CHARACTERIZATION OF NOVEL ANTICOAGULANTS FROM HEMATOPHAGOUS ARTHROPODS

TAN WEI LING, ANGELINA

(B.Sc. (Hons.), National University of Singapore)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

NUS GRADUATE SCHOOL FOR INTEGRATIVE SCIENCES AND ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

_____

Tan Wei Ling, Angelina

8th September 2014

ACKNOWLEDGMENTS

I would like to thank my PhD supervisor, Prof. R. Manjunatha Kini, for his support during my entire PhD candidature. He was very enthusiastic and welcomed me into his lab when I first approached him in my first year. Throughout the years, he has always taught us not be limited by boundaries and walls, and to break the wall when it gets in our way! His pep talks and stories were a constant motivation during this whole time.

I would like to express my gratitude to our collaborators from the Slovak Academy of Sciences, Bratislava - RNDr. Mirko Slovak and RNDr. Maria Kazimirova, who has been the source of ticks for all these experiments. Thank you for your hard work and efforts and always striving to meet our endless requests for the ticks. Thank you for your hospitality when I was at your institute and showing me how tick collection is performed in the wild. It was indeed an eye-opening hands-on trip. Another collaborator of ours - Dr. Jose Ribeiro, thank you for taking time out to personally teach me how raw reads are processed and all the bioinformatics involved. Thank you for patiently answering my endless questions regarding the tables, and all the calculations.

I would also like to thank members of Protein Science Laboratory who have worked with me and have brought many joys, whether during benchwork, lab meetings, nonsense in the student room or chai pani session etc. Special thanks to Cho Yeow, my senior, who got me acquainted with the world of ticks and blood coagulation. Buddies Shiyang, Amrita, Girish and Sheena, you guys never fail to brighten my day in the lab, and I am glad for the strong friendship

forged after all these years together. To Sindhuja and Janaki, thank you for supporting me by providing a listening ear whenever I needed to “release tension”. I will always remember all the silly things we have done outside of the lab. Lastly, thanks to all else who worked along me in the lab and made the lab a pleasant environment to work in: Ryan, Bhaskar, Bidhan, Summer, Angie, Chen Wan, Ritu, Norrapat, Feng Jian, Varuna, and all the honours students who have spent time with us. To Bee Ling and Liyuan, our lab officers, thank you for supporting our lab, and always helping us out with a big smile on your faces. Thank you for helping us process our orders timely without fail, keeping the lab orderly and sane, and always entertaining all sorts of nonsense from us.

I would like to thank my PhD buddies – Veronica and Carol. I would never have lasted through without both of you guys. Thank you for your constant support, encouragement, feedbacks and suggestions. Thank you also for applying the right pressure on me, not too much but not too little as well, and always at the right time.

I would also like to thank NGS for awarding me with the PhD scholarship, allowing me this opportunity to further my studies in NUS. Thank you for providing the funding for overseas attachments. A big thank you to Irene who has always cheered me on whenever I showed up at the NGS office.

A special thanks to my family members, who have always supported the decisions I have made in life. This whole period wasn’t easy for me, Mom and Dad, thank you for cutting me some slack when I was not having the best of days. To my Chieh, thank you for your support, encouragements, and thank you for letting me feel I am actually doing something of worth. I know I have

always fanned you off when you have asked me to teach you a thing or two about anticoagulants, and I hereby promise I'll teach you now. Thank you Gerald, for being the best chauffeur in the world, allowing me to have that extra sleep in the mornings on the way to school.

Last but definitely not the least, a loving thanks to my fiancé, Hansel. Thank you for going through thick and thin, ups and down together with me, Thank you for being patient with me, being understanding, and being my pillar of support.

TABLE OF CONTENTS

Summary

Ticks are hematophagous arthropods that rely exclusively on blood for their survival. During feeding, ticks inject into their hosts, a complex salivary cocktail that induces vasodilation, and impedes platelet aggregation, blood clotting and host immunity, thus overcoming host responses. These pharmacological mediators may also enhance the efficiency of pathogen transmission. Although both male and female ticks feed on blood, the manner that they feed off their host differs in length of time and volume taken up.

Firstly, to investigate the difference in salivary composition between male and female *Rhipicephalus pulchellus*, we profiled the salivary gland extracts in terms of its anticoagulant properties. While the female salivary glands extracts displayed excellent inhibition towards key blood coagulation factors FXa and thrombin, that of males showed poor inhibition properties. Further, we also established that the salivary protein content between the two genders differs.

In order to obtain information on the salivary transcriptome of *R. pulchellus*, we sequenced two cDNA libraries from pools of adult males and females salivary glands at different feeding time points, using the Illumina HiSeq protocol. *De novo* assembly of a total of 241,229,128 paired-end reads lead to the extraction of 50,460 coding sequences (CDS). In addition, we generated the proteomes of male and female *R. pulchellus* separately, which yielded a total of 454 and 2,063 proteins, respectively, which were identified by one or more peptides with at least 95% confidence.

A comparison between the male and female tick sialome revealed male- and female-specific transcripts. From the proteome, 169 and 1,777 proteins were found exclusively in males and females respectively. We hypothesize that certain classes of proteins which were highly expressed in the male glands may be involved in reproduction as males use their mouthparts to introduce their spermatophores into the females' genital pore during copulation. In addition. we have analyzed Kunitz-type serine protease inhibitors in detail and we report five new subclasses of bilaris proteins. qPCR data suggests that male and female *R. pulchellus* selectively express certain subclasses of these proteins.

The analyses of the sialomes of male and female ticks independently allow us to understand the various strategies used by each gender which enables them to feed successfully off their hosts. It has opened up opportunities to discover new salivary proteins and determine candidate male salivary proteins that may assist reproduction. Knowledge of the salivary protein repertoire of ticks may also lead to vaccine targets to disrupt feeding and/or parasite transmission as well as lead to the discovery of novel pharmacological agents.

List of Tables

List of Figures

List of Abbreviations

Single and three letter abbreviations of amino acids were followed as per the IUPAC-IUBMB Joint Commission on Biochemical Nomenclature.

Chemicals and reagents

ACN	Acetonitrile
APS	Ammonium persulfate
BSA	Bovine serum albumin
$CaCl_2$	Calcium chloride
CNBr	Cyanogen bromide
dNTPs	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
FA	Formic acid
FXa	Factor Xa
HCl	Hydrochloric acid
NaCl	Sodium chloride
PBS	Phosphate buffered saline
SDS	Sodium dodecyl sulfate
S2222	Benzoyl-IIe-Glu (Glu-γ -methoxy)-Gly-Arg-p-nitroanilide (pNA) hydrochloride (HCl)
S2238	H-D-Phe-pipecolyl (Pip)-Arg-pNA•2HCl
TBS	Tris buffered saline
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	N,N,N',N'-Tetramethylethylenediamine
TF	Tissue factor
TFA	Trifluoroacetic acid
Tris	Tris(hydroxymethyl)-aminomethane

Units

Å	Angstrom
Bp	Base-pair
Da	Daltons
h	Hour
kbp	Kilo base-pair
kDa	Kilo daltons
M	Molar
mg	Milli-gram
min	Minute
ml	Milli-litre
mM	Milli-molar
mm	Milli-metre
ng	Nano-gram
nl	Nano-litre
nM	Nano-molar
nm	Nano-metre
rpm	Revolutions per minute
µg	Micro-gram
µl	Micro-litre
µM	Micro-molar
µm	Micro-metre
V	Volt
ºC	Degree Celsius
%	Percent

Others

ADP	Adenosine diphosphate
AFXa	Anticoagulant-factor Xa

APTT	Activated partial thromboplastin time
BPTI	Bovine pancreatic trypsin inhibitor
CDD	Conserved domain database
CDS	Coding sequence
EST	Expressed sequence tags
FXaI	FXa-inhibitor
GP	Glycoprotein
HPLC	High performance liquid chromatography
IG	Immunoglobulin
IGFBP	Insulin growth factor binding proteins
iTRAQ	Isobaric tags for relative and absolute quantitation
ML	MD-2-related lipid-recognition
MS	Mass spectrometry
ORF	Open reading frame
PAR	Protease-activated receptors
PCR	Polymerase Chain Reaction
qRT-PCR	Quantitative Real Time – Polymerase Chain Reaction
RQ	Relative quantitation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SG	Salivary glands
SGE	Salivary gland extracts
TAP	Tick anticoagulant peptide
TE	Transposable elements
TEP	Thioester containing proteins
TIL	Trypsin inhibitor-like
t-PA	Tissue-type plasminogen activator
vWF	von Willebrand factor

CHAPTER 1

Introduction

1.1 Hemostasis

1.1.1 Overview

The cardiovascular system is one of the most important organ system and is crucial for the survival of an organism. Within this system, blood circulates and transports nutrients, oxygen and other compounds to all parts of the body. In vertebrates, it is a closed and high-pressured system. Thus, when there is a breach, if large enough, it can be life-threatening. It is hence critical for the body to seal this breach to prevent excessive loss of blood. Hemostasis is a process that maintains the integrity of this system when damage occurs. It is regulated by three basic mechanisms, namely vasoconstriction, platelet aggregation and blood coagulation.

1.1.2 Vasoconstriction

When injury to the blood vessels occur, the constriction of the blood vessels is the first response. Endothelial dilating agents, such as nitric oxide, adenosine and prostacyclins, which are present under normal conditions, are reduced. In addition, adenosine diphosphate (ADP), serotonin and thromboxane are released and they act on the vascular smooth muscle cells to trigger constriction of the vessels (Becker et al., 2000). Endothelin, a potent vasoconstrictor, is also produced and released by endothelial cells. This first phase of hemostasis aims to reduce and even stop the flow of blood.

1.1.3 Platelet aggregation

The second phase of hemostasis is the aggregation of blood platelet cells. Endothelial cells normally produce nitric oxide and prostacyclin I_2 which suppresses platelets adhesion and aggregation. During vessel injury when the endothelium is disrupted, collagen is exposed to the circulating blood, thus triggering the activation of platelets. Figure 1.1 illustrates a platelet and its receptors and agonists that binds to it. Collagen binds to the glycoprotein (GP) VI on the platelet while collagen-bound von Willebrand factor (vWF) binds to GPIb/V/IX and integrin $\alpha_{IIb}\beta_3$, which is the most important adhesive receptor for platelet aggregation (Jennings, 2009; Versteeg et al., 2013). Integrin $\alpha_2\beta_1$, also plays a role in platelet adhesion and anchoring, supporting platelet interaction via other platelet receptors (Clemetson and Clemetson, 2001; Nieswandt et al., 2011). The above mediates the adhesion of the platelets to the site of injury. Thereafter, platelet activation causes the release of ADP from granules with the platelets and signals the aggregation of other free platelets to the site of injury (Versteeg et al., 2013). During this process of platelet activation, the platelets undergo a change of shape to become more rounded.

