each PBM time point were collected in triplicates (two individuals per triplicate). Samples were sonicated in 100 µL Tris-HCl-CaCl₂ buffer, pH

7.5, until completely homogenized and centrifuged at 12,000 rpm for 5 min at 4 °C. 10 µL supernatant from each sample was added to 90 µL of Tris-HCl-CaCl₂ (Horn et al., 2009) followed by 200 µL of 4 mM BApNA. Serial dilutions of trypsin (Sigma) were similarly using 10 µL trypsin, 90 µL of Tris-HCl-CaCl₂, and 200 µL of 4 mM BApNA were used to make a standard curve. Samples and standards were then placed on a shaker for 15 min at 25 °C, loaded onto a 96 well plate (100 µL sample per well), and absorbance at 405 nm was recorded on a Spectramax M5 microplate reader. For adults, midguts were dissected and pooled from two females per sample and samples were prepared and measured as above.

RNA Interference (RNAi)

dsRNA was synthesized for three serine proteases (SP): SP1 (ISCW021184), SP2 (ISCW006427), and SP4 (ISCW007492). Total RNA was extracted from unfed or 1-day PBM tick midguts and cDNA synthesis as described above. Primers were designed with a T7 promoter sequence (5'-TAATACGACTCACTATAGGG-3') on the 5' end of both forward and reverse primers (Table 1). PCR conditions were the same as mentioned above in the RT-PCR section. Gel bands were extracted using the QIAquick gel extraction kit (Qiagen, Hilden, Germany) and used as a template for dsRNA synthesis using the T7 Megascript kit (Invitrogen, Carlsbad, CA, USA). Newly synthesized dsRNA was purified using phenol-chloroform and ethanol precipitated. Sixteen unfed adult female ticks per gene were injected with 1 µL of dsRNA (2 µg/µL). Injections were performed with a u-200 insulin syringe on the ventral right side between the 3rd and 4th leg of the tick. Control ticks were injected with 1 µL of RNase/DNase/Protease-free water (11 ticks). Injected ticks were immediately placed in a holding container at 95%

RH and observed for 2 h for recovery before storage. Ticks were allowed to recover for 7 d before placing them on New Zealand white rabbits (see adult feeding procedure above). Detached ticks were collected daily, weighed, photographed immediately after dropping off, and stored in individual containers in the colony incubator at 20 °C and 95% RH. A batch of females from each treatment was kept for fecundity assessment. Females were observed daily for mortality and egg-laying. Egg mass was weighed once females stopped laying eggs.

For the BApNA assay, six dsRNA-injected adult females from each of the three serine protease RNAi treatments were collected 1-day PBM. Control ticks were collected at the same time. Two midguts per sample were dissected (N = 3) and the assay was carried out as described above.

Hemoglobin Degradation Assay

Midguts were dissected individually from control and RNAi females collected 1-day PBM, washed, and homogenized with a pestle in 0.1 M sodium acetate, 1% CHAPS, and 2.5 mM DTT [8]. Midgut extracts were centrifuged at 16,000× g for 10 min at 4 °C, filtered with a 0.22 µm polyethersulfone (PES) membrane syringe filter (Olympus, Sigma-Aldrich, MO), and stored at −80 °C until used for assays. Protein concentration was measured using the BCA protein assay kit (Thermo Fisher, Waltham, MA, USA). A quantity of 0.5 µg protein extract was used to digest 10 µg of bovine hemoglobin in 25 mM Na-citrate-phosphate (pH 7.5), 2.5 mM DTT, and 25 mM NaCl. Bovine trypsin (Sigma Aldrich, St. Louis, MO, USA) was used as a control for hemoglobin digestion. Aliquots were taken out at 0-, 10-, 20-, and 30-min post-incubation. 0.03% fluorescamine (Biotium, Fremont, CA, USA) in acetone was added to the midgut extract-hemoglobin

reaction to quantify the newly formed amino-terminal ends (Sorgine et al., 2000). Fluorescence was measured using a Spectramax M5 microplate reader at an excitation of 370 nm and emission of 485 nm wavelengths. Measurements were performed in triplicate. For trypsin activity inhibition, midgut extracts were pre-incubated with 0.1 mM PMSF (Research Products International, Mt Prospect, IL, USA) for 15 min at 37 °C before adding hemoglobin.

Statistical Analysis

All experiments were replicated a minimum of three times with different biological cohorts. One-way ANOVA and Dunnett's multiple comparisons were used for the analysis of trypsin activity data. For the rest, a t-test with Welch's correction was used in GraphPad (GraphPad Software, La Jolla, CA, USA). Mortality was recorded each day in control as well as three other genes of interest and then Kaplan–Meier survival analysis was performed using a Log-rank (Mantel-Cox) test to perform a pairwise comparison using control survival as reference.

3.4 Results

Transcript Expression

Transcripts of 10 proteases identified as the main proteolytic enzymes for hemoglobin degradation in *I. scapularis* (Gulia-Nuss et al., 2016) were examined in the adult female midgut. Out of these 10, four were serine proteases (ISCW021184, ISCW006427, ISCW010371, and ISCW007492), two were cathepsin L (ISCW024899 and ISCW000076), and one each were cathepsin C (ISCW03494), cathepsin D (ISCW023880), legumain (ISCW015983), and leucine aminopeptidase (ISCW023735). All sequences were confirmed by Sanger sequencing. Expression was determined at

different time points: partially engorged adult females (collected 5 days post-host attachment but before rapid engorgement), and at 1, 2, 7, and 14 days PBM (fully engorged and actively digesting blood to provision developing eggs). Out of 10 genes tested, six were expressed in unfed samples whereas four genes [ISCW023880 (cathepsin D), ISCW024899 (cathepsin L), ISCW007492 (serine protease), and ISCW023735 (leucine aminopeptidase)] were only expressed during feeding or PBM (Figure 1a,b). Cathepsin D expression was highest at 1-day PBM and decreased afterward. One cathepsin L paralog (ISCW024899) was expressed in all blood-fed stages tested, from partially engorged to 14 days PBM, and the other cathepsin L (ISCW000076) was expressed in unfed and partially fed females with almost no expression in fully engorged females. Legumain was expressed at low levels in unfed females, then expression increased during and after feeding. Cathepsin C was also expressed at low levels in unfed females and expression was higher during and after a blood meal; peak expression was at 2 and 7 days PBM and decreased at 14 days PBM. Serine proteases 1 and 2 (SP 1 and 2; ISCW021184 and ISCW006427) had a similar expression pattern in unfed, partially fed females, and at 14 days PBM. Serine protease 3 (SP3; ISCW010371) had low expression in unfed and 14 days PBM samples. Serine protease 4 (SP4; ISCW007492) was not detected in unfed females, but the expression was detected during and after blood-feeding with the highest expression at 1-day PBM. Leucine aminopeptidase expression was not detected in unfed females and peak expression was in partially engorged females. Expression decreased afterward and no expression was detected at 14 days PBM (Figure 1a,b).

Active Trypsin in Tick life stages

Trypsin activity increased significantly in larvae after a blood meal (Figure 2a). The peak trypsin activity period was 1–3 days PBM and levels decreased gradually afterward to nearly unfed levels at 21–28 days PBM (Figure 2a). In our colony, larvae begin molting into nymphs within 3–4 weeks; therefore, the 28-day time-point coincides with molting. In nymphs, a similar pattern was observed. Trypsin levels increased after a blood meal and were highest at 3 days PBM. Subsequently, levels decreased and by 28 days trypsin levels were similar to those in unfed nymphs (Figure 2b).

In adult midguts, no trypsin activity was detected in unfed or partially fed ticks. Trypsin activity was highest after drop-off from the host (1-day PBM) and decreased gradually. By 14 days PBM, trypsin activity returned to the unfed levels (Figure 2c). Under our rearing conditions, ticks start laying eggs 7–10 days PBM; therefore, most blood digestion occurs during the first two weeks PBM.

Effect of Serine Proteases knockdown on Tick Blood-Feeding and Physiology

RNAi knockdown was successful for three serine proteases (SP1, SP2, and SP4) and knockdown persisted until at least 2 days PBM (a total of 15 days) (Figure 3a). Knockdown of serine proteases by RNAi increased tick mortality. Approximately 10% of ticks died in all treatments during the recovery period post-injection. Once attached to the host, mortality was significantly higher in serine protease depleted ticks compared to the controls. Seventy-seven percent of control ticks survived and were fed to repletion in comparison to SP1 (55%), SP2 (11%), and SP4 (11%) (Figure 3b).

Most mortality occurred between 2–6 days post-attachment (Figure 3b). Serine protease knockdown also resulted in a significantly lower volume of blood ingested, as

indicated by engorgement weight. Control females weighed an average of 200 mg whereas SP1 knockdown weighed 140 mg (29% reduction), SP2 knockdown weighed 110 mg (45% reduction), and SP4 knockdown weighed 78 mg (61% reduction) (Figure 3c). Fecundity was also reduced, as measured by egg mass weight. Egg clutch weight in controls was, on average, 100 mg, but was only 54 mg in SP1 and 22 mg in SP4. SP2 RNAi females did not produce any eggs (Figure 3d). Significantly less overall trypsin activity was detected in the midgut of RNAi females compared to controls at day 1 PBM (peak trypsin activity). Control ticks had an average of 400 μ g trypsin compared to 95 (76% reduction), 240 (40% reduction), and 140 μ g (65% reduction) in SP1, SP2, and SP4 RNAi females, respectively (Figure 3e).

Hemoglobin Degradation by Tick Midgut Extract

Bovine hemoglobin incubated with tick midgut extracts at 7.5 pH resulted in free amino-terminal ends indicative of hemoglobin digestion by the midgut extract. Serine protease RNAi decreased this activity, further suggesting that these proteases are involved in hemoglobin degradation and therefore blood digestion. SP1 and SP2 knockdown resulted in significantly lower hemolytic activity compared to the control at 30 min, with a 29% and 25% reduction in hemoglobin breakdown activity (fluorescent activity), respectively (Figure 4a). SP4 knockdown had the greatest effect on hemoglobin breakdown and was significantly lower than the control throughout the assay, with a 52% reduction in midgut extract hemolytic activity at 30 min incubation (Figure 4a).

To confirm the trypsin-like protease activity in the midgut tissue extract of fully engorged *I. scapularis* females, we incubated midgut extract from water-injected ticks with a trypsin inhibitor, PMSF, prior to the addition of hemoglobin. Midgut extract without

PMSF incubation was used as a control. Incubation with PMSF inhibited the midgut extract activity by 55% (Figure 4b). The hemoglobin degradation activity was significantly lower in PMSF incubated samples starting at 10 min, but was even more evident at 30 min as there was no increase in activity with time compared to the non-PMSF midgut sample (Figure 4b).