A second pathway, triggered by tissue-factor and independent of collagen, can also initiate platelet activation. This pathway is dependent on thrombin, and is part of the blood coagulation cascade which will be elaborated in detail in the next section. Thrombin, which is activated through this pathway, interacts with the protease-activated receptors (PAR) on the surface of platelets thereby activating them (Furie and Furie, 2008). This results in the release of ADP, serotonin and thromboxane A_2, which further activates other platelets (Brass, 2003).

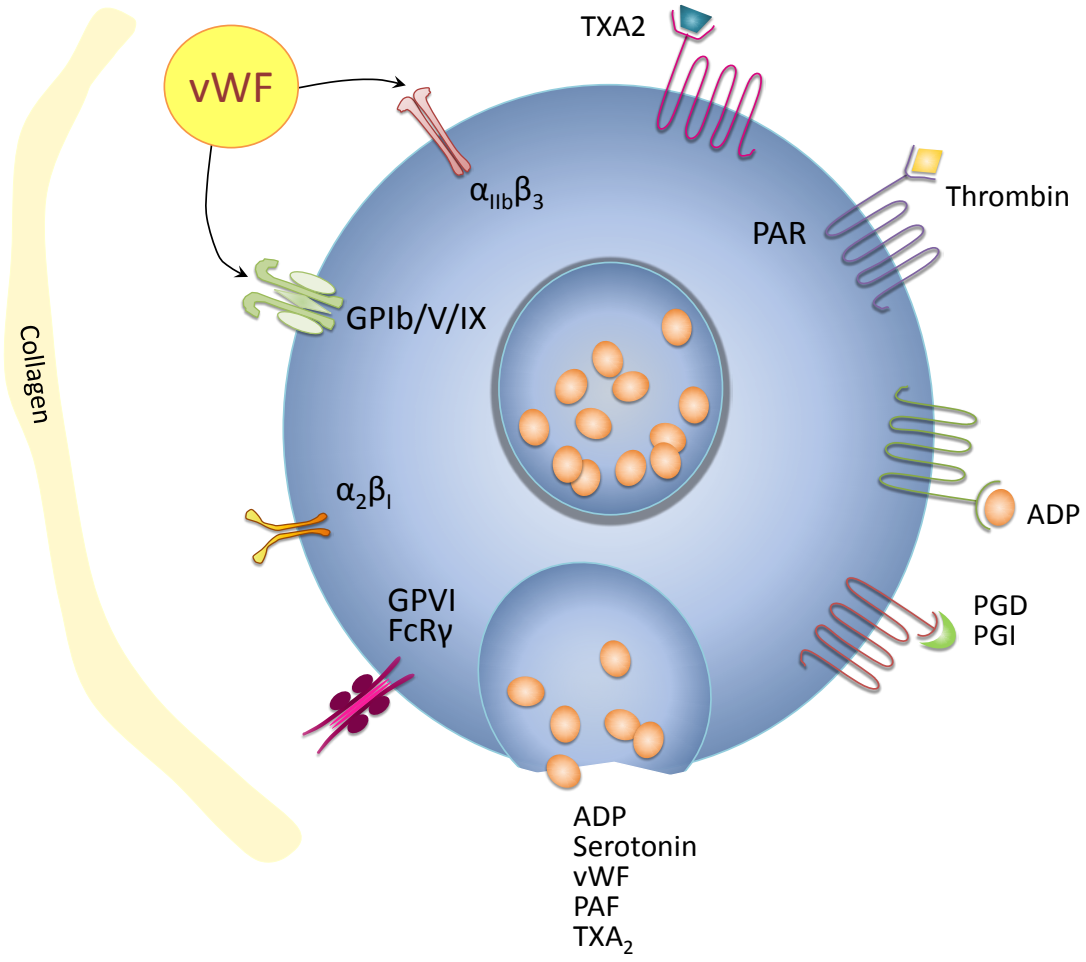

Figure 1.1 Platelet activation. Platelets have various receptors for recruitment and activation for platelet plug formation. Upon activation, the granules release agonists such as ADP, serotonin and TXA_2 which further amplifies and propagate the activation of other platelets.

1.1.4 Blood coagulation

The third phase of hemostasis is the formation of a fibrin clot to act as a mesh to prevent the outflow of red blood cells. Although it has been long thought that this process only occurs after the formation of the platelet plug (Davie et al., 1991), it has been recently revealed that both phases work concurrently to seal the site of injury and arrest bleeding (Furie and Furie, 2008). There are three main phases of blood coagulation: initiation, amplification and propagation (Monroe and Hoffman, 2006) which will be elaborated in detail below.

1.1.4.1 Initiation

Classically termed as the extrinsic pathway of blood coagulation, the initiation phase is activated when there is injury to blood vessels. Subendothelial cells (e.g. smooth muscle cells and fibroblasts) constitutively express tissue factor (TF) on their membranes. When endothelial cells are ruptured and TF on the subendothelial cells are exposed to the bloodstream, the coagulation cascade is initiated. TF, the key initiator of the extrinsic pathway, binds to blood coagulation factor VII (FVII) and activates it. Together, they form the extrinsic tenase complex (TF/FVIIa) which in turns activates FIX and FX into FIXa and FXa respectively. This allows FXa to bind to FVa to form the prothrombinase complex which converts prothrombin into thrombin.

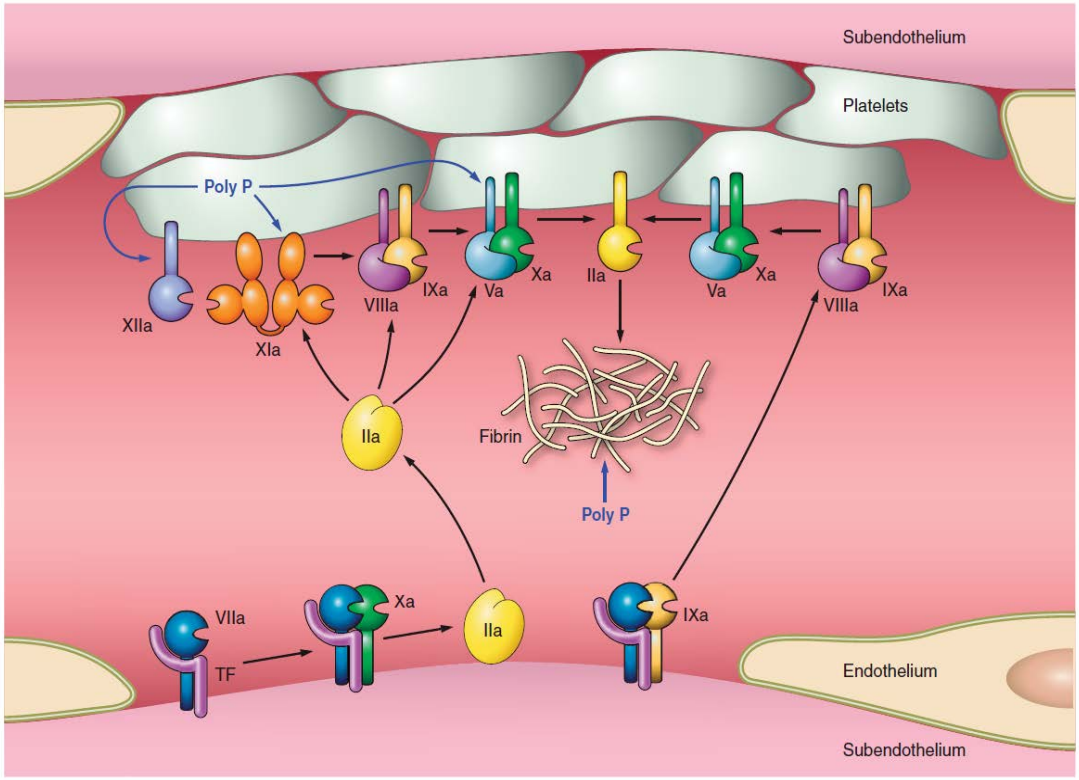

Figure 1.2 Blood coagulation cascade. Tissue factor from subendothelium cells are exposed to the bloodstream. It binds to FVIIa and together, activates FX into FXa. This complex activates thrombin (FIIa) which in turns activate FXI, FVIII, and FV. This feedback results a burst of thrombin, leading to the generation of fibrin monomers. (Adapted from Versteeg et al. (2013))

1.1.4.2 Amplification

Thrombin plays a strong role in hemostasis. As thrombin is generated and accumulated from the initiation phase, it activates many coagulation factors, which in turn leads to the generation of even more thrombin. Firstly, it activates FV into FVa which is a feedback mechanism that amplifies the prothrombinase activity. FVIII and FIXa are also activated by thrombin, which together on the surface of platelets, converts FX to FXa. These processes result in a large burst of thrombin (Furie and Furie, 2008). Thrombin also activates platelets as mentioned in Section 1.1.3 to further enhance platelet aggregation (Versteeg et al., 2013).

1.1.4.3 Propagation

With a large amount of thrombin now at the site of injury, the coagulation cascade proceeds on to its final step which is the formation of a fibrin clot. Thrombin converts fibrinogen into fibrin monomers. It also activates FXIII into FXIIIa, a transglutaminase which is responsible for the cross-linking of fibrin into a polymerized fibrin clot (Ariens et al., 2002; Versteeg et al., 2013). This formation of the cross-linked fibrin polymer, together with the platelet plug, forms a mesh such that the outflow of blood from the vessels is arrested.

1.1.5 Fibrinolysis

The site of injury has been patched by the platelet and fibrin clot to prevent bleeding. However, this is only a temporary structure formation to allow

the complete regeneration of the blood vessel walls. Once it has been fully restored, the fibrin clot is no longer required and thus will be removed through a process known as fibrinolysis. This is essential for the restoration of normal blood flow. The key enzyme responsible for the dissolution of the fibrin clot is plasmin, a serine protease. This enzyme circulates in the blood as an inactive zymogen plasminogen. The main activator of plasminogen is tissue-type plasminogen activator (t-PA) which is synthesized and secreted from endothelial cells (Dobrovolsky and Titaeva, 2002). Trypsin-like protease urokinase is also another enzyme that is able to activate plasminogen, albeit to a lesser degree. With plasmin now formed, it is able to degrade the fibrin clot and restore blood flow.