3.5 Discussion

In the present work, we examined *I. scapularis* midgut protease expression including four trypsin-like serine protease transcripts. These serine proteases had a different temporal expression: two were expressed in unfed and partially engorged females, one only in unfed, and one in all blood-feeding stages tested but not in unfed midguts (Figure 1a). Since transcript abundance is not a direct measure of enzyme activity and higher transcript abundance does not always result in translation, we measured active trypsin levels. The trypsin assay showed that unfed *I. scapularis* midguts did not have trypsin activity (Figure 2c). However, trypsin activity increases after detachment from the vertebrate host and peaks at 1–3 days PBM in all life stages (Figure 2a–c). Trypsin cleaves p-nitroaniline off BApNA yielding a yellow substrate that was measured at 405 nm. This is a trypsin-specific reaction that does not occur with cathepsins or chymotrypsin (Hayakawa et al., 1980). In *Aedes aegypti* mosquitoes, induction of trypsin biosynthesis after the blood meal is a two-phase process. The first phase of trypsin biosynthesis involves the initiation of translation of an mRNA transcript that is already present, producing early trypsin. The second phase, 7–9 h PBM, is activated by the synthesis of a new mRNA transcript that codes for late trypsin (Felix et

al., 1991). Our transcript and enzyme activity data suggest that similar to mosquito early trypsin, the mRNA for *I. scapularis* trypsin is present in unfed ticks.

Most previous studies have focused on the mechanism of hemoglobin degradation in ticks during the early stages of feeding and suggest that when ticks are actively feeding on the host, the main peptidases for hemoglobin digestion are: (1) clan CA cathepsins B, C, and L; (2) clan CD asparaginyl endopeptidase (legumain); and (3) clan AA cathepsin D. Other activities detected were attributed to mono-peptidases, namely a serine carboxypeptidase and a leucine metallo-aminopeptidase, within midgut digestive vesicles (Lara et al., 2005). Previous studies have also suggested hemoglobin receptor-mediated endocytosis (Santiago et al., 2017, Lara et al., 2005) occurs in the digestive vesicles. Digestive vesicles then lead to further breakdown of these peptides by creating an acidic environment suitable for cathepsin and legumain activity. However, in this study, we investigated blood digestion after repletion and off the host, a relatively unexplored blood digestion phase in ticks. Our data strongly suggest that serine proteases are involved in blood digestion in the PBM phase. Midgut extracts lysed hemoglobin *in vitro* and pre-incubation with trypsin inhibitor reduced this hemolysis activity (Figure 4a,b). The knockdown of three serine proteases individually resulted in lower levels of active trypsin in the BApNA assay (Figure 3e). Serine protease knockdown also resulted in reduced hemoglobin degradation activity *in vitro* (Figure 4a). Other studies in replete ticks have also suggested that trypsin proteases might be involved in blood digestion. Ribeiro (Ribeiro et al., 1988) showed that midgut homogenates of *I. scapularis*, (formerly *I. dammini*) lysed erythrocytes from rabbits, rats, hamsters, and guinea pigs. The midgut homogenate activity was optimal at an alkaline pH (pH 7.5–8.5), suggestive of trypsin-

like serine proteases. This activity was not detected in unfed ticks as well as ticks attached for up to 2 days to a host but rather increased during the latter phase of feeding. It was hypothesized that this initial activity helped process the blood meal by releasing the contents of erythrocytes for further enzymatic hydrolysis, possibly in the digestive vesicles (Ribeiro et al., 1988). Two serine proteases in *Haemaphysalis longicornis* ticks were identified and characterized, and expression of both serine proteases was induced by blood-feeding (Mulenga et al., 2001). In another study, two genes encoding trypsin-like serine proteases, HISP2 and HISP3, in *H. longicornis* were also proposed to be involved in blood digestion (Miyoshi et al., 2007). One of these HISP genes was further characterized and was found to be secreted in the midgut lumen (Miyoshi et al., 2008). Disruption of HISP-specific mRNA by RNAi prevented degradation of host erythrocyte membranes, indicating that HISP plays a crucial role in hemolysis in the midgut of ticks (Miyoshi et al., 2008). Hemolysin-like material was also demonstrated in the midgut lumen of ixodid ticks by immuno-localization (Miyoshi et al., 2008, Hughes, 1954). An RNAseq study comparing blood-fed and serum-fed *I. ricinus* midgut transcriptomes showed that the number of genes encoding serine proteases were markedly up-regulated in the late stage of feeding (Perner et al., 2016) and the possibility of active serine proteases during the off-host stage of blood digestion was suggested. Given these data in other tick species and our results in *I. scapularis*, we propose a modified model of blood digestion in ticks (Figure 5). We suggest that the existing model of blood digestion, which occurs in digestive cells by cathepsins and aminopeptidases, applies during the early digestive phase when the tick is still feeding, whereas in replete females, a revised model is warranted which includes trypsin-like serine proteases that are important for

hemolysis and initial degradation of blood proteins and that digestion may take place in both the midgut lumen and digestive vesicles.

A remarkable property of certain insect midguts is a very high luminal pH, especially in lepidopteran larvae (pH 9–12). In contrast, the midgut of the yellow fever mosquito, *Ae. aegypti*, is acidic (pH 6.0) before a blood meal, which then increases to an alkaline range (pH 7.5) after a blood meal (Nepomuceno et al., 2017). The pH of mite midguts also strongly affects their digestive processes. For instance, the midgut contents of acarid mites ranged from pH 4 to 7 (Erban et al., 2010). *Ixodes scapularis/dammini* midgut homogenate derived from females in the rapid feeding phase had high proteolytic activity at pH 7.5–8.5 (Ribeiro et al., 1988). Both BApNA and hemoglobin degradation assays in this study were carried out at pH 7.5, suggesting that these serine proteases are active at an alkaline pH and provide indirect evidence of an alkaline midgut lumen environment. We attempted to measure pH by homogenizing midguts and using a universal pH paper; our results suggested an alkaline pH, but a refined method of measurement is needed (data not shown). For instance, the pH in midguts of 12 species of the stored product and house dust mites was determined based on the color changes of pH indicators fed to the organisms and looking at pH change microscopically. Unfortunately, this is not feasible with ticks due to the dark cuticle and blood meal coloration. However, microelectrodes are frequently used to determine midgut pH in insects (Brune et al., 1995, Harrison, 2001, Zimmer et al., 2005, Gross et al., 2008), and we plan to utilize these in future experiments.

The knockdown of three serine proteases resulted in ingestion of lower blood volume (Figure 3b) which correlates with lower fecundity (Figure 3d). In *H. longicornis*,

knockdown of three serine proteases, HISP1, 2, and 3, also resulted in a significant reduction in body weight compared to the control group (Miyoshi et al., 2007). In contrast, *I. ricinus* Cathepsin D knockdown did not affect mortality, weight, oviposition, and larvae hatching (Sojka et al., 2012) but knockdown of *I. ricinus* cathepsin L1 (IrCL1) resulted in decreased weight gain in partially engorged females injected with IrCL1 dsRNA relative to the controls (Franta et al., 2011). A cysteine protease, longipain, knockdown in *H. longicornis* also resulted in significantly lower body weight than that of the control ticks (Tsuji et al., 2008). High mortality in SP2 and SP4 knockdown females' post-attachment (Figure 3c), combined with reduced feeding, suggest additional roles in *I. scapularis* physiology that need to be further investigated.

Here, we provide direct evidence of serine proteases as active digestive enzymes that can break down blood proteins in replete ticks. Future experiments will include the use of recombinant *I. scapularis* serine proteases for blood protein digestion assays in vitro. This initial exploration examined the most prominent *I. scapularis* digestive enzymes, which only included 3 out of a putative 63 serine proteases present in the genome (Mulenga et al., 2011). In future studies, a more expansive screen will yield additional information on the dynamics of *I. scapularis* enzymes important in blood digestion.

3.6 References

Almazán, C.; Bonnet, S.; Cote, M.; Slovák, M.; Park, Y.; Šimo, L. A Versatile Model of Hard Tick Infestation on Laboratory Rabbits. *J. Vis. Exp.* **2018**, 140, 57994.

Brune, A.; Emerson, D.; Breznak, J.A. The termite midgut microflora as an oxygen sink-microelectrode determination of oxygen and pH gradients in midguts of lower and higher termites. *Appl. Environ. Microbiol.* **1995**, 61, 2681–2687.

Calvo, E.; Pham, V.M.; Ribeiro, J.M.C. An insight into the sialotranscriptome of the non-blood feeding *Toxorhynchites amboinensis* mosquito. *Insect Biochem. Mol. Biol.* **2008**, 38, 499–507.

Erban, T.; Hubert, J. Determination of pH in regions of the midguts of acarid mites. *J. Insect Sci.* **2010**, 10, 42.

Felix, C.R.; Betschart, B.; Billingsley, P.F.; Freyvogel, T.A. Post-feeding induction of trypsin in the midgut of *Aedes aegypti* L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochem.* **1991**, 21, 197–203.

Franta, Z.; Sojka, D.; Frantova, H.; Dvorak, J.; Horn, M.; JindrichSrba, P.T.; Mares, M.; Schneider, E.; Craik, C.S.; McKerrow, J.H.; et al. IrCL1—The haemoglobinolytic cathepsin L of the hard tick, *Ixodes ricinus*. *Int. J. Parasitol.* **2011**, 41, 1253–1262.

Gross, E.M.; Brune, A.; Walenciak, O. Midgut pH, redox conditions and oxygen levels in an aquatic caterpillar: Potential effects on the fate of ingested tannins. *J. Insect Physiol.* **2008**, 54, 462–471.

Gulia-Nuss, M.; Nuss, A.B.; Meyer, J.M.; Sonenshine, D.E.; Roe, R.M.; Waterhouse, R.M.; Sattelle, D.B.; Fuente, J.; Ribeiro, J.M.; Megy, K.; et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nat. Commun.* **2016**.