1.2 Hematophagous animals

Blood-feeding animals depend on blood meals for their survival. Some examples of hematophagous animals include mosquitoes, leeches, bed bugs, ticks, nematodes, tsetse flies and triatomines. During the blood-feeding process, hematophagous animals puncture through the skin of the host with their mouthparts and lacerate blood vessels, causing haemorrhage and destroying damaged host cells (Mans and Neitz, 2004). This leads to the activation of defence mechanisms in the host, mainly involving the hemostatic and immune systems of the host. The hemostatic system comprises of vasoconstriction, platelet aggregation and blood coagulation as detailed in Section 1.1, that aim to stop excessive loss of blood from the host. On the other hand, the immune system of the host causes pain and inflammation on the site of feeding, which leads to a grooming response by the host, aiding in awareness and subsequent removal of the parasite. In addition, humoral immunity responses result in parasite rejection, and in some cases, the death of the parasite.

In order to overcome these host responses and feed successfully, the saliva of hematophagous animals contain a cocktail of anti-clotting, anti-platelet, vasodilatory, anti-inflammatory and immunomodulatory molecules, which they infuse into the host (Champagne, 2005; Francischetti, 2010; Francischetti et al., 2009b; Lehane, 2005; Ribeiro and Francischetti, 2003; Valenzuela, 2004). In many cases, the saliva amounts are extremely minute, compared to the size of the host. However, they are still able to fully achieve their intended biological functions, indicating that they are extremely potent.

With the blood-feeding habit having independently evolved at least 20 times (Francischetti et al., 2009b; Mans and Francischetti, 2010), distinct strategies, protein scaffolds and mechanisms have emerged to counter host responses. These salivary biomolecules have undergone both convergent and divergent evolution, resulting in distinct protein scaffolds exhibiting the same function and similar protein scaffolds with diverse functions. Thus large numbers of biologically active proteins with great complexity and diversity could be found in these salivary cocktails, where the saliva of each animal may contain several hundreds to thousands of components (Francischetti et al., 2009b; Mans and Neitz, 2004).

1.2.1 Anticoagulants from hematophagous animals

We are particularly interested in the anticoagulant components within the saliva of hematophagous animals as they provide us with an excellent source of novel anticoagulant proteins and peptides which differ in structure and mechanism of action. Herein lies a short review of the current anticoagulants that have been reported in literature. Based on the mechanism of action of these anticoagulants, they can be classified into thrombin inhibitors, FXa inhibitors, extrinsic tenase complex inhibitors and intrinsic tenase complex inhibitors.

1.2.1.1 Thrombin inhibitors

There are several classes of thrombin inhibitors isolated from hematophagous animals. Each class of thrombin inhibitors show distinct protein folds as they have evolved through convergent evolution. They also inhibit thrombin through a variety of mechanisms.

Hirudin

Hirudin, which was isolated from the medicinal leech, *Hirudo medicinalis*, is probably the most prominent and well-characterized thrombin inhibitor and was reported more than 100 years ago. A 65 amino acid residue protein, hirudin is specific for thrombin, with a K_i of 22 fM (Markwardt, 1994). The key residue, Tyr64, is found to be sulphated and has been identified to play a significant role in interactions with thrombin. When this residue was removed, the protein binds to thrombin 10 times weaker (Stone and Hofsteenge, 1986). Hirudin's first three N-terminal residues bind to the hydrophobic pocket of thrombin's active site in a non-canonical form, through hydrogen bonds. The C-terminus on the other hand is rich in acidic residues and binds to exosite-I of thrombin through specific electrostatic interactions (Grutter et al., 1990; Liu et al., 2007; Rydel et al., 1990; Rydel et al., 1991; Vitali et al., 1992). Thus, hirudin specifically inhibits thrombin as a tight-binding inhibitor via these interactions with the active site and exosite-I of thrombin. Apart from hirudin, there have also been many inhibitors from other species of leeches that belong to this family of proteins that were subsequently isolated (Scacheri et al., 1993; Steiner et al., 1992).

Haemadin

Haemadin is yet another molecule which has also been isolated from leeches. Found in *Haemadipsa sylvestris,* the Indian leech, haemadin is 57 amino acid residues in length, only a few residues shorter than hirudin (Strube et al., 1993). It is a slow and tight-binding inhibitor of thrombin, with a K_i of 210 fM. Although the sequence similarity between haemadin and hirudin is low, they share a common three-dimensional fold, where the globular N-terminal core is stabilised by three disulphide linkages, with an extended C-terminal tail (Richardson et al., 2000). Just like hirudin, the first three N-terminal residues of haemadin bind to active site of thrombin in a non-canonical form. However, haemadin's acidic C-terminus binds to exosite-II of thrombin, instead of exosite-I as in the case of hirudin (Richardson et al., 2002; Richardson et al., 2000).

Kunitz-type thrombin inhibitors

The Kunitz-type inhibitors are a family of proteins that belong to the serine proteinase inhibitors (Laskowski and Kato, 1980). They have a characteristic reactive-site loop which binds and runs anti-parallel to the enzyme active site residues. This type of inhibition is commonly found in ticks. However, based on their sequences, there appears to be two different subclasses of these inhibitors which originates from the two different families of tick – Ixodidae (hard ticks) and Argasidae (soft ticks). From the hard ticks, inhibitors that have been isolated include amblin from *Amblyomma hebraeum* (Lai et al., 2004), boophilin from *Boophilus microplus* (Macedo-Ribeiro et al., 2008) and haemalin from *Haemaphysalis longicornis* (Liao et al., 2009). With

two tandem Kunitz domains, these proteins generally have a lower affinity for thrombin as compared to those from the soft ticks (see below), with a K_i of 20 nM for amblin (Lai et al., 2004) and 1.8 nM for boophilin (Macedo-Ribeiro et al., 2008). On the other hand, Kunitz-type inhibitors from soft ticks include ornithodorin from *Ornithodoros moubata* (van de Locht et al., 1996a), savignin from *O, savignyi* (Mans et al., 2002; Nienaber et al., 1999) and monobin from *Argas monolakensis* (Mans et al., 2008b). They are slow, tight-binding and competitive inhibitors of thrombin, with a K_i of 4.89 pM for savignin (Nienaber et al., 1999) and 7 pM for monobin (Mans et al., 2008b).

The three-dimensional structures of two Kunitz-type thrombin inhibitors have been well-studied, namely ornithodorin and boophilin. Ornithodorin, which comprises of two tandem Kunitz domains, binds to thrombin's active site cleft through its N-terminal residues (van de Locht et al., 1996a). Although ornithodorin contains two reactive site loops, neither of them are interacts with thrombin's active site. Amino acid residues Ser1, Leu1, Asn2 and Val3 of ornithodorin run towards Ser195 of thrombin, forming a parallel β-sheet arrangement with thrombin's Ser214-Gly219. Since, the physiological substrate of thrombin, fibrinogen, binds in a way to form an antiparallel β-sheet arrangement with Ser214T-Gly219T of thrombin, this mechanism of inhibition is termed non-canonical.

Similar to ornithodorin, boophilin binds and inhibits thrombin in a non-canonical manner engaging thrombin's active site with its N-terminal domain (Macedo-Ribeiro et al., 2008). The guanidinium group of Arg17 in boophilin anchors to the S1 pocket of the enzyme. This Arg side chain forms two

hydrogen bonds with the carboxyl group of Asp189 of thrombin at the bottom of the S1 pocket. Two additional hydrogen bonds are also formed between the terminal group of Arg17 and the main chain carbonyls of Gly219 and Phe227 of thrombin. This feature distinguishes boophilin from ornithodorin, which does not possess an Arg at this position. The boophilin residues Asn18, Gly19, Arg22, and Phe39 are also involved in interactions with different subsites of the serine protease to facilitate boophilin binding.

Kazal-type thrombin inhibitors

Another common family of serine proteinase inhibitors are the Kazal-type inhibitors (Laskowski and Kato, 1980). While the classical Kazal domains has the first two cysteine residues separated by seven to eight residues, the non-classical Kazal domains has only one to two spacer residues. Generally, thrombin inhibitors belonging to this family of proteins contain tandem non-classical Kazal domains and bind to thrombin in a slow, tight-binding and competitive mode. In these inhibitors, the first domain binds to the active site of thrombin canonically while the second domain, together with inter-domain linkers, binds to exosite-I. Examples of exogenous Kazal-type thrombin inhibitors include rhodniin from the triatomid bug, *Rhodnius prolixus* (Friedrich et al., 1993; van de Locht et al., 1995), dipetalogastin from the blood-sucking bug, *Dipetalogaster maximus* (Mende et al., 1999) and infestin from the assassin bug, *Triatoma infestans* (Campos et al., 2002). Kinetically, the K_i for rhodniin, dipetalogastin (domain 3–4) and infestin (domain 1–2) are 0.2 pM

(Friedrich et al., 1993), 0.05 pM (Mende et al., 1999) and 25 pM (Campos et al., 2002), respectively.

Variegin

Recently, a novel class of thrombin inhibitors have been identified from the tropical bont tick, *Amblyomma variegatum*. Named variegin, this fast and tight-binding thrombin inhibitor is one of the smallest thrombin inhibitors found in nature, with a length of only 32 amino acid residues (Koh et al., 2007). Structurally, variegin is flexible as it lacks secondary structures. Despite this and its small size, variegin binds highly specifically to thrombin, with K_i of 10.4 pM (Koh et al., 2007). Binding is achieved through thrombin's exosite-I with variegin's C-terminus and the active site with its middle region. Its N-terminus is found to be responsible for its fast binding kinetic properties. Interestingly, comparing structure and function, variegin resembles bivalirudin (Warkentin et al., 2008), which is the product of a human-designed, bivalent thrombin inhibitor. It is fascinating to envisage the independent development of bivalirudin through rational drug design (Warkentin et al., 2008) and compare it that of nature's own design of variegin through evolution and natural selection.