- Gulia-Nuss, M.; Robertson, A.E.; Brown, M.R.; Strand, M.R. Insulin-like peptides and the target of rapamycin pathway coordinately regulate blood digestion and egg maturation in the mosquito *Aedes aegypti*. *PLoS ONE* **2011**, *6*, e20401.
- Harrison, J.F. Insect acid-base physiology. *Annu. Rev. Entomol.* **2001**, *46*, 221–250.
- Hayakawa, T.; Kondo, T.; Yamazaki, Y.; Linuma, Y.; Mizuno, R. A simple and specific determination of trypsin in human duodenal juice. *Gastroenterol. Jpn.* **1980**, *15*, 135–139.
- Horn, M.; Nussbaumerova, M.; Sanda, M.; Kovarova, Z.; Srba, J.; Franta, Z.; Sojka, D.; Bogyo, M.; Caffrey, C.R.; Kopacek, P.; et al. Hemoglobin Digestion in Blood-Feeding Ticks: Mapping a Multi-peptidase Pathway by Functional Proteomics. *Chem. Biol.* **2009**, *16*, 1053–1063.
- Hughes, T.E. Some histological changes which occur in the midgut epithelium of *Ixodes ricinus* females during gorging and up to oviposition. *Ann. Trop. Med. Parasitol.* **1954**, *48*, 397–404.
- Lara, F.A.; Lins, U.; Bechara, G.H.; Oliveira, P.L. Tracing heme in a living cell: Hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *J. Exp. Biol.* **2005**, *208*, 3093–3101.
- Miyoshi, T.; Tsuji, N.; Islam, M.K.; Alim, M.A.; Hatta, T.; Huang, X.; Fujisaki, K. A set of serine proteinase paralogs are required for blood-digestion in the ixodid tick *Haemaphysalis longicornis*. *Parasitol. Int.* **2008**, *57*, 499–505.

- Miyoshi, T.; Tsuji, N.; Islam, M.K.; Huang, X.; Motobu, M.; Alim, M.A.; Fujisaki, K. Molecular and reverse genetic characterization of serine proteinase-induced hemolysis in the midgut of the ixodid tick *Haemaphysalis longicornis*. *J. Insect Physiol.* **2007**, *53*, 195–203.
- Mulenga, A.; Erikson, K.A. Snapshot of the *Ixodes scapularis* degradome. *Gene* **2011**, *482*, 78–93.
- Mulenga, A.; Sugimoto, C.; Ingram, G.; Ohashi, K.; Misao, O. Characterization of two cDNAs encoding serine proteinases from the hard tick *Haemaphysalis longicornis*. *Insect Biochem. Mol. Biol.* **2001**, *31*, 817–825.
- Nepomuceno, D.B.; Santos, V.C.; Araujo, R.N.; Pereira, M.H.; Sant'Anna, M.R.; Moreira, L.A.; Gontijo, N.F. pH control in the midgut of *Aedes aegypti* under different nutritional conditions. *J. Exp. Biol.* **2017**, *220*, 3355–3362.
- Nuss, A.B.; Mathew, M.G.; Gulia-Nuss, M. Rearing *Ixodes scapularis*, the Black-legged Tick: Feeding Immature Stages on Mice. *J. Vis. Exp.* **2017**, *123*, e55286.
- Page, M.J.; Di Cera, E. Serine peptidases: Classification, structure and function. *Cell. Mol. Life Sci.* **2008**, *65*, 1220–1236.
- Perner, J.; Provazník, J.; Schrenková, J.; Urbanová, V.; Ribeiro, J.M.; Kopáček, P. RNA-seq analyses of the midgut from blood- and serum-fed *Ixodes ricinus* ticks. *Sci. Rep.* **2016**, *6*, 36695.

Puente, X.S.; Sanchez, L.M.; Midgutierrez-Fernandez, A.; Velasco, G.; Lopez-Otin, C.A. Genomic view of the complexity of mammalian proteolytic systems. *Biochem. Soc. Trans.* **2005**, *33*, 331–334.

Rascón, A.A., Jr.; Gearin, J.; Isoe, J.; Miesfeld, R.L. In vitro activation and enzyme kinetic analysis of recombinant midgut serine proteases from the Dengue vector mosquito *Aedes aegypti*. *BMC Biochem.* **2011**, *12*, 43.

Ribeiro, J.M. The midgut hemolysin of *Ixodes dammini* (Acari: Ixodidae). *J. Parasitol.* **1988**, *74*, 532–537.

Santiago, P.B.; de Araujo, C.N.; Motta, F.N.; Praca, Y.R.; Charneau, S.; Dourado Bastos, I.M.; Santana, J.M. Proteases of haematophagous arthropod vectors are involved in blood-feeding, yolk formation and immunity—A review. *Parasites Vectors* **2017**, *10*.

Sharma, A.; Pooraiouby, R.; Guzman, B.; Vu, P.; Gulia-Nuss, M.; Nuss, A.B. Dynamics of Insulin Signaling in the Black-Legged Tick, *Ixodes scapularis*. *Front. Endocrinol.* **2019**, *10*.

Sojka, D.; Franta, Z.; Frantová, H.; Bartosová, P.; Horn, M.; Váchová, J.; O'Donoghue, A.J.; Eroy-Reveles, A.A.; Craik, C.S.; Knudsen, G.M.; et al. Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). *J. Biol. Chem.* **2012**, *287*, 21152–21163.

Sojka, D.; Franta, Z.; Horn, M.; Hajdusek, O.; Caffrey, C.R.; Mares, M.; Kopacek, P. Profiling of proteolytic enzymes in the midgut of the tick *Ixodes ricinus* reveals an

evolutionarily conserved network of aspartic and cysteine peptidases. *Parasites Vectors* **2008**, 1.

Sonenshine, D.E. *Biology of Ticks*; Oxford University Press: Oxford, UK, 1992; Volume 1, pp. 122–157.

Sorgine, M.H.; Logullo, C.; Zingali, R.B.; Paiva-Silva, G.O.; Juliano, L.; Oliveira, P.L. A Heme-binding Aspartic Proteinase from the Eggs of the Hard Tick *Boophilus microplus*. *J. Biol. Chem.* **2000**, 275, 28659–28665.

Tsuji, N.; Miyoshi, T.; Battsetseg, B.; Matsuo, T.; Xuan, X.; Fujisaki, K. A cysteine protease is critical for *Babesia spp.* transmission in *Haemaphysalis* ticks. *PLoS Pathog.* **2008**, 4, e1000062.

Zimmer, M.; Brune, A. Physiological properties of the midgut lumen of terrestrial isopods (Isopoda: Oniscidea): Adaptive to digesting lignocellulose? *J. Comp. Physiol. B* **2005**, 175, 275–283.

3.7 Figures and Tables

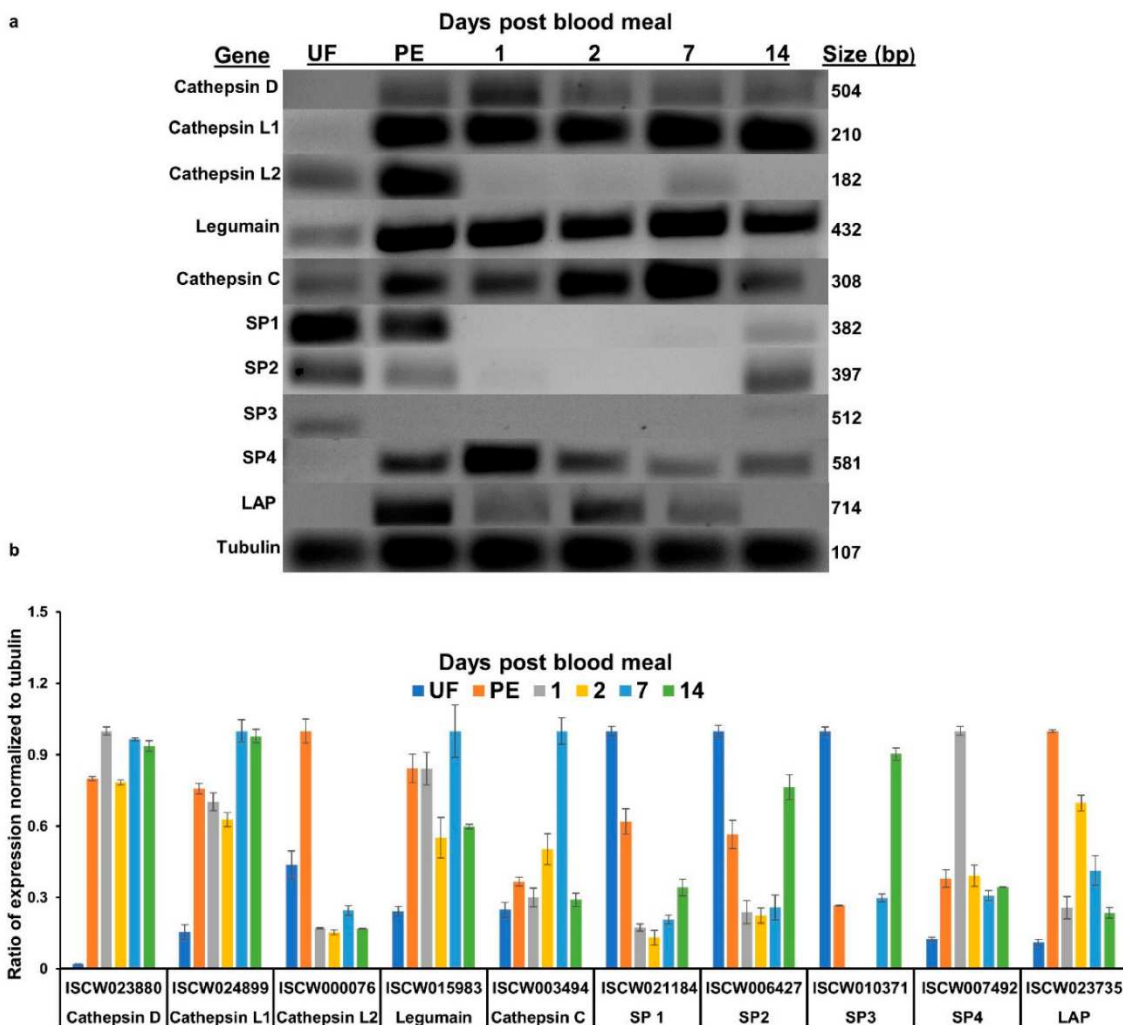


Figure 1. Transcript expression of proteases in the midgut of adult female *Ixodes scapularis*. (a) Representative RT-PCR of ten *I. scapularis* proteases, representing different protease families, at different time points during blood feeding and digestion. Total RNA was extracted from a pool of 2–4 midguts at each time point and an equal amount of cDNA was used for RT-PCR. (b) Densitometry of the RT-PCR gel electrophoresis images to estimate relative transcript abundance scaled to the

housekeeping gene tubulin. UF = unfed female midgut, PE = partially engorged (females were pulled from the host 5 days post-attachment), PBM = post-blood meal (fully engorged females dropped off the host). SP = serine protease; LAP = leucine aminopeptidase.