1.2.1.2 FXa inhibitors

Kunitz-type FXa inhibitors

Two anticoagulants, tick anticoagulant peptide (TAP) (Waxman et al., 1990) and FXa-inhibitor (FXaI) (Gaspar et al., 1996), were isolated from the soft ticks *Ornithodoros moubata* and *O. savignyi* respectively. These two inhibitors belong to the Kunitz-type FXa inhibitors. Unlike the Kunitz-type thrombin inhibitors, TAP and FXaI contains only a single Kunitz domain with 60 and 56 amino acid residues respectively. Both these two anticoagulants are slow, tight-binding and competitive inhibitors of FXa (Neeper et al., 1990; Waxman et al., 1990).

TAP inhibits FXa through a non-canonical active site inhibition, where its first three N-terminal residues Tyr1, Asn2, and Arg3 are in multiple contacts with the FXa active site and catalytic triad (Wei et al., 1998). Tyr1 is located in the P1 specificity pocket, forming a hydrogen bond with Ser195 of FXa. Another interaction region close to the active site is called the “secondary binding site". The secondary binding determinant of TAP consists of two distinct segments of peptides, Asp47 to Tyr49 and Asp54 to Ile560, of which the former segment interacts with Arg222 and Lys224 of FXa, and the latter peptide segment is known to interact with Arg143, Glu146, Lys147 and Arg149 of FXa. During the formation of the TAP-FXa complex, an initial slow-binding occurs at the secondary binding site, which, induces a rearrangement in the N-terminal residues of TAP to lock into the active site of FXa.

Ascaris-type FXa inhibitors

The Ascaris family of serine proteinase inhibitors are distinguished by 10 cysteine residues that form a unique disulphide pattern in a single domain (Grutter, 1994). Several Ascaris-type FXa-binding proteins have been reported from the hookworms *Ancyclostoma canium* (NAP5/6 and NAPc2/3/4) (Cappello et al., 1995; Mieszczanek et al., 2004b; Stassens et al., 1996) and *A. ceylonicum* (AceAP1) (Harrison et al., 2002). While NAP5/6 and AceAP1 binds to the active site of FXa (Harrison et al., 2002; Stassens et al., 1996), NAPc2/3/4 binds to the exosite (Mieszczanek et al., 2004b; Stassens et al., 1996). With NAPc2/3/4 and AceAP1 exploiting the binding of FXa to inhibit the FVIIa-TF complex, they are considered as extrinsic tenase complex inhibitors (see below for details). NAP5 inhibits FXa canonically at the active site, with a K_i of 43 pM (Stassens et al., 1996). Similar to TAP, it also interacts with the Na^+-binding site and autolysis loops of FXa (Rios-Steiner et al., 2007).

Antistasin-like FXa inhibitors

Many leeches have anticoagulants that contain the antistasin-like domain. This class of FXa inhibitors include antistasin (Nutt et al., 1988; Tuszynski et al., 1987), ghilanten (Brankamp et al., 1990) and therostasin (Chopin et al., 2000). Antistasin is a 119 amino acid protein that was isolated from the Mexican leech, *Haementeria officinali*. It contains two tandem antistasin-like domains which has 10 cysteines that form five intra-domain disulphide bridges. Antistasin is a slow, tight-binding and competitive inhibitor, with a K_i of 0.3-0.6 nM. Inhibition of the active site of FXa is achieved through

the canonical reactive-site loop residing in the N-terminal domain (Dunwiddie et al., 1989). As for therostasin, although this canonical reactive-site loop and the domain signature are conserved, its overall sequence similarity with antistasin is low.(Chopin et al., 2000) It is noted that therostasin is more potent than antistasin, with a K_i of 34 pM.

Serpin family of FXa inhibitors

Serpins are a superfamily of proteins (45-55 kDa) that undergo huge conformational changes and form a covalent complex with the target proteinase, thereby inhibiting it irreversibly (Otlewski et al., 2005). An example of a FXa inhibitor that belongs to this class of proteins is found in the saliva of the yellow fever mosquitos, *Aedes aegypti*. The crude salivary gland extract of female *A. aegypti* was found to inhibit the FXa active site reversibly and non-competitively (Stark and James, 1995). Fifty four kDa in size, this protein, named as anticoagulant-factor Xa (AFXa), contains post-translational modifications (four *N*-linked glycosylation sites) which are likely to be important for its activity. However, as compared to typical serpins, AFXa has a shorter reactive site-loop and different hinge residues.

1.2.1.3 Extrinsic Tenase Complex Inhibitors

There are two main classes of extrinsic tenase complex inhibitors that have been isolated from hematophagous animals. Both classes act through a similar, but not identical, mechanism as the physiological tissue factor pathway inhibitor (TFPI). The first class of inhibitors are the Kunitz-type inhibitors. Two such inhibitors have been isolated from the hard tick, *Ixodes scapularis* – ixolaris (Francischetti et al., 2002b) and penthalaris.(Francischetti et al., 2004) Ixolaris, 15.7 kDa in size, has two tandem Kunitz domains. It was postulated that the second Kunitz domain binds first to FX/Xa, then followed by the FVIIa-TF complex via the first Kunitz domain (Francischetti et al., 2002b). Ixolaris binds to FX and FXa with affinities between 0.5–10 nM, but not at the FXa active site (Monteiro et al., 2008b). The surface amino acid residues of Factor X/Xa that are involved in the binding of ixolaris overlap largely with the heparin binding proexosite/exosite (Monteiro et al., 2008b; Monteiro et al., 2005a). This binding of ixolaris to FX and FXa impairs their interactions with FVIIIa and prothrombin, respectively (Monteiro et al., 2008b; Monteiro et al., 2005a). Penthalaris, on the other hand, has five Kunitz domains, as compared to only two in ixolaris. Penthalaris inhibits the FVIIa-TF complex in the same way as ixolaris, making use of FX or FXa as a scaffold. The contribution of the three additional Kunitz domains of penthalaris to the interaction/function has yet to be elucidated (Francischetti et al., 2004).

The second class of extrinsic tenase complex inhibitors are the Ascaris-type inhibitors, isolated from the hookworms *Ancyclostoma canium* (Mieszczanek et al., 2004b) and *A. ceylanicum* (Mieszczanek et al., 2004a;

Mieszczanek et al., 2004b). They are single-domain Ascaris-type inhibitors, such as NAPc2 and AceAP1. NAPc2 binds to the FXa heparin binding exosite with its extended C-terminus, before binding to the FVIIa-TF complex (Murakami et al., 2007). In contrast, AceAP1 uses FXa as a scaffold by binding to both its active site and exosite and forms a quaternary complex with the FVIIa-TF complex (Mieszczanek et al., 2004a; Mieszczanek et al., 2004b).

1.2.1.4 Intrinsic Tenase Complex Inhibitors

Till date, Nitrophorin-2 (or Prolixin-s), isolated from *Rhodnius prolixus* (kissing bug), is the only intrinsic tenase complex inhibitor that has been identified from animals (Ribeiro et al., 1995). Approximately 20 kDa in molecular weight, it belongs to the lipocalin family. By targeting the Gla-domain of FIX/FIXa, Nitrophorin-2 binds specifically to FIX/FIXa with a K_d of about 13 nM. This binding interferes with FIX activation (by both the FVIIa-TF complex and FXIa) and with FIXa activity in the intrinsic tenase complex (Isawa et al., 2000). The surfaces on Nitrophorin-2 involved in FIX/FIXa binding have been suggested to be within its B-C and E-F strands as well as its C-terminus (Andersen and Montfort, 2000).

1.3 Ticks

As described in the above sections, ticks have been proven to be an enormous source of bioactive molecules from their saliva. They have long been an area of interest in the isolation of novel compounds due to their diversity. We take a closer look at two aspects of ticks that are closely related to their saliva - their feeding behaviours and their sexual reproduction.

1.3.1 Feeding behaviours of ticks

Some adult hematophagous arthropods are either facultative or obligatory blood-feeders. In other words, some only feed on blood when there is a requirement for it or when blood meals are available, while others depend entirely on blood meals for their survival. However, in certain species (e.g., mosquitoes and horse-flies) the blood-feeding habit differs between sexes; females are facultative blood-feeders while the males do not feed on blood at all (Lehane, 2005). This difference in feeding habit between the sexes is in fact reflected in their salivary components. For example, in the mosquito *Anopheles gambiae*, anti-clotting and anti-platelet proteins are expressed exclusively in females and assist them in blood-feeding, while anti-microbials, and maltases and other enzymes that aid in digestion of plant sugar meals, are expressed in both sexes (Arca et al., 2005; Calvo et al., 2006).

On the other hand, in some obligatory blood-feeders (e.g., ticks, tsetse flies, and kissing bugs), although both sexes feed on blood, there are differences in their feeding behaviors. For example, female ticks feed on a

larger volume of blood as compared to male ticks, and their body weight can increase to more than 100-fold after feeding (Chen et al., 2012a; Chen et al., 2012b; Ma et al., 2013; Zheng et al., 2011). This is in contrast to males, which rarely engorge to higher than two-fold their body weight. The relationship between the salivary gland components and sexual differences in feeding or sexual behavior of the ticks is not well understood and hence, of interest.

1.3.2 Sexual reproduction in ticks

The saliva of male ticks has been reported to play a role during the course of copulation in ticks. It has been observed that during copulation, after the male tick's spermatophore is completely formed, a droplet of saliva appears on the chelicerae of the male (Feldman-Muhsam et al., 1970). The male tick then uses its chelicerae to grasp the spermatophore to insert into the genital aperture of the female tick. As it does this, the saliva from the chelicerae comes into contact with the spermatophore and spreads over the upper portion of it (Feldman-Muhsam et al., 1970). The mouthparts of the male tick also actively penetrate the female's genital pore while pushing-in the spermatophore (Kiszewski et al., 2001). It is speculated that the saliva may assist this process by lubricating it or preventing it to adhere to the tick surfaces while transiting into the female genital pore, thus facilitating its transfer (Feldman-Muhsam et al., 1970). It is possible that in addition to this lubricating or anti-stickiness property of saliva, male salivary molecules may play a role traditionally played by arthropod male accessory glands to increase the female's fecundity (Kaufman, 2004). Male ticks remain with their mouthparts inside the female's

genital pore for several hours after successful spermatophore transfer, perhaps as a mechanism of guarding the females from other males and thus avoid sperm competition (Kiszewski et al., 2001).