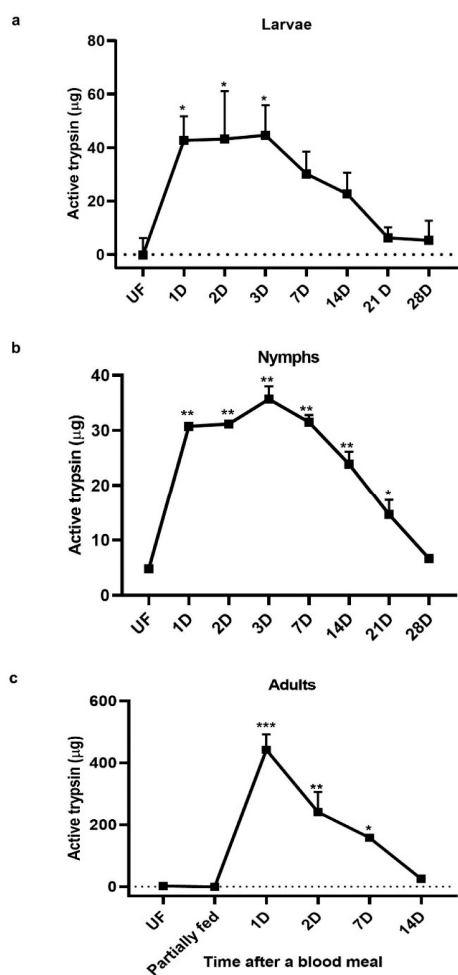


Figure 2. BApNA assay of trypsin activity in *I. scapularis* developmental stages. (a) Larvae and (b) Nymphs were collected before and after feeding on a mouse and whole bodies of larvae or nymphs were homogenized for use in the BApNA assay to estimate trypsin activity. (c) Adult females were collected before, during (5 days post-attachment, partially engorged), or at intervals after feeding on a rabbit. Dissected midguts were

homogenized for use in the BApNA assay. One-way ANOVA and Dunnett's multiple comparisons were used for statistical analysis. * = 0.01; ** = 0.001; *** = 0.0001

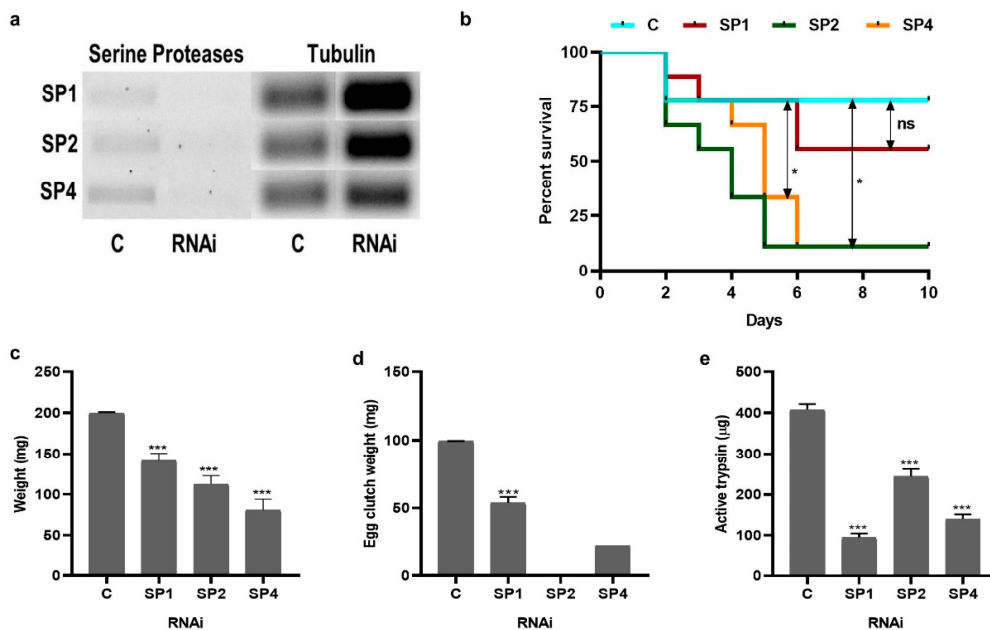


Figure 3. Effect of serine protease knockdown on adult female *I. scapularis* feeding, blood digestion, survival, and reproduction. (a) Representative RT-PCR of three *I. scapularis* serine proteases demonstrating knockdown. (b) Percent mortality during blood-feeding, until all ticks dropped off. (c) Wet weight of ticks measured immediately after dropping off the host. (d) Egg clutch weight, as determined after females stopped laying eggs for 2 consecutive days. (e) Active trypsin levels measured by the BApNA assay in midguts dissected from replete females, 1-day PBM. C: control ticks injected with water, RNAi: dsRNA injected ticks. SP: serine protease, SP1 = ISCW021184, SP2 = ISCW006427, and SP4 = ISCW007492. An unpaired t-test with Welch's correction was used for comparing control and a treatment (SP1, SP2 and SP4) using Graphpad Prism v8. *** p < 0.0001.

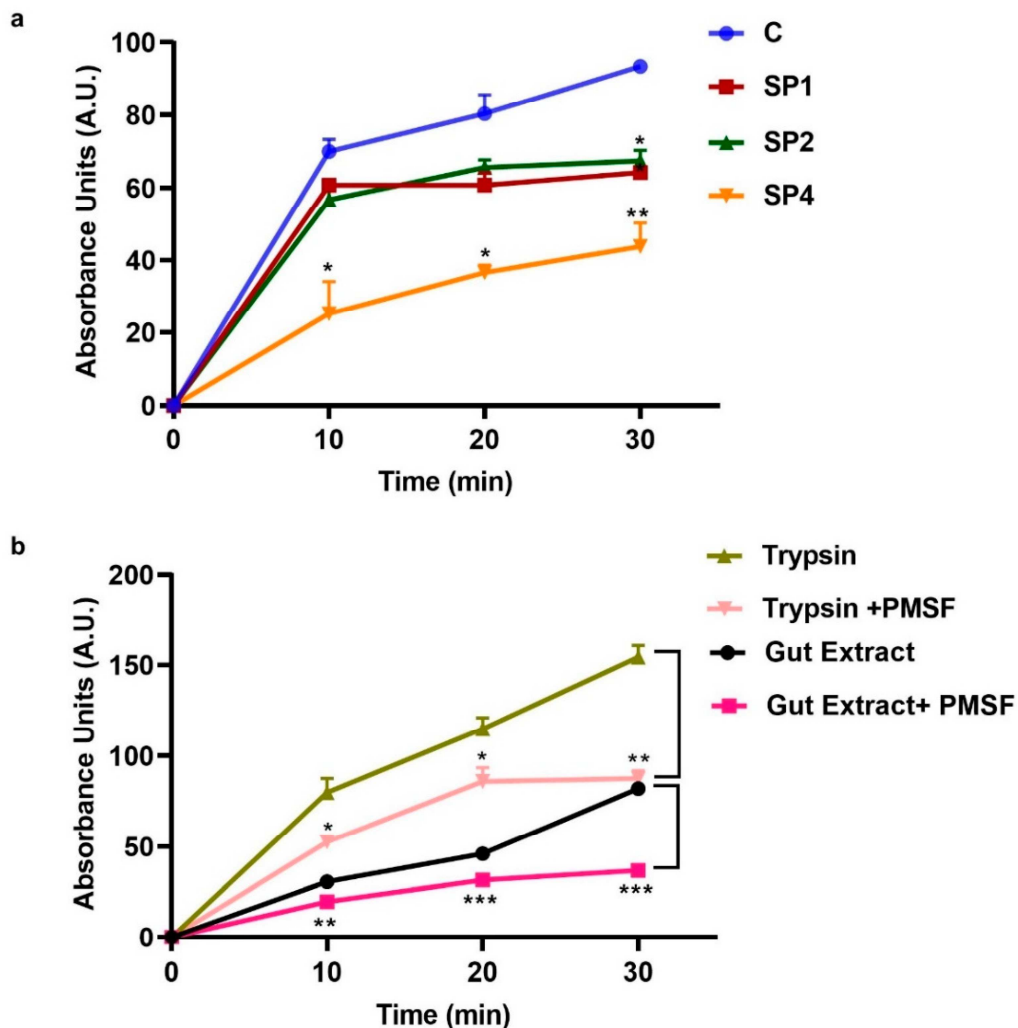


Figure 4. Hemoglobin degradation by *I. scapularis* midgut extract. (a) Effect of serine proteases knockdown on hemoglobin degradation. Homogenized midgut extract from fully engorged (1-day PBM) Control (water-injected) or SP1, SP2, and SP4 knockdown ticks was incubated with bovine hemoglobin and degradation was estimated over time by absorbance of fluorescamine (mean \pm standard deviation, normalized to fluorescence intensity at 0 min). (b) Hemolytic activity of the tick midgut extract was significantly inhibited by pre-incubation with the serine protease-specific inhibitor PMSF. Control

hemolytic reaction with bovine trypsin were similarly inhibited by PMSF. Significance was calculated separately for trypsin and trypsin + inhibitor; and gut extract and gut extract + inhibitor. An unpaired t-test with Welch's correction with 95% confidence interval was used. * $p < 0.05$; ** $p < 0.01$, *** $p < 0.0001$.