1.4 Aim and scope of the thesis

Heart attack and stroke are the major causes of death and debilitation globally, and accounts for more than 20% of the death in Singapore in 2012 (Ministry of Health, 2012). These cardiovascular diseases are associated with unwanted clot formation in the bloodstream and thus, the prevention and treatment requires the use of anticoagulants and antiplatelet therapies. Coumarins and heparin are the most well-known clinically used anticoagulants. However, they have their limitations and side effects due to their non-specific mode of actions. These limitations have resulted in the demand for highly specific novel methods to prevent blood coagulation and to regulate complications in atherosclerosis, heart diseases, stroke and cancer, without being associated with the negative side effects of drugs. The need for the development of new anticoagulants, which target specific coagulation enzymes or a particular step in the clotting process is vital.

The salivary gland extracts of ticks have proven to be an exploitable source of novel biomolecules in which many lead compound could be isolated from. In this thesis, we look at the salivary components of two ticks - the zebra tick, *Rhipicephalus puchellus* and the marsh tick, *Dermacentor reticulatus*. *R. pulchellus* is native to Eastern Africa (Walker, 1955) and is responsible for the high infestation in cattle and buffaloes (Kariuki et al., 2012). On the other hand, D. reticulatus is found in Europe and central Africa and is the vector of the Russian spring summer encephalitis virus and the Flavivirus. These particular tick species were made available to us by our collaborators in Slovakia. As preliminary testing was done on the saliva of the ticks as described in detail later, the story unfolded itself, bringing a particular interest into the salivary

composition differences between male and female ticks. As such, subsequent experiments were carried out in a manner where male and female *R. pulchellus* were analyzed independently.

The specific aims of this thesis are:

1. To investigate the anticoagulant profile of *R. pulchellus* and *D. reticulatus* ticks
2. To generate the sialotranscriptome of male and female *R. pulchellus*
3. To generate the proteome of male and female *R. pulchellus*
4. To delineate the salivary composition differences between male and female *R. pulchellus*
5. To explore the Kunitz-type protease inhibitor class of proteins

On the whole, the scope of this thesis covers the methods used to achieve these aims, the results obtained from the study, as well as a discussion of the inferences made. Herein, the salivary composition of the *R. pulchellus* ticks is presented for the very first time. In addition, we present a different perspective of looking at the sialomes of these ticks by examining each gender independently. Hence, this thesis opens up a huge repertoire of proteins that may potentially be lead molecules for drug development, and also gives an insight into the biology of how ticks feed.

CHAPTER 2

Materials and Methods

2.1 Salivary glands and extracts

Salivary gland and salivary gland extracts of ticks were obtained from collaborators RNDr. Mirko Slovak and RNDr. Maria Kazimorova from the Slovak Academy of Sciences, Slovakia. The *R. pulchellus* tick species was the kind gift of Dr. Milan Kozánek (Institute of Zoology SAS, Slovakia) who collected them from West Tsavo, Kenya (determined by M.Slovák). The ticks were reared under laboratory conditions (Slovak et al., 2002) in the Institute of Zoology, SAS, Slovakia. Ticks used in the experiments resulted from the fourth breeding generation. Briefly, the ticks were maintained at a temperature of 24 ± 2 ºC in desiccators filled with concentrated KCl solution, with 85-90% relative humidity and a photoperiod of 16:8 h (L:D). White New Zealand rabbits were used as hosts for all stages and also for feeding of adult ticks for the given intervals mentioned in the experiments below. The usage of animals in these experiments was approved by the State Veterinary and Food Administration of the Slovak Republic (permit numbers 928/10-221 and 1335/12-221). Ticks were removed from the rabbits, according to the number of days they have been feeding for, and their salivary glands were dissected in ice-cold sterile phosphate buffered saline (PBS), pH 7.2, and washed three times in PBS.

For the preparation of salivary gland extracts, glands were first homogenized in 150 µl of PBS, using a handheld homogenizer in a micro tube. The homogenized mixture was subsequently centrifuged at 15,000 rpm, after which the supernatant was harvested as the salivary gland extracts. This was repeated a second time with the pellet from the centrifuged sample, and the supernatants were pooled, to increase protein yield from the extraction. The

salivary gland extracts were dried with either a freeze drier or a centrifugal evaporator for long term storage.

For preparation of intact salivary glands for subsequent RNA work, glands were immersed into RNAlater reagent (Qiagen, Hilden, Germany) immediately after dissection. Tissues were kept in 4ºC for a minimum of 48 h to ensure penetration of RNAlater solution into the tissue. Thereafter, they were kept frozen in -30ºC until ready for processing.

2.2 Purification and anticoagulation activity testing

2.2.1 Protein quantification

Salivary gland extracts of 6-day fed male and female ticks were reconstituted with distilled water and filtered with a 0.45 µm Minisart® RC4 regenerated cellulose membrane syringe filter (Sartorius Stedium Biotect, Goettingen, Germany). Protein quantification was performed by the Bradford assay. The Bradford reagent concentrate (Bio-rad Laboratories, California, USA) was diluted five times with distilled water. 200 µl of the diluted reagent was mixed with 10 µl of sample in a 96-well microtitre plate, and incubated at room temperature for 5 min. Absorbance was read at 595 nm with a microplate reader (Tecan M200, Männedorf, Switzerland). The concentration was estimated with a standard curve using BSA.

2.2.2 Enzymatic inhibition assays

Activities of crude extracts and partially purified fractions were determined with enzymatic inhibition assays where they were tested for their ability to inhibit either thrombin or FXa amidolytic activity. This was monitored by using a chromogenic substrate, specific to the enzyme. All assays were of a 40 µl reaction, performed in a 384-well microtiter plate at room temperature. 10 µl of enzyme (6.6 nM thrombin / 5.3 nM FXa (Haemtech, Vermont, USA) in 50 mM Tris, 120 mM NaCl, 5 mM $CaCl_2$ and 1% BSA) and 10 µl of protein were pre-incubated in the reaction well for 10 minutes. Subsequently, 20 µl of substrate (S2238, 200 µM, for thrombin; S2222, 2 mM, for FXa) (Chromogenix, Milano, Italy) was added into the mixture and the kinetics of the reaction was monitored immediately with a microplate reader (Tecan M200, Männedorf, Switzerland) at an absorbance of 405 nm. Both substrates upon being cleaved by its corresponding enzyme, releases a chromogenic compound, p-nitroaniline, which can be detected at 405 nm. The initial velocity of this reaction was monitored. Percentage inhibition of the enzyme by the protein was calculated relative to the initial velocity of the control reaction without protein (10 µl PBS) which was taken to be 0%.

$$Percentage\ inhibition\ = \frac{v_{(control)} - v_{(sample)}}{v_{(control)}} \times 100\%$$

2.2.3 Purification of crude salivary extracts

2.2.3.1 Size exclusion chromatography

Salivary gland extracts were first partially separated by size exclusion chromatography with a Superdex 75 10/300 GL (GE Healthcare, Uppsala, Sweden). 50 mM Tris, pH 7,4, was used for equilibration of the column and the salivary glands, approximately 100 pairs, were reconstituted in 1 ml of deionized water for injection into the column. Protein elution was monitored at a UV wavelength of 280 nm and 215 nm. The flow rate used was 0.8 ml/min and the proteins eluted were collected as 0.8 ml fractions.

2.2.3.2 FXa affinity chromatography

Fractions from the size exclusion chromatography were subjected to affinity chromatography with FXa which was coupled to CNBr-Activated Sepharose 4B (GE Healthcare, Uppsala, Sweden) according to the manufacturer's protocol. 1 ml of FXa coupled sepharose beads were used for the purification in a 10 ml poly-prep column (Bio-rad, California, USA). Protein samples were allowed to incubate with the beads for 20 minutes, with gentle rocking. The wash step was performed with 40 ml of TBS and protein elution was performed with increasing concentrations of NaCl (0.5 M, 1.0 M, 1.5 M and 2.0 M) in 20 mM HCl. Tris 1M, pH 8.0 solution was immediately added to the eluate so as to buffer the pH of the solution.

2.2.3.3 Reverse phase HPLC

Active fractions were subsequently purified by a reverse phase HPLC with the Agilent Zorbax 300SB-C18 column, 1.0 x 150 mm, 3.5 µm (Agilent, California, USA). The column was equilibrated with 15% acetonitrile (ACN) with 0.1% Trifluoroacetic acid (TFA). Proteins were eluted at a flow rate of 40 µl/min with a gradient of 100% ACN with 0.1% TFA. The elution was monitored at a wavelength of 215 nm and 280 nm. Fractions were collected according to the eluted peaks. The fractions were lyophilized and reconstituted for anticoagulant activity testing.

2.3 Transcriptomics

2.3.1 cDNA library construction

Salivary glands from the following time points were pooled as follows: Unfed, 1, 3 and 7 h, 1, 2, 3, 4, 5, 6 and 7 days. Six adult ticks from each gender were used per time point, except for 1-day fed where 12 ticks were used. Male and female SGs were collected and processed as two separate samples. Tick SGs were prepared as in Section 2.1. The extracted mRNA was fragmented using a Covaris E210 (Covaris, Woburn, MA). Library amplification was performed using eight cycles to minimize the risk of over-amplification. Unique barcode adapters were applied to each library. Individual libraries were quantitated by qPCR and then pooled in an equimolar ratio before sequencing on a HiSeq 2000 (Illumina) with ver. 3 flow cells and sequencing reagents. Two lanes of the HiSeq machine were used. To avoid lane bias, the two libraries

were run together in both lanes. Raw data were processed using RTA 1.12.4.2 and CASAVA 1.8.2.

2.3.2 Transcriptome assembly and bioinformatics

Reads obtained from the Illumina sequencing were assembled with the Abyss software (Birol et al., 2009; Simpson et al., 2009) with various k values (every even number from 50 to 96). Because the Abyss software tends to miss highly expressed contigs (Zhao et al., 2011), we have also run the Trinity assembler (Grabherr et al., 2011) on the raw data. The resulting assemblies were joined by an iterative blast and cap3 assembler (Karim et al., 2011). Coding sequences (CDS) were extracted using an automated pipeline, based on similarities to known proteins, or by obtaining coding sequences from the larger open reading frame (ORF) of the contigs containing a signal peptide. A non-redundant set of the coding and their protein sequences were mapped into a hyperlinked excel spreadsheet. Signal peptide, transmembrane domains, furin cleavage sites and mucin-type glycosylation were determined with software from the Center for Biological Sequence Analysis, Denmark (Duckert et al., 2004; Julenius et al., 2005; Nielsen et al., 1999; Sonnhammer et al., 1998). Detailed bioinformatic analysis of the pipeline can be found in a reference publication (Karim et al., 2011). To map the raw Illumina reads to the coding sequences and determine their sex bias, raw reads from each library were blasted to the coding sequences using blastn with a word size of 25 (-W 25 switch) and allowing recovery of up to three matches. The three matches were used if they had less than two gaps and if their scores were equal to the

best score. The resulting blast file was used to compile the number of reads each CDS received from each library, and also to count the number of hits at each base of the CDS, allowing for the determination of the average CDS coverage, maximum coverage and minimum coverage. These results can be statistically tested by a X^2 test (using the number of reads per CDS), the results of which are reported significant if $P<0.05$ and no CDS had an expected value of 5 or less.

2.3.3 Sequence analyses

Sequences were aligned with Clustal X (Jeanmougin et al., 1998)., and phylogenetic analyses were done with Mega (Tamura et al., 2011). In addition, CLC sequence viewer Ver. 6 (CLC Bio Inc., Massachusetts, USA).was used to generate the consensus sequence and for visualization.

2.4 Proteomics

2.4.1 Tryptic Digestion

50 µg (1 mg/ml) of salivary gland extracts used as the starting material. Reduction was first performed where 100 mM tris(2-carboxyethyl)phosphine (TCEP) was added to the extracts to a final concentration of 5 mM and allowed to incubate at 60ºC for 1 h. This was followed by alkylation with 1 M iodoacetamide in 100 mM ammonium bicarbonate (final concentration: 40 mM iodoacetamide). Trypsin (Promega, Wisconsin, USA) (1 µg/µl in 50 mM acetic

acid) was added to the mixture at a ratio of 1 µg trypsin : 30 µg protein. Incubation of the mixture was done at 37ºC for 16 h, with light shaking.

2.4.2 SDS-PAGE

Complete digestion of the sample was confirmed by running 1 µl of pre-digested and post-digested sample on a 12% (w/v) SDS-PAGE gel. SDS-PAGE gels were made with the following recipe in Table 1.

Table 1. SDS-PAGE composition

Reagent	Stacking Gel 4%	Resolving Gel 12%
Distilled water (ml)	6.36	4.35
40% (w/v) Acrylamide/Bis (ml)	1 ml	3
1.5 M Tris-Cl, pH 8.8 (ml)	-	2.5
0.5 M Tris-Cl, pH 6.8 (ml)	2.52	-
10% (w/v) SDS (µl)	100	100
TEMED (µl)	10	5
10 % (w/v) APS (µl)	50	50
Total Volume (ml)	**10**	**10**

SDS loading dye (62.5 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% glycerol, 50 mM DTT, 0.01% (w/v) bromophenol blue) was added to the sample and heated at 100ºC for 10 min. After allowing the mixture to cool down to room temperature, samples were loaded into the SDS-PAGE gel. Electrophoresis was carried out 180 V in a mini-PROTEAN® cell (Bio-rad Laboratories,

California, USA), powered by the PowerPac™ Basic Power Supply (Bio-rad Laboratories, California, USA). Gels were subsequently rinsed briefly in distilled water and then stained with a silver-stain kit (Thermo Fisher Scientific Inc., Massachusetts, USA) according to the manufacturer's instructions.

2.4.3 Sample clean-up

Digested samples were desalted with a Sep-Pak plus short cartridge (C18, 55-105 µm) (Waters, Massachusetts, USA). The cartridge was first conditioned with 10 ml 100% ACN, and then equilibrated with 10 ml 2% ACN. Sample mixture was diluted with 2% ACN to approximately 5 ml before passing it through the cartridge twice. A wash step was performed with 10 ml 2% ACN and the peptides were eluted with 10 ml 80% ACN. Sample was freeze dried.

2.4.4 Mass spectrometry

The dried sample mixture was reconstituted with 50 µl of 5% ACN and 0.05% formic acid (FA) for LC MS/MS analysis. For each sample, 2 µg was injected into a reverse phase column (75 µm × 150 mm, ReproSil-Pur C18-AQ, 3 µm, 120 Å (Eksigent, California, USA)). The mobile phase A used was 2% ACN, 0.1% FA and mobile phase B was 98% ACN, 0.1% FA. The gradient (in terms of mobile phase B percentage) for the liquid chromatography was as follows: 5% for 1 min, 5-12% for the 2 min, 12-30% for 120 min, 30-90% for 1 min and lastly held at 90% for 11 min, with a flow rate of 300 nl/min. MS analysis was performed with the TripleTOF™ 5600 system (AB SCIEX).

Data acquisition was performed in information dependent mode, where the mass range was set to 400-1250 m/z with an accumulation time of 250 ms per spectrum. A maximum of 20 precursor ions were subjected to MS/MS analysis, with dynamic exclusion enabled for a duration of 15 s. Proteins were identified using the ProteinPilot™ Software v. 4.5 (AB SCIEX) which uses the Paragon™ Algorithm. ID focus was set to biological modification (refer to suppl. for the full list of modifications as set by default within the program), with cysteine alkylation by iodoacetamide, digestion by trypsin, and search effort set to thorough. The database used for the search was the proteins derived from the transcriptome of R. pulchellus, containing 50,460 protein sequences. The false discover rate (FDR) analysis was also performed simultaneously with the Proteomics System Performance Evaluation Pipeline (PSPEP) feature in the software. It adopts a decoy database search strategy in which the protein database sequences were reversed and searched against. The reported FDR for both male and female proteomes for proteins with at least 95% confidence were less than 1% based on the global FDR fit. Only proteins with at least one peptide of >95% confidence are reported and hits which corresponded to the reversed sequences were removed.

2.5 Quantitation of differential expression

2.5.1 RNA isolation and first-strand cDNA synthesis

Salivary glands in RNAlater were thawed and weighed. Salivary glands from male (12.8 mg) and female (20.5 mg) ticks were used for RNA extraction. Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany)

according to the manufacturer's protocol. Total RNA (1 µg from each sample) was used for first-strand cDNA synthesis using the Reverse Transcription & cDNA Synthesis Kit (Clontech) with the SMART Moloney Murine Leukemia Virus (MMLV) Reverse Transcriptase (Clontech), following the manufacturer's instructions.

2.5.2 Primer design

Primers were designed with the help of Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) which also checks for primer specificity. Generally, primer size was set to be between 18-25 bp, with Tm between 65°C to 70°C and GC% within 40-60%. The PCR product size was between 150 – 350 bp, with an optimal of 250 bp. The list of primers used are listed in Table 2.

Table 2. List of primers for qPCR

Gene	Forward Primer Sequence (5'-3')	Length (bp)	Reverse Primer Sequence (5'-3')	Length (bp)	Product Size (bp)
RpSigp-759502	TGCCTTTTCCGTTCCTTGCGTGCC	24	ACGGGGACTTTGGTGCGCATTCC	23	273
RpSigp-907535	ACTGTGAACGGACGAGCCGCA	21	CTCCGCTGGCCACTCATCCCA	21	345
RpSigp-907535 (2)	CGGCACCACACTGCAGTCCACCT	23	ACCGTTCTCCGCTGGCCACTCAT	23	284
RpSigp-673073	TGGCCGTAGTTTCCCTCGGCGTC	23	AGGCCATGCTCCAGGGGTGTTTT	23	265
RpSigp-921556	GTGGGCAAATCCCAACGCCATGCT	24	TCTCCACCAAGACGCAACGTCCTCT	25	142
RpIx75-699640	ACCACCAACCGTGGAAGGGTGC	22	CTTCGGCCGTACGGGTCTGGGT	22	255
RpIx75-477142	AGTGCGAAGCTTCTTGCCCGCC	22	TCTTCGGCCGTACGGGTCTGGG	22	134
RpIx75-909901	GCAGTACTCGGCAAACGCCCA	21	TGTGGTGAAGGCCCGCATCCAGT	23	345
RpIx75-485129	GCAGCGCCACTGCTTCAAGGAACC	24	AATCCTTCGGTTTCGGCGCAGGC	23	219
RpSigp-842790	GCCAGTGGCCTGCTACATGAAACCA	25	TGCCAGGAGGACGCAGACACGG	22	212
RpSigp-843116	TCCGTGTCTGCGTCCTCCTGGC	22	TGCACTTGTCGCAAGACGGGAACA	24	155
RpIx75-680109	ACGCGGCCGCTGCAGAATGT	20	CCAACCCTTGGGACACCGGCCA	22	250
RpIx75-657025	CGCCTTCGCTGCTGGCTACC	20	TGGGGGCAAGCAGGCATTGT	20	171
RpIx75-947624 (Beta-actin)	CGCCCTTCCCCACGCCATCC	20	GAACAGGGCCTCGGGGCAGC	20	280

2.5.3 Polymerase chain reaction amplification

Primers were checked for specificity and ability to amplify the target gene by performing a PCR amplification on the cDNA of male and female *R. pulchellus*. For PCR reactions, the KAPA Taq PCR kit (KAPA Biosystems, Massachusetts, USA) was used, in which contained the KAPA Taq polymerase, Buffer B for the polymerase and dNTPs. The reaction mix used was as followed in Table 3.

Table 3. PCR reaction mix

Reagent	Volume (µl)
Distilled water	7.56
KAPA Taq polymerase	0.04
10X Buffer B	1
10 mM dNTPs	0.2
cDNA (150 ng/µl)	1
10 µM Forward Primers	0.1
10 µM Reverse Primers	0.1
Total Volume (ml)	**10**

The PCR cycling parameters used were as follows: initial denaturation of 95 °C for 3 min; 35 cycles of 95 °C for 30 s, 63 °C for 30 s and 72 °C for 30 s; and final extension of 72 °C for 1 min.

2.5.4 DNA gel electrophoresis

The PCR products were visualized with a 1.5% (w/v) agarose gels. Briefly, molecular biology grade agarose (1st Base Laboratories, Selangor,

Malaysia) was dissolved in a 1 X TAE buffer (40 mM Tris-acetate and 1 mM EDTA, pH 8) and brought to a boil. GelRed solution (10,000X) added into the molten agarose gel solution with a dilution of 1:50,000 (v/v). DNA samples were mixed with a 6X loading dye (Fermentas, Massachusetts, USA) before loading into the agarose gels. A 100 bp or 1 kbp GeneRuler DNA ladder (Fermentas, Massachusetts, USA) was also run together with the samples in a separate well to provide an estimation of size of the DNA samples. Gels were generally run at 90 V for about 30 min. After separation, DNA bands were visualised with a transilluminator.