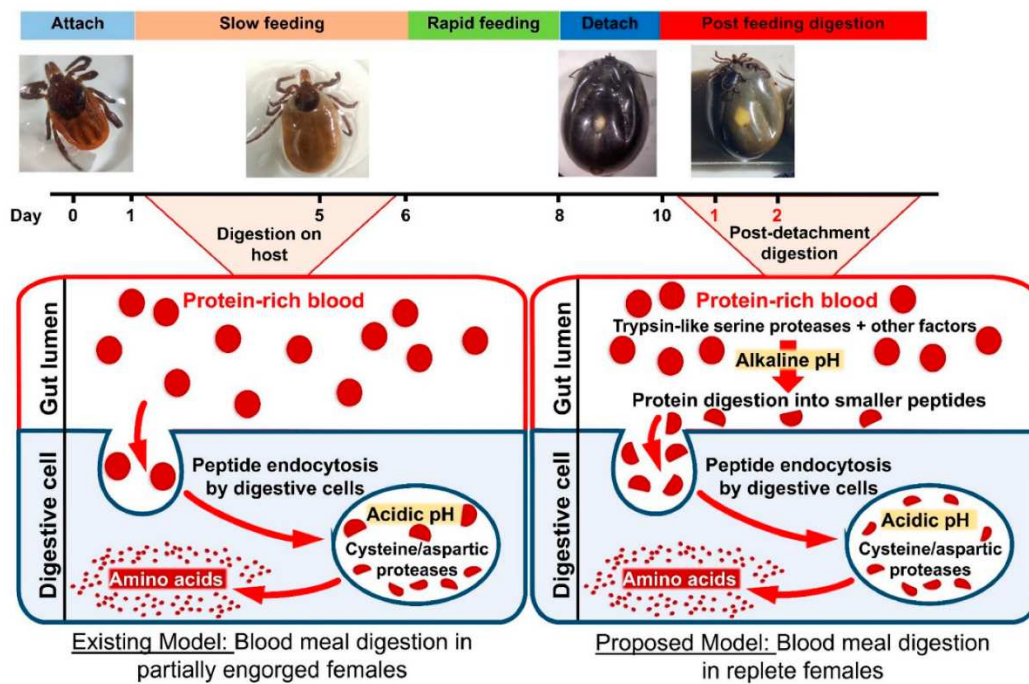


Figure 5. Revised model of blood digestion in replete females. Proposed model suggesting that in replete females (1–2 days PBM), ingested blood proteins are digested by the trypsin-like serine protease enzymes.

Table 1. A list of primer pairs and the annealing temperature used for each primer pair.

Gene Name and Accession Number	Primer Pair Sequence	Annealing Temp °C
Cathepsin D (Aspartic) ISCW023880/ XM_002416473.1		
IscapCathDFwd	CCCTTCCGTGTGGTGTTG	55
IscapCathDRev	AGTAGCCCTTGGTTGAGACAG	55
Cathepsin L (Cysteine) ISCW024899/ XM_002416305.1		
IscapCathLFwd	GACTTCCAGATGTACCAGGGC	55
IscapCathLRev	GAAGGATGCGGAAGTAGCCG	55
Cathepsin L (Cysteine) ISCW000076/ XM_002404428.1		
IscapCathL2Fwd	AAGTGGCCCCACTGCAACTC	55
IscapCathL2Rev	TTACCCGTAACCGCAGGAATG	55
Cathepsin C (Cysteine) ISCW03494/ XM_002400742.1		
IscapCathCFwd	CGTAACTACGTGTCCCCTG	57
IscapCathCRev	TAGTTGCCGACGTAATGCC	57
Legumain (Apartic) ISCW015983/ XM_002402043.1		
IscapLegumainFwd	CCCCTGGAGTGGTCATCAAC	55
IscapLegumainRev	TAAGTGTTTCGGAGGGCGTC	55
Serine Protease 1 ISCW021184/ XM_002405400.1		

Gene Name and Accession	Primer Pair Sequence	Annealing
Number		Temp °C
IscapSP1Fwd	AGCCTAATCAATCAAGGGCG	58
IscapSP1Rev	GACCAGTTTAGGGATGCGAG	58
Serine Protease 2 ISCW006427/ XM_002435219.1		
IscapSP2Fwd	ATCCACGTTGGGAACCTTTC	58
IscapSP2Rev	CAATGGTCAAACGCCTTTC	58
Serine Protease 3 ISCW010371/ XM_002402819.1		
IscapSP3Fwd	TCTACGAGTTCCTGGGACAG	58
IscapSP3Rev	GGACCAGGGAATAATCGTCG	58
Serine Protease 4 ISCW007492/ XM_002404245.1		
IscapSP4Fwd	GCTTCGTCGAAAAAGCTCAC	58
IscapSP4Rev	CAACTCTCGGCGATCTCTTC	58
Leucine Aminopeptidase ISCW023735/ XM_002416067.1		
IscapLAPFwd	ACGCCATTCTCTCACCAAG	55
IscapLAPRev	TTCGGACCCACTGCATTCTC	55
Tubulin ISCW005137/ XM_002402966.1		
IscapB-TubFwd	TGAATGACCTGGTGTCCGAG	55–58
IscapB-TubRev	GCAAAGCTGTTCAAGCCTCT	55–58

Population Genetic structure of the Western black-legged tick, *Ixodes pacificus*, in the state of California

4.1 Abstract

The epidemiology of vector-borne zoonoses depends on the movement of vectors and their hosts. However, little information is available on vector movement patterns for *Ixodes pacificus* ticks that transmit many infectious agents, including the causative agent of Lyme disease. Because of their life history traits and small size, tick dispersal may be frequent but limited in distance. To elucidate the dispersal of ticks and pathogen diversity, we investigated the population structure of *I. pacificus* within the state of California. Ticks were collected from ten counties in north, south, and eastern California. Genomic DNA was isolated and using Restriction-associated DNA sequencing (RADSeq), genome-wide single nucleotide polymorphism (SNP) markers were identified in individual ticks from different populations. Using the SNP markers, we identified two common ancestors. Our data suggest migration routes for these ancestral populations. We also developed a bioinformatics pipeline to identify *Borrelia burgdorferi*, the causative agent of Lyme disease, using genomic data. Our data suggests that tick dispersal is likely frequent and therefore high gene flow occurs between the populations sampled. The genomic data suggested up to one out of ten *Borrelia* infection prevalence at some locations in California, suggesting vector movements may play a key role in pathogen emergence and needs to be considered for predicting pathogen circulation and for establishing effective control strategies.

4.2 Introduction

Throughout the previous decade, genotyping-by-sequencing approaches have increased our understanding of non-model organisms and wild animals (Hohenlohe et al., 2017). Restriction site associated DNA (RAD) sequencing has increased in popularity because it generates a massive single nucleotide polymorphism (SNP) dataset for species which lack genomic resources, and even species which contain genomes as it is more affordable than sequencing the entire genome (Hohenlohe et al., 2012). These data allow us to resolve the genetic differences among many populations, as well as the power of investigating the main drivers of genetic diversity and population structure (Narum et al., 2013).

Population genetics studies are integral for understating the biology of tick vectors by estimating how they disperse and their potential to spread pathogens (Araya-Anchetta., et al, 2014). Investigating tick population structure is critical because, among all arthropods, ticks vector the widest variety of pathogens, leading to a serious public health issue and economic loss in livestock production (Hill and Wikel, 2005, Jongejan and Uilenberg, 2004, Pagel Van Zee et al., 2007, Parola and Raoult, 2001). A large focus has been placed on the effectiveness of tick control for reducing pathogen transmission, but this work cannot be done without the consideration of tick population dynamics and dispersal patterns. While ticks transmit many pathogens, the spirochete *Borrelia burgdorferi*, the causative agent of Lyme disease is most notable. Over 400,000 Lyme disease cases are reported annually in the US alone (CDC, January 13, 2021). In the USA, the Lyme disease spirochete is mainly transmitted by *Ixodes scapularis* in the east and *Ixodes pacificus* in the west. While many studies have investigated populations of *I.*

scapularis (Keirans et al., 1996; McLain et al., 1995, Qiu et al., 2002; Sakamoto et al., 2014; Kelly et al., 2014; Krakowetz et al., 2011; Gulia-Nuss et al., 2016), much less is known of the genetic diversity in *I. pacificus*.

Ixodes pacificus is responsible for the transmission of *B.burgdorferi* in the western United States. *Ixodes pacificus* has a broad host range, which includes over 100 species of reptiles, birds, and mammals (McVicar et al., 2022), increasing the potential for gene flow (Lane et al. 1991). The recent increase in deer range in California and increased frequencies of wildfires have resulted in tick hosts in new locations (Furnas et al., 2020). In a study in the 1990s, low to moderate gene flow was observed in this tick species using eight polymorphic allozymes in ten populations from California (Kain et al., 1997); however, these data might now differ due to increased human activities and other natural factors such as wildfires.

In the present study, 96 individual *I. pacificus* adults collected from ten counties in California were compared genetically using single nucleotide polymorphism (SNP) analyses. Haplotype analysis suggested two different geographical origins and ancestral populations in California, which may be related to the transportation of ticks by deer or by migratory passerines using different flyways. Determination of the origins of the endemic populations of *I. pacificus* in California, as well as those predicted to establish, has important implications for understanding the risk of human exposure to tick-borne pathogens.

4.3 Materials and Methods

Sample Collection

Ixodes pacificus adults were collected from 10 different counties: Mendocino, Nevada, Placer, El Dorado, Sacramento, Alameda, San Mateo, Santa Clara, Santa Cruz, and Santa Barbara in the state of California (Fig. 1). Ticks were either collected by our research team, our collaborators at the public health and vector control districts of California or mailed by the local citizens. The geographic positioning coordinates were recorded for each field-collected sample and this information was also requested from citizen-collected specimens. Soon after collection, ticks were surface cleaned with 70% ethanol and stored at -80°C until further processing.

DNA Extraction

Genomic DNA was extracted from individual ticks using the protocol by Gulianuss et al. (2016). Ticks were mechanically homogenized in 1.5 mL tube with extraction buffer (500 mM Tris, pH 9.0, 20 mM NaCl), 20% SDS, and proteinase K and incubated for 6-12 hrs at 55 °C. After incubation, samples were centrifuged to remove debris, RNase A (20 mg/mL concentration) was added and incubated at room temperature for 20 min followed by incubation at 70 °C for 10 min Phenol:chloroform:isoamyl alcohol (Invitrogen) was added to digested tissue, vortexed vigorously, and centrifuged at 13.3K x g for 20 min. The upper aqueous phase was then collected and transferred to a new tube before adding chloroform only (Thermo Fisher), vortexed vigorously, then centrifuged. The aqueous phase was collected again, transferred to a new tube, and 3 M sodium acetate (pH 5.2) was added. The sample was then precipitated overnight with 95% ethanol, centrifuged and the ethanol was decanted. Pellet (DNA) was washed with 70% ethanol, air dried and resuspended in Low Tris-EDTA (Low TE) buffer and incubated at

55 °C for 10 min. To assess the concentration and ensure the quality of DNA, an aliquot was electrophoresed on 1.2% agarose gel.

RADSeq library preparation

The published protocol of Ali et al., (2016) was used for RADSeq library preparation. 100 ng genomic DNA (gDNA) per tick sample was added to the well of a 96-well plate. The volume was adjusted to 10 uL using Low Tris-EDTA (TE) buffer. 2 uL of SbfI enzyme digestion master mix (8 ul H₂O, 120 ul NEB buffer, and 12 ul SbfI HF restriction enzyme) was added to each well, and digested at 37°C for 60 min, followed by 80°C for 20 min. 2 ul of the BestRAD adaptor (Table 1) with a different barcode for each sample were added to each well and 2uL ligation master mix (128 uL water, 40 uL NEBuffer 4, 16 uL 100 mM rATP (Fermentas R0441), and 16 uL DNA Ligase (NEB M0202L) was added to anneal the adaptors to the restricted DNA. Sample was incubated at 20 °C for 12 h and 65 °C for 20 min. master mix in 1.5mL tube. Out of 16 ul total volume in each well, 8 uL was taken out in a 1.5 ml tubes and pooled with other samples (each individual sample now has an adapter ligated to identify later). The pooled samples were sonicated with a probe sonic dismembrator model 100 (Thermo Fisher) at level 3 (8 cycles: 30 sec on, 90sec off) while keeping the sample on ice. An aliquot of the sonicated sample was electrophoresed on a 1.2% agarose gel to visualize 250-500bp fragments. Streptavidin magnetic Dynabead (Invitrogen 11205D) beads were used to size select and remove DNA without adapters. Then NEBNext Ultra DNA Library prep kit for Illumina (NEBE7970S/L) was followed with no modifications, the final elution was 22 uL In Low TE. The library was amplified using 10-12 PCR cycles and sent for Illumina sequencing to the Genomics Core at UC Davis. The library was

split into two equal volumes and run on two lanes of Illumina HiSeq 2500 in single end mode.

Identification of Single Nucleotide Polymorphisms (SNPs)

Raw Illumina single-end reads were first quality assessed using FastQC (Andrews, 2010) and MultiQC (Ewels et al., 2016). Low-quality bases were trimmed using trimmomatic version 0.39. Illumina reads were corrected for barcodes and restriction site and demultiplexed using the “process_radtags.pl” script of STACKS (version 2.55) with the code (process_radtags -i gzfastq --adapter-1 CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC -f /path/to/RAW_data/*.fastq.gz -o sample_output/ -b path/to/barcodes.txt -c -r --bestrad -q --barcode_dist_1 2 -e sbf). Quality trimmed reads were used to assemble a *de novo* reference file. Assembly from all reads was conducted using the Ustacks, Cstacks, and Sstacks pipelines (Catchen et al., 2013). Tsv2Bam was used to convert files into binaries, along with samtools (Li et al., 2009). Gstacks were used to identify SNPs across loci.

F-statistics and Phylogenetic Analysis

Populations program from STACKS was run in two manners: 1) With a single SNP per loci for F-statistics (Gulia-Nuss et al., 2016) and 2) All SNPs in loci for phylogenetic analysis and migration patterns analysis. Criteria common for both F-statistics and phylogenetic analysis in Populations program within STACKS package were: a minimum of 60% individuals within a population, a minimum of two populations to report a locus, and a minimum stack depth of 10 per locus.

Migration Routes Analysis

SNPs from Populations program were used to output TreeMix file (Pickrell and Pritchard, 2012). TreeMix file contains SNPs from all counties and was used to estimate the historical relationships among populations. TreeMix automatically subsamples one SNP per locus to reduce the effect of linkage on the results. This is followed by the inference on the number of admixture events to identify migration routes through the allele frequency of haplotypes.

Pathogen detection from RADseq data

Sequences from our RADseq data that did not align with our *de novo* reference file were used to align with pathogen genomes to identify tick-borne pathogens. Trimmomatic was used to remove barcodes from demultiplexed individuals. HISAT2 was implemented to align all samples to the *B. burgdorferi* CA11 genome, the most common strain found in the Western United states (Schutzer et al., 2011).

Samples which contained signatures of *B. burgdorferi* in RADseq samples were then verified with PCR as per Xu et al. (2019) with *B. burgdorferi* outer surface protein A (ospA) primers (Forward – ATAGGTCTAATATTAGCCTTAATAGCAT, Reverse – AGATCGTACTTGCCGTCTT) for a product size of 140bp. Briefly, 1 μ L DNA samples were used as a template in a 10 μ L reaction. RT-PCR conditions for samples were: initial denaturation at 98 °C for 5 min, 98 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s, repeated for 35 cycles, and a final extension at 72 °C for 5 min. 10 μ L PCR product was separated by electrophoresis on a 1.2% agarose gel along with a DNA ladder (Apex DNA Ladder I; Genesee, San Diego, CA, USA) and visualized by using ethidium bromide-free dye (Amresco, Solon, OH, USA).

4.4 Results

The RADseq technique was used for genome-wide SNP discovery and examination of genetic diversity within and among 10 *I. pacificus* populations from the northern coast, central coast, southern coast, and along the Eastern Sierra in California. F-statistics were used to assess genetic distance.

Read Quality analysis

Two Illumina lanes provided 94.45 and 97.76 million reads. A total of 38,101 loci, 781,571 SNPs, and a total of 34,404 variant sites were identified (Table 1). A total of 18,490 private alleles were identified in total (Table 2). Unique private alleles per county ranged from 219 (Santa Barbara) to 2,937 (El Dorado County) (Table 2).

Assessment of genetic variation inter and intra-counties

F-statistics were employed from the Populations program within the STACKS pipeline to assess genetic distances. In population genetics, F-statistics (also known as fixation indices) describe the statistically expected level of heterozygosity in a population; specifically, the expected degree of reduction in heterozygosity when compared to Hardy–Weinberg expectation. The F_{IS} is mean deficiency of observed heterozygotes among individuals with respect to the sub-populations (individuals within a population; here within a county) whereas F_{ST} is mean deficiency of expected heterozygotes among sub-populations with respect to that expected for the total population, which in this case is a measure of population differentiation in populations from all 10 counties.

F_{IS} values ranged from 0.00076 (Alameda) to 0.00162 (El Dorado) (Table 3).

Nucleotide diversity in all SNP loci (π) was assessed at variant and all positions within

each county. For variant SNPs unique to each county, the lowest nucleotide diversity was 0.073 from Santa Barbara, and the highest was 0.1080 from El Dorado. Considering all SNPs identified within a county, the range was from 0.00051-0.00115 with Santa Barbara being the lowest and El Dorado having the most diversity (Table 4).

F_{ST} values were assessed to understand the level of differentiation between all counties. F_{ST} ranged from 0.0415-0.0657, maintaining within the range of low ($F_{ST}<0.05$) to moderate ($F_{ST} = 0.05-0.15$) genetic variation. No subpopulations had high or very high genetic variation. The lowest genetic variation was observed among either the counties near the bay area (Alameda, Santa Clara, San Mateo, and Santa Cruz) or closer to the Sierras (Nevada, Sacramento, Eldorado, Placer). The highest variation was observed between counties that are geographically distanced (Santa Barbara and Mendocino) (Table 5).

The population structure of *I. pacificus* was analyzed to identify the relationship of these counties with haplotype data from the genome-wide set of SNPs and phylogenetic tree was plotted with TreeMix. We identified two separate ancestral clades based on 871,000 haplotypes. Ticks from Santa Clara, Santa Cruz, San Mateo, Mendocino, and Alameda formed one clade, whereas ticks from Nevada, Placer, El Dorado, Sacramento, and Santa Barbara grouped in the second clade (Figure 2).

Migration routes were also assessed using the allelic frequency in our dataset to identify movements of tick populations. TreeMix uses allele frequencies from a number of populations and returns the maximum likelihood tree for the set of populations and attempt to infer a number of admixture events. Admixture proportions are given a weight of 0-0.5, 0.5 meaning 50% of population has the alleles similar to ancestral population. .

Genome-wide allelic frequency data suggests one major route of migration or gene flow from Alameda County south towards Santa Clara (migration weight of 0.5), and two minor routes (migration weight of 0.3), from Sacramento County towards El Dorado County, and from Placer County towards the Northern coast in Mendocino County (Figure 3).

***Borrelia burgdorferi* identification in RADseq samples**

Using the HISAT2 alignment pipeline, we found two ticks (one each from El Dorado and Santa Clara counties) out of our 96 individuals (~2% total and 10% per county) that aligned to the *B. burgdorferi* genome (Table 6). The El Dorado positive sample aligned 17 times to different regions of *B. burgdorferi* genome; whereas the Santa Clara sample aligned 8 times. These samples were verified using *B. burgdorferi* ospA specific primers with RT-PCR and both were PCR positive. (Figure 4).

4.5 Discussion

This work represents the first genome-wide SNP analysis of *I. pacificus* populations from ten counties in California (Figure 1). Our data agrees with the previous work by Kain et al. (1997) where they predicted two groups/clades in California: A northern group of populations (Alameda co, CA) and a southern group (Mendocino, Santa Clara, and San Diego counties) (Kain et al., 1997). However, our data differs in the prediction of these groups. Alameda, Mendocino, and Santa Clara populations were related to one ancestral group whereas the second group was formed by the populations from Nevada, Placer, El Dorado, Sacramento, and Santa Barbara. Kain et al. (1997) suggested that either a rapid range expansion or high rates of gene flow could explain the pattern of allozyme variation. However, they proposed that the most plausible explanation for the pattern of

allozyme variation in *I. pacificus* was a high rate of gene flow. Similar to the previous study, our data also suggests a high gene flow based on low to moderate genetic variation in counties and low variation within a county. Dispersal or migration is an integral part of population genetics and evolution because the extent of gene flow influences a range of variables including genetic variation. Dispersal and gene flow are particularly important to populations that are prone to localized extinctions and recolonizations such as ticks. Our F_{IS} data suggests high gene flow and random mating in tick populations because of low fixation index within populations (close to zero) suggesting that tick populations are able to move around. Nucleotide diversity is used to measure the degree of polymorphism within a population (Nei et al., 1979). This measure is defined as the average number of nucleotide differences per site between two DNA sequences in all possible pairs in the sample population, and is denoted by π . In our data, nucleotide diversity was very low possibly either due to a relatively small long-term effective population size or severe bottleneck during evolution, most likely due to a relatively small long-term effective population size rather than any severe bottleneck. Low nucleotide diversity was also noted in population structure of elephant tick in Kenya (King'ori et al., 2022) and the brown dog tick in Colombia (Páez-Triana et al., 2021). Habitat fragmentation results in decreased habitat patch size and consequences of this common phenomenon include, amongst others, modified community composition and structure, restrictive movement, and smaller population sizes (Fahrig, 2003). Both population size and movement significantly affect population genetic diversity. As population size decreases, genetic stochasticity increases, resulting in increased allele fixation with each generation due to higher genetic drift. Anthropogenic activities potentially lead to fragmentation of tick

habitats and given the long life cycle of ticks (2-5 years), the low nucleotide diversity suggests recolonization events by migration.