2.5.5 DNA sequencing

In order to verify the sequence of the amplified product, the PCR product was excised from the gel, and a clean-up of the sample was performed with a commercially available kit. The QIAquick gel extraction kit (Qiagen, Hilden, Germany) was used according to the manufacturer's protocol. Thereafter, a cycle sequencing PCR was carried out with a reaction mix as in Table 4, with the using BigDye Terminator version 3.1 cycle sequencing reagents (Applied Biosystems Life Technologies, Massachusetts, USA).

Table 4. Cycle sequencing reaction mix

Reagent	Volume (µl)
Distilled water	0.5
DNA	2
10 µM Primer	0.5
BigDye Terminator	2
Total Volume (ml)	**5**

The cycling protocol was as follows: Initial 96 °C for 1 min, 30 cycles of 96 °C for 10 s, 52 °C for 5 s and 60 °C for 2 min. The cycle sequencing products were purified with Axygen™ AxyPrep Mag™ DyeClean kit (Fisher Scientific, Massachusetts, USA) according to the manufacturer's protocol. Sequencing was the performed on Sanger sequencer machine ABI 3130 XL (Applied Biosystems Life Technologies, Massachusetts, USA).

2.5.6 Quantitative Real Time – Polymerase Chain Reaction

The qRT-PCR analyses were performed using "StepOne™ Real-Time PCR Systems" (Applied Biosystems Life Technologies, Massachusetts, USA) and analysed using the "StepOne™ Software" (v2.1; Applied Biosystems Life Technologies, Massachusetts, USA). Each PCR reaction had a total final volume of 10 µl comprising of 5 µl of KAPA SYBR® FAST ABI Prism® 2X qPCR Master Mix (KAPA Biosystems, Massachusetts, USA), 25 ng of cDNA, and forward and reverse primers (125 nM each). Each reaction was set up in technical duplicates i.e. two reactions for the target gene and two for the internal control from the same cDNA. The experiment was confirmed with a biological replicate. PCR protocol used was follows: Initial denaturation and hot-start activation at 95°C for 20 s, followed by 40 cycles of denaturation at 95°C for 3 s and annealing/extension at 63°C for 30 s. Melting curve analysis was done from 60°C to 95°C at the end of each PCR run, for detecting non-specific PCR product and/or primer-dimer co-amplification. For normalization, β-actin was used as the internal control gene. Results were expressed as relative

quantitation (RQ) in reference to mRNA of the female gene for each set of genes, using the ΔΔCt method.

2.6 Recombinant expression

2.6.1 Primer design

Primers were designed to include a Kozak translation initiation sequence and an ATG start codon to ensure proper initiation of translation. Thus, the forward primers include a sequence “AAA ATG G” at the 5’ end. Reverse primers were lengthened with a “GGG” sequence at their 5’ ends. The length of all primers were capped at a maximum of 40 residues long and the list of primers and their sequences are as in Appendix 2.

2.6.2 Cloning

The genes of interest were amplified from the cDNA and ligated into a pIB/V5-His vector (Life Technologies, Massachusetts, USA) as per the manufacturer’s protocol. They were then transformed into TOP10 One Shot® (Life Technologies, Massachusetts, USA) competent cells. The competent cells were thawed on ice and 2 µl of the ligation mixture was added to the cells and incubated on ice for 15 min. This was followed by a heat shock step at 42ºC for 30 s and immediately cooled on ice. 250 µl of SOC medium (room temperature) was added to the cells and this mixture was allowed to incubate for 1 h at 37ºC with horizontal shaking at 200 rpm. Cells (100 µl) were then plated on LB agar plates containing 100 µg/ml ampicillin and incubated for 16 h at 37ºC.

2.6.3 Colony screening

Colonies were picked and inoculated into 5 ml of LB broth containing 100 µg/ml of ampicillin. The tubes were incubated for 16 h at 37 ºC at a shaking speed of 200 rpm. Plasmids were isolated from the culture using the GeneAll™ plasmid extraction kit (GeneAll Biotechnology, Seoul, Korea) following the recommended protocol from the manufacturer. DNA sequencing of the plasmid was performed as previously described in Section 2.5.5 with the OpIE2 forward primer, 5´-CGCAACGATCTGGTAAACAC-3´, and OpIE2 reverse primer, 5´-GACAATACAAACTAAGATTTAGTCAG-3´. Plasmids were checked for orientation and sequence of insert.

2.6.4 Transfection

The High Five™ insect cell line (Life Technologies, Massachusetts, USA) was used for expression. A total of 2 X 10^6 cells were seeded into a 60 mm round dish. The dish was rocked side to side for 2 min and incubated thereafter for 15 min to allow attachment of cells onto the bottom of the dish. The transfection mixture was prepared by adding 1 ml of Express Five serum-free media (Life Technologies, Massachusetts, USA) to 1 µg of construct. 20 µl of Cellfectin® Reagent (Life Technologies, Massachusetts, USA) was added and the mixture was mixed gently for 10 s and incubated at room temperature for 15 min. The media was aspirated from the cells and the transfection mixture was added drop-wise into the dish. The dish was incubated at room temperature for 4 h,

with side-to-side rocking motion. 2 ml of serum-free media was then added to the dish and it was incubated at 27ºC for 4 days.

2.6.5 Activated partial thromboplastin time

The activated partial thromboplastin time was used to screen for anticoagulant properties of the expressed proteins. Citrated human plasma was (50 µl) was incubated with equal amount of supernatant at 37ºC for 10 min. 50 µl of pre-warmed APTT reagent (Helena Bioscience, Gateshead, UK) was then added and the mixture was incubated further for 3 min. Calcium chloride (25 mM, 50 µl) was added to the mixture and the clotting time was measured at an absorbance of 650 nm.

CHAPTER 3

Results

3.1 Anticoagulant activity of tick saliva

3.1.1 *Rhipicephalus pulchellus*

3.1.1.1 Protein quantification of crude salivary gland extracts

Salivary glands were dissected from male and female adult ticks and the protein contents were extracted from them (Section 2.1). The SGEs of both male and females ticks were reconstituted in distilled water and the protein amount present in each mixture was quantified with the Bradford Assay (Section 2.2.1). Table 5 summarizes the concentration and amount of protein that each sample contained. It is important to note that female SGEs contain approximately four times more protein than male SGEs, for each salivary gland pair.

Table 5. Protein quantification

Sample	Number of glands (pairs)	Concentration (mg/ml)	Amount (mg)	Amount per gland (µg)
Males	58	1.53	0.459	7.91
Females	30	2.95	0.885	29.50

3.1.1.2 Activity of crude salivary gland extracts

The crude extracts were evaluated for their anticoagulation potential using chromogenic enzyme inhibition assays (Section 2.2.2). Equivalent amounts of proteins (15 µg) from male and female ticks were assayed for their abilities to inhibit both FXa and thrombin. Figure 3.1 shows the percentage of inhibition each sample had towards the two enzymes. Female SGEs were able to inhibit both FXa and thrombin well, with 89% and 86% inhibition respectively.

On the other hand, male SGEs showed much lower inhibition (20% and 17%, respectively) towards both enzymes.

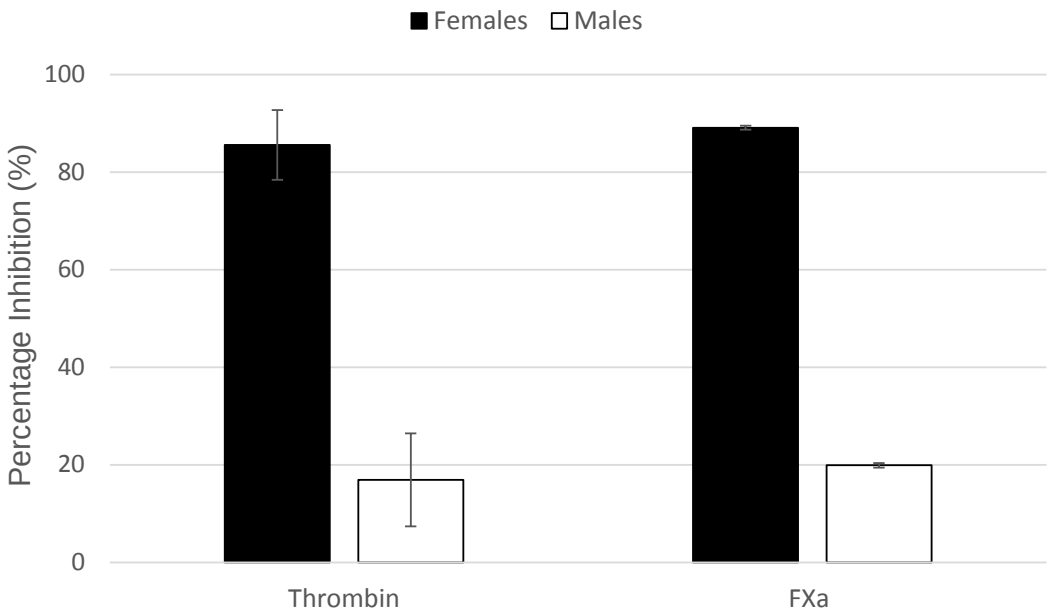

Figure 3.1. Anticoagulant activity of crude *R. pulchellus* salivary gland extracts. The amidolytic activities of thrombin and FXa were measured in the presence of 15 µg of female or male crude salivary gland extracts. Female extracts were able to inhibit both thrombin and FXa to higher than 80% while male extracts inhibited the enzymes lower than 20%.

3.1.1.3 Purification of salivary gland extracts

R. pulchellus SGEs (200 pairs) from each gender were fractionated by size-exclusion chromatography (Section 2.2.3.1), and each fraction was assayed for anti-FXa and anti-thrombin activities (Section 2.2.2). The protein elution profiles of both male and female SGEs are shown in Figure 3.2. Comparing both profiles and taking into account that the protein content for females is roughly four times that of males, it is evident that they are distinct. This indicates that their salivary compositions vary from each other. Inhibitor elution profiles of female SGEs show two major inhibitors of FXa (FXaI-1 and

FXaI-2), and one major inhibitor of thrombin (TI-1). Interestingly, in male SGEs, only a small peak of FXaI-2 was found; FXaI-1 and TI-1 were clearly absent. This corroborates the finding in Section 3.1.2 with poor inhibition of FXa and thrombin by male SGEs (Figure 3.1), indicating distinct differences in the composition of male and female *R. pulchellus* SGEs, particularly in their anticoagulant protein content.