Phylogenetics are important to understand the distribution and movement of individuals within species. We identified two distinct ancestral clades using haplotype data (Figure 2). The Northern clade included the northern coastal counties near the San Francisco Bay Area including the northernmost in our RADseq library, Mendocino County. The second clade comprised of populations from the Sierra Mountains and our southernmost county, Santa Barbara (Figure 2). A possible factor limiting gene flow could be the geographic barriers in California. Much of the Central Valley reaches extreme temperatures in the south, constant drought in the north, and has minimal routes of travel for large mammals as it is heavily used for agriculture. The possible routes for animal migration is via the northern corridor of the state (Faunt et al., 2016). We determined one major and two minor migration routes based on allelic frequency using TreeMix.. Our data suggests strong migration from Alameda County towards Santa Clara (a migration weight of 0.5); this may be attributed to the lack of geographical barriers, as well as an increased in the deer population throughout California in recent years (Brazeal et al., 2017, Furnas et al., 2020). Our data also suggest a minor migration from Eastern California (Sierra) to Northwest (Mendocino). As the deer herds move from Sierras towards the northern coastal regions (Mendocino) (CDFW Proposal, 2016), they might carry ticks with them and therefore, we see a minor migration pattern in that direction with a weight of 0.3 (Figure 3).

Finally, RADseq data was used to identify the presence of *Borrelia burgdorferi* in our samples. We identified one infected tick each from 10 total individuals from two

counties (El Dorado and Santa Clara) resulting in 10% in *B. burgdorferi* positivity rate for each of these counties. While most studies suggest ~2% *B. burgdorferi* prevalence in ticks in California (McVicar et al., 2022); the California public health department suggests up to 15% positivity rate in some focal areas. More work is needed to understand the *I. pacificus* populations in California to close the gaps. Larger sample sizes are required in order to get a better assessment of pathogen prevalence. These data will help bring together the multiple factors involved in developing management plans to control the expansion of this medically important tick species.

4.6 References

- Ali, Omar A., Sean M. O'Rourke, Stephen J. Amish, Mariah H. Meek, Gordon Luikart, Carson Jeffres, and Michael R. Miller. RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics* 202, no. 2 (2016): 389-400.
- Andrews, Simon. FastQC: a quality control tool for high throughput sequence data. (2010).
- Araya-Anchetta, Ana, Joseph D. Busch, Glen A. Scoles, and David M. Wagner. Thirty years of tick population genetics: a comprehensive review. *Infection, Genetics and Evolution* 29 (2015): 164-179.
- Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, no. 15 (2014): 2114-2120.

Brazeal, Jennifer L., Terri Weist, and Benjamin N. Sacks. Noninvasive genetic spatial capture-recapture for estimating deer population abundance. *The Journal of Wildlife Management* 81, no. 4 (2017): 629-640.

California Department of Fish and Wildlife (CDFW). In estimation of abundance of multiple deer herds using fecal DNA north central region. *California Department of Fish and Wildlife* proposal, Large Mammal Advisory Committee, Fiscal year 2015-2016

Catchen, Julian, Paul A. Hohenlohe, Susan Bassham, Angel Amores, and William A. Cresko. Stacks: an analysis tool set for population genomics. *Molecular ecology* 22, no. 11 (2013): 3124-3140.

Ewels, Philip, Måns Magnusson, Sverker Lundin, and Max Käller. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, no. 19 (2016): 3047-3048.

Faunt, Claudia C., Michelle Sneed, Jon Traum, and Justin T. Brandt. Water availability and land subsidence in the Central Valley, California, USA. *Hydrogeology Journal* 24, no. 3 (2016): 675-684.

Fahrig L. Effects of habitat fragmentation on biodiversity. *Ann Rev Ecol Evol Syst.* 2003;34:487–515. doi: 10.1146/annurev.ecolsys.34.011802.132419

Frankham R. Genetics and extinction. *Biol Conserv.* 2005;126:131–140. doi: 10.1016/j.biocon.2005.05.002.

Furnas, Brett J., Russ H. Landers, Rhonda G. Paiste, and Benjamin N. Sacks.

Overabundance of black-tailed deer in urbanized coastal California. *The Journal of Wildlife Management* 84, no. 5 (2020): 979-988.

Gulia-Nuss, Monika, Andrew B. Nuss, Jason M. Meyer, Daniel E. Sonenshine, R.

Michael Roe, Robert M. Waterhouse, David B. Sattelle et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nature communications* 7, no. 1 (2016): 1-13.

Hill, Catherine A., and Stephen K. Wikel. The *Ixodes scapularis* Genome Project: an opportunity for advancing tick research. *Trends in parasitology* 21, no. 4 (2005): 151-153.

Hohenlohe, Paul A., Julian Catchen, and William A. Cresko. Population genomic analysis of model and nonmodel organisms using sequenced RAD tags. In *Data production and analysis in population genomics*, pp. 235-260. Humana Press, Totowa, NJ, 2012.

Jongejan, Frans, and G. Uilenberg. The global importance of ticks. *Parasitology* 129, no. S1 (2004): S3-S14.

Kain, Douglas E., Felix AH Sperling, and Robert S. Lane. Population genetic structure of *Ixodes pacificus* (Acari: Ixodidae) using allozymes. *Journal of medical entomology* 34, no. 4 (1997): 441-450.

Kayondo, Hassan W., Alfred Ssekagiri, Grace Nabakooza, Nicholas Bbosa, Deogratius Ssemwanga, Pontiano Kaleebu, Samuel Mwalili et al. Employing phylogenetic tree shape statistics to resolve the underlying host population structure. *BMC bioinformatics* 22, no. 1 (2021): 1-20.

Keirans, James E., H. Joel Hutcheson, Lance A. Durden, and J. S. H. Klompen. *Ixodes (Ixodes) scapularis* (Acari: Ixodidae): redescription of all active stages, distribution, hosts, geographical variation, and medical and veterinary importance. *Journal of Medical Entomology* 33, no. 3 (1996): 297-318.

Kelly, Rebecca R., David Gaines, Will F. Gilliam, and R. Jory Brinkerhoff. Population genetic structure of the Lyme disease vector *Ixodes scapularis* at an apparent spatial expansion front. *Infection, Genetics and Evolution* 27 (2014): 543-550.

Kim, Daehwan, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology* 37, no. 8 (2019): 907-915.

King'ori EM, Obanda V, Nyamota R, Remesar S, Chiyo PI, Soriguer R, Morrondo P. Population genetic structure of the elephant tick *Amblyomma tholloni* from different elephant populations in Kenya. *Ticks Tick Borne Dis.* 2022 May;13(3):101935. doi: 10.1016/j.ttbdis.2022.101935. Epub 2022 Mar 6. PMID: 35325688.

Krakowetz, Chantel N., L. Robbin Lindsay, and Neil B. Chilton. Genetic diversity in *Ixodes scapularis* (Acari: Ixodidae) from six established populations in Canada. *Ticks and tick-borne diseases* 2, no. 3 (2011): 143-150.

Lane, R. S., J. Piesman, and W. Burgdorfer. Lyme borreliosis: relation of its causative agent to its vectors and hosts in North America and Europe. *Annual review of entomology* 36, no. 1 (1991): 587-609.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. The sequence alignment/map format and SAMtools. *Bioinformatics* 25, no. 16 (2009): 2078-2079.

McLain, Denson Kelly, Dawn M. Wesson, James H. Oliver Jr, and Frank H. Collins. Variation in ribosomal DNA internal transcribed spacers 1 among eastern populations of *Ixodes scapularis* (Acari: Ixodidae). *Journal of Medical Entomology* 32, no. 3 (1995): 353-360.

McVicar, Molly, Isabella Rivera, Jeremiah B. Reyes, and Monika Gulia-Nuss. Ecology of *Ixodes pacificus* Ticks and Associated Pathogens in the Western United States. *Pathogens* 11, no. 1 (2022): 89.

Narum, Shawn R., C. Alex Buerkle, John W. Davey, Michael R. Miller, and Paul A. Hohenlohe. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular ecology* 22, no. 11 (2013): 2841.

Nei, Masatoshi, and Wen-Hsiung Li. Mathematical model for studying genetic variation In terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76, no. 10 (1979): 5269-5273.

Páez-Triana, L., Muñoz, M., Herrera, G. et al. Genetic diversity and population structure of *Rhipicephalus sanguineus sensu lato* across different regions of Colombia. *Parasites Vectors* 14, 424 (2021). <https://doi.org/10.1186/s13071-021-04898-w>

Palamara, Pier Francesco, and Itsik Pe'er. Inference of historical migration rates via haplotype sharing. *Bioinformatics* 29, no. 13 (2013): i180-i188.

Parola, Philippe, and Didier Raoult. Ticks and tickborne bacterial diseases in humans: an emerging infectious threat. *Clinical infectious diseases* 32, no. 6 (2001): 897-928.

Pickrell, Joseph, and Jonathan Pritchard. Inference of population splits and mixtures from genome-wide allele frequency data. *Nature Precedings* (2012): 1-1.

Qiu, Wei-Gang, Daniel E. Dykhuizen, Michael S. Acosta, and Benjamin J. Luft. Geographic uniformity of the Lyme disease spirochete (*Borrelia burgdorferi*) and its shared history with tick vector (*Ixodes scapularis*) in the northeastern United States. *Genetics* 160, no. 3 (2002): 833-849.

Sakamoto, Joyce M., Jerome Goddard, and Jason L. Rasgon. Population and demographic structure of *Ixodes scapularis* Say in the eastern United States. *PLoS One* 9, no. 7 (2014): e101389.

Van Zee, J. Pagel, N. S. Geraci, F. D. Guerrero, S. K. Wikel, J. J. Stuart, V. M. Nene, and C. A. Hill. Tick genomics: the Ixodes genome project and beyond. *International journal for parasitology* 37, no. 12 (2007): 1297-1305.

Williams, Scott C., Kirby C. Stafford III, Goudarz Molaei, and Megan A. Linske. Integrated control of nymphal *Ixodes scapularis*: Effectiveness of white-tailed deer reduction, the entomopathogenic fungus *Metarhizium anisopliae*, and fipronil-based rodent bait boxes. *Vector-Borne and Zoonotic Diseases* 18, no. 1 (2018): 55-64

4.7 Tables and Figures



Figure 1. California map of *I. pacificus* collections. Ticks were collected or received for RADseq processing from counties highlighted in green.

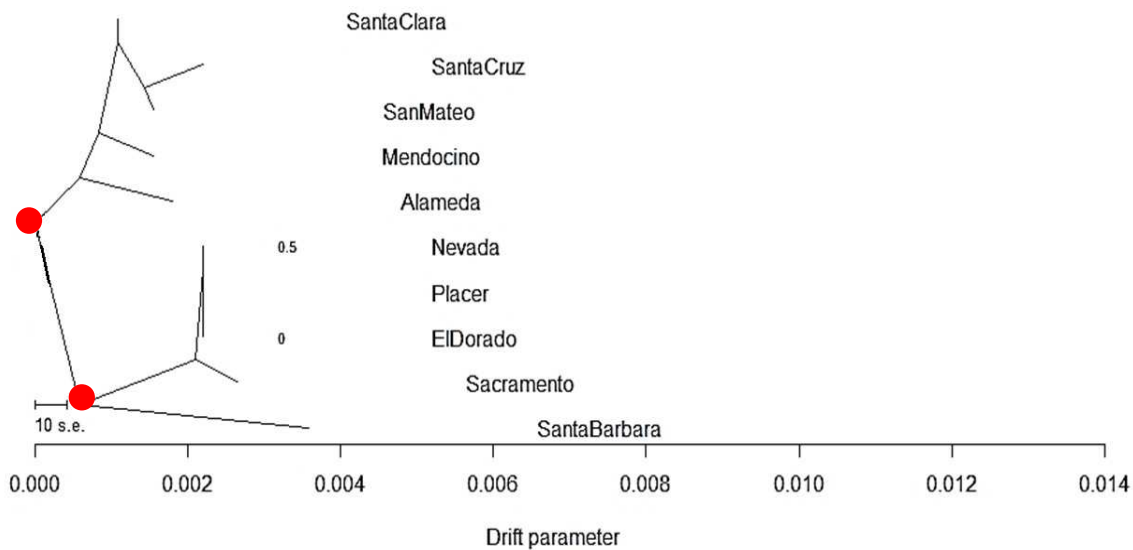


Figure 2. Phylogenetic analysis from haplotype data suggests two common ancestors. Haplotypes from genome-wide data give insight into two ancestral populations (red dots). These separate into a northern and southern clade.

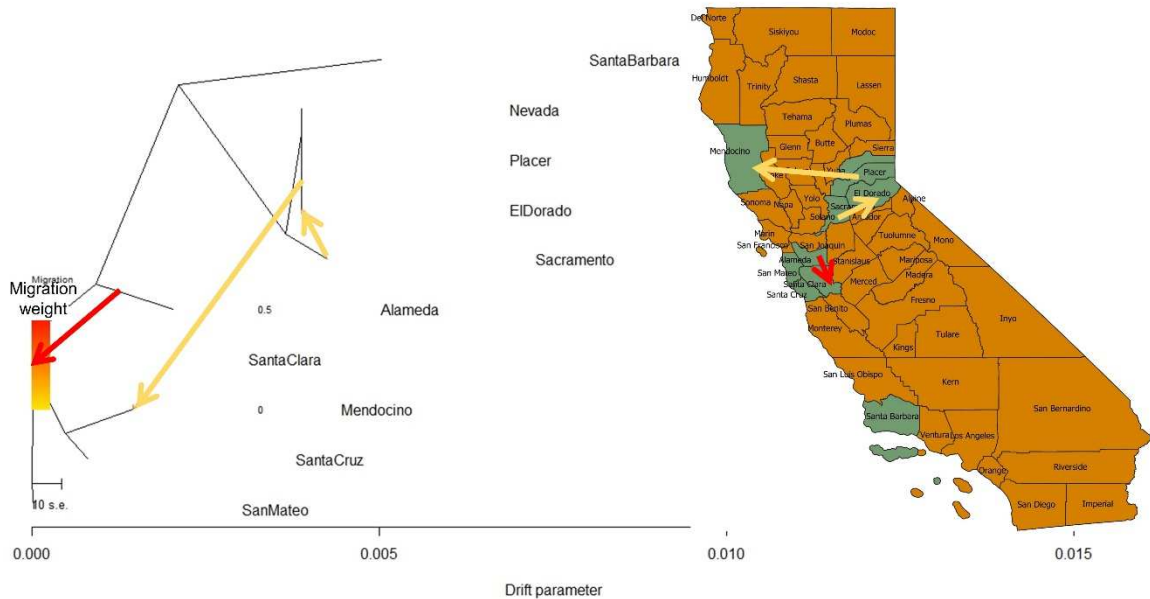


Figure 3. Genome-wide allelic frequency data suggests one major and two moderate migration patterns. Major (red) and moderate (yellow) tick migration analyzed from allelic frequency. Migration of tick populations from western Sierra's to the northern coast could be attributed to mule deer expansion.

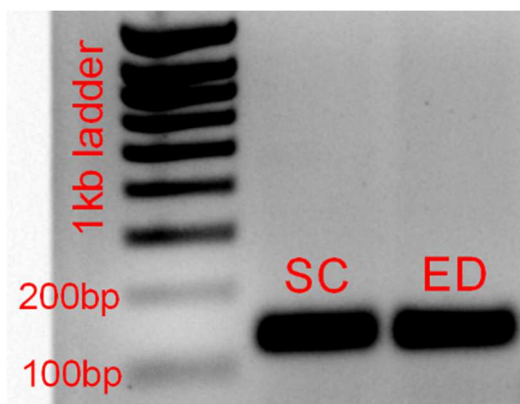


Figure 4. PCR product of DNA samples positive for *Borrelia burgdorferi*. One sample each from Santa Clara (SC) and El Dorado (ED) counties showed the predicted band size of 140bp.

Table 1. Summary of statistics across samples. Raw reads received back from Illumina sequencing platform.

Sample	IPAC_L1_S1	IPAC_L2_S2
Unique Reads	94,448,868	97,763,834
Phred Quality Score	>30	>30
Loci	38,101	
Total SNPs	781,571	
Variant Sites	34,404	

Table 2. Identification of private alleles per county. The number of alleles unique to each county were identified from genome-wide SNP data.

County	Private Alleles
Alameda	660
Santa Clara	2434
Placer	2288
El Dorado	2937
San Mateo	2257
Santa Barbara	219
Nevada	2567
Sacramento	1929
Santa Cruz	1578
Mendocino	1621
Total	18,490

Table 3. Genetic variation among populations of *Ixodes pacificus* collected from different counties in the state of California. F_{st} values are shown as a measure of differentiation between all subpopulations (counties). $F_{st} \leq 0.05$, low genetic variation (tan shading); $F_{st} = 0.05-0.15$ moderate genetic variation (green shading).

	Alameda	Santa Clara	Placer	El Dorado	San Mateo	Santa Barbara	Nevada	Sacramento	Santa Cruz	Mendocino
Alameda		0.0442	0.0548	0.0539	0.0472	0.0639	0.0533	0.0560	0.0527	0.0633
Santa Clara			0.0717	0.0567	0.0471	0.0558	0.0540	0.0557	0.0518	0.0597
Placer				0.0432	0.0579	0.0545	0.0415	0.0454	0.0615	0.0657
El Dorado					0.0598	0.0516	0.0427	0.0445	0.0609	0.0662
San Mateo						0.0564	0.0564	0.0587	0.0475	0.0593
Santa Barbara							0.0518	0.0564	0.0603	0.0632
Nevada								0.0448	0.0569	0.0637
Sacramento									0.0607	0.0672
Santa Cruz										0.0602

Low genetic variation (<0.05)
 Moderate variation (0.05-0.15)
 High variation (0.15-0.25)
 Very High variation (>0.25)

Table 4. Genetic variation between each individual subpopulation of *Ixodes pacificus* from counties in California. F_{is} values were computed to identify genetic variation between individuals (n) collected from each county. A negative F_{is} value = outbreeding and positive F_{is} value = inbreeding.

County	n	F_{is}
Alameda	10	0.00076
Santa Clara	10	0.00133
Placer	10	0.00119
El Dorado	10	0.00162
San Mateo	10	0.00133
Santa Barbara	10	0.00043
Nevada	10	0.00148
Sacramento	10	0.00143
Santa Cruz	10	0.00108
Mendocino	6	0.00091

Table 5. Nucleotide diversity at variant and all positions within each population.

Computed degree of polymorphism with within each subpopulation (county).

County	Variant positions	All positions (variant and fixed)
Alameda	0.08292	0.00087
Santa Clara	0.1035	0.00105
Placer	0.10034	0.00101
El Dorado	0.10803	0.00115
San Mateo	0.10422	0.0011
Santa Barbara	0.07323	0.00051
Nevada	0.10446	0.00112
Sacramento	0.10036	0.00107
Santa Cruz	0.09352	0.00098
Mendocino	0.10477	0.00105

Table 6. Identification of samples positive for *Borrelia burgdorferi*. Identification of the total number of positive individuals for *B. burgdorferi in silico*, verified with PCR.

County	Total number of Individuals	<i>Borrelia burgdorferi</i> positive
El Dorado	10	1, 10%
Santa Clara	10	1, 10%

Conclusions and future directions

Future Directions

With a more complete picture of the blood ingestion and digestion profile in *Ixodes scapularis*, we have identified possible targets to develop recombinant proteins. Currently, we wish pursue gene targets from our RNAi data that did not produce any eggs or fell off early from host for developing an anti-tick vaccine. We would also like to pursue other protein families identified in this work and conduct RNAi studies to identify other possible candidates for tick control. Also understanding the effect of knockout or knockdown on some of these identified targets on pathogen proliferation would help us better understand key proteins important in the tick-host-pathogen interface.

In part-2, identification of gene loci associated with SNPs may help understand vector competency of ticks from different regions of California. This project has developed into an even bigger multi-state project which will allow us to get better insight into the spatial distribution, host prevalence, and allelic specificity by geographic location throughout the Western United States.

Concluding Remarks

With the completion of the projects described herein, we have presented new targets which may be used for anti-tick vaccine for *Ixodes* species. We have also shown through our omics data that serine proteases are the most abundant proteases in tick gut after a full meal. We further characterized three female specific and blood meal induced serine proteases and showed their ability to breakdown haemoglobin and loss of this activity in gut when these genes were knockdown, confirming their role in blood digestion. Our population genomics data provides the first structure analysis using genome-wide SNP

analysis to learn more about this tick species of medical importance on the west coast, which will undoubtedly lead to further discoveries in these realms of study.