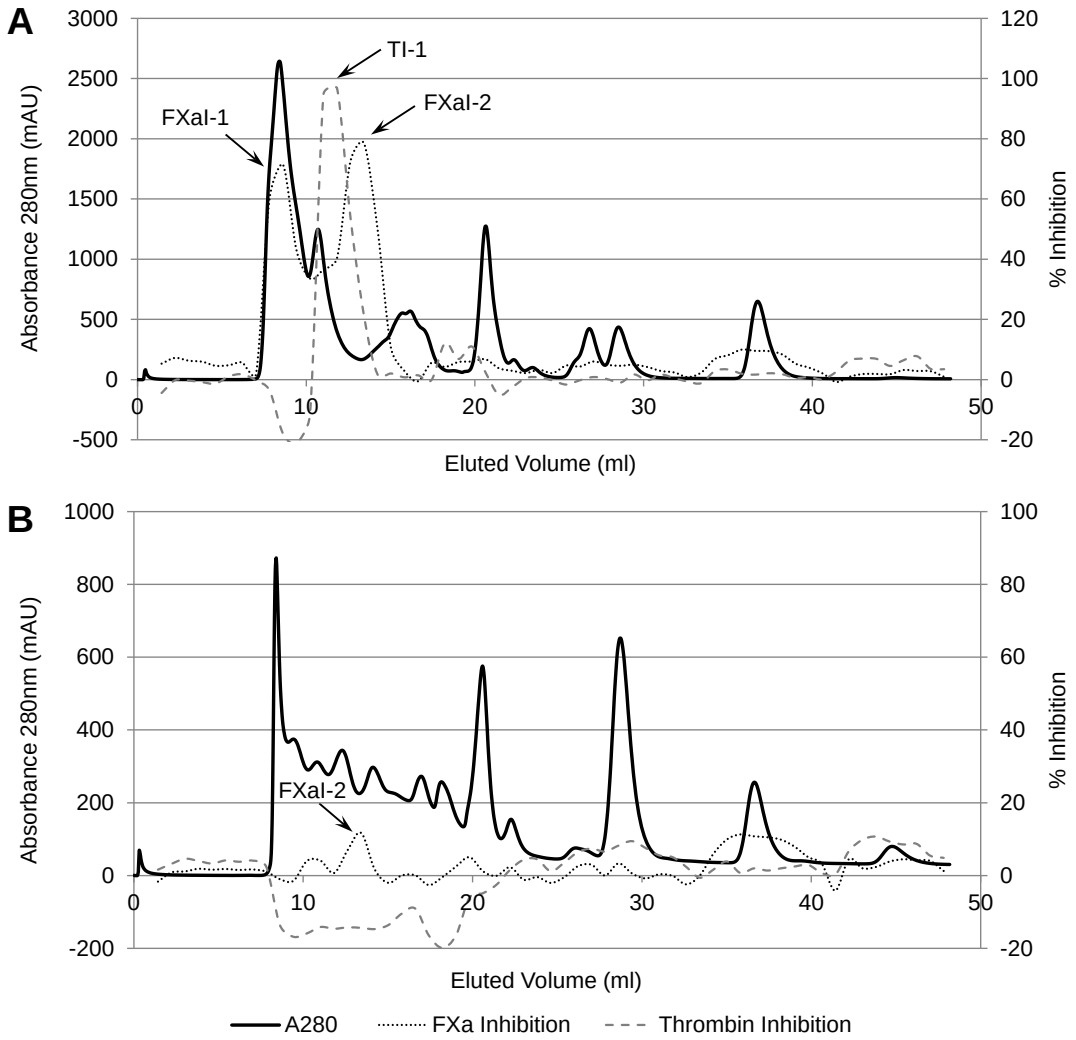

Figure 3.2. Anticoagulant profile of *R. pulchellus* SGEs. Crude salivary gland extracts from female **(A)** and male **(B)** ticks were subjected to fractionation by size exclusion chromatography. Fractions were tested for their ability to inhibit the amidolytic activity of FXa and thrombin. Female extracts evidently inhibited FXa (FXaI-1 and FXaI-2) and thrombin (TI-1) while male extracts only inhibited FXa to a small extent (FXaI-2).

3.1.2 *Dermacentor reticulatus*

3.1.2.3 Purification of salivary gland extracts

D. reticulatus SGEs (females 81 pairs; males 61 pairs) from each gender were fractionated by size-exclusion chromatography (Section 2.2.3.1), and each fraction was assayed for anti-FXa and anti-thrombin activities (Section 2.2.2). The protein elution profiles of both male and female SGEs are shown in Figure 3.3. Comparing both, it is evident that similar to that of *R. pulchellus*, they are distinct. This indicates that their salivary compositions vary from each other. Inhibitor elution profiles of female SGEs show four major inhibitors of FXa (DRFXaI-1, DRFXaI-2, DRFXaI-3 and DRFXaI-4). Interestingly, in male SGEs, only one inhibitor DRFXaI-2 was found. No distinct thrombin inhibitor was found in the saliva of this tick species.

Out of the four active regions from female *D. reticulatus,* the two most active ones inhibitors, DRFXaI-2 and DRFXaI-3 were selected for further purification via reversed phase LC with a microbore (Section 2.2.3.3). Purification of the DRFXaI-2 revealed that there were more than five active inhibitors (Figure 3.4A), while purification of DRFXaI-3 (purple) contained only one active peak (Figure 3.4B). Based on its elution on gel filtration column, the estimated molecular sizes of the FXa inhibitors are approximately 50 kDa (DRFXaI-2) and 10 kDa (DRFXaI-3) respectively.

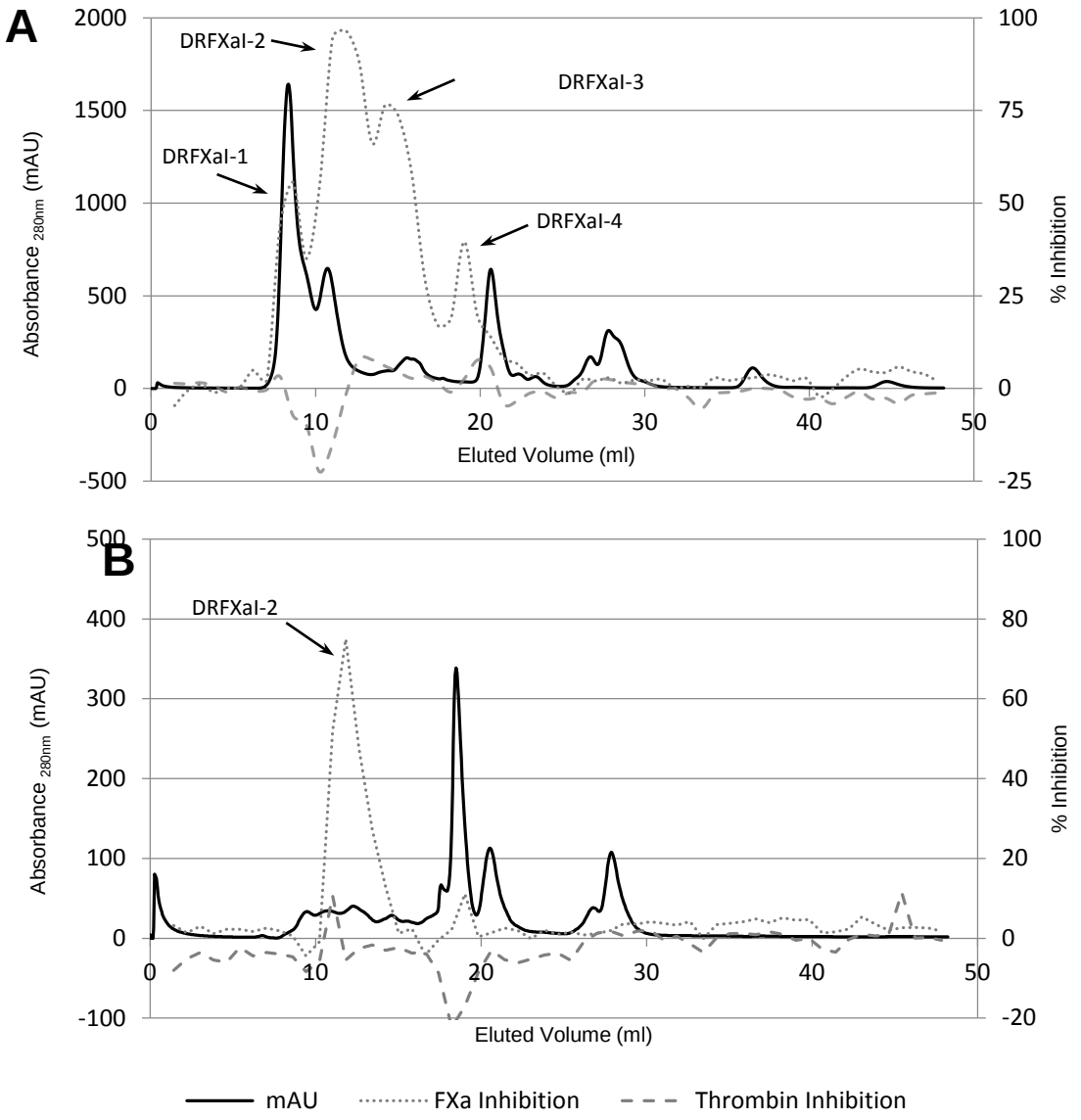

Figure 3.3. Anticoagulant profile of *D. reticulatus* SGEs. Crude salivary gland extracts from female **(A)** and male **(B)** ticks were subjected to fractionation by size exclusion chromatography. Fractions were tested for their ability to inhibit the amidolytic activity of FXa and thrombin. Female extracts evidently inhibited FXa (DRFXaI-1, DRFXaI-2, DRFXaI-3 and DRFXaI-4) while male extracts inhibited FXa to a smaller extent (FXaI-2).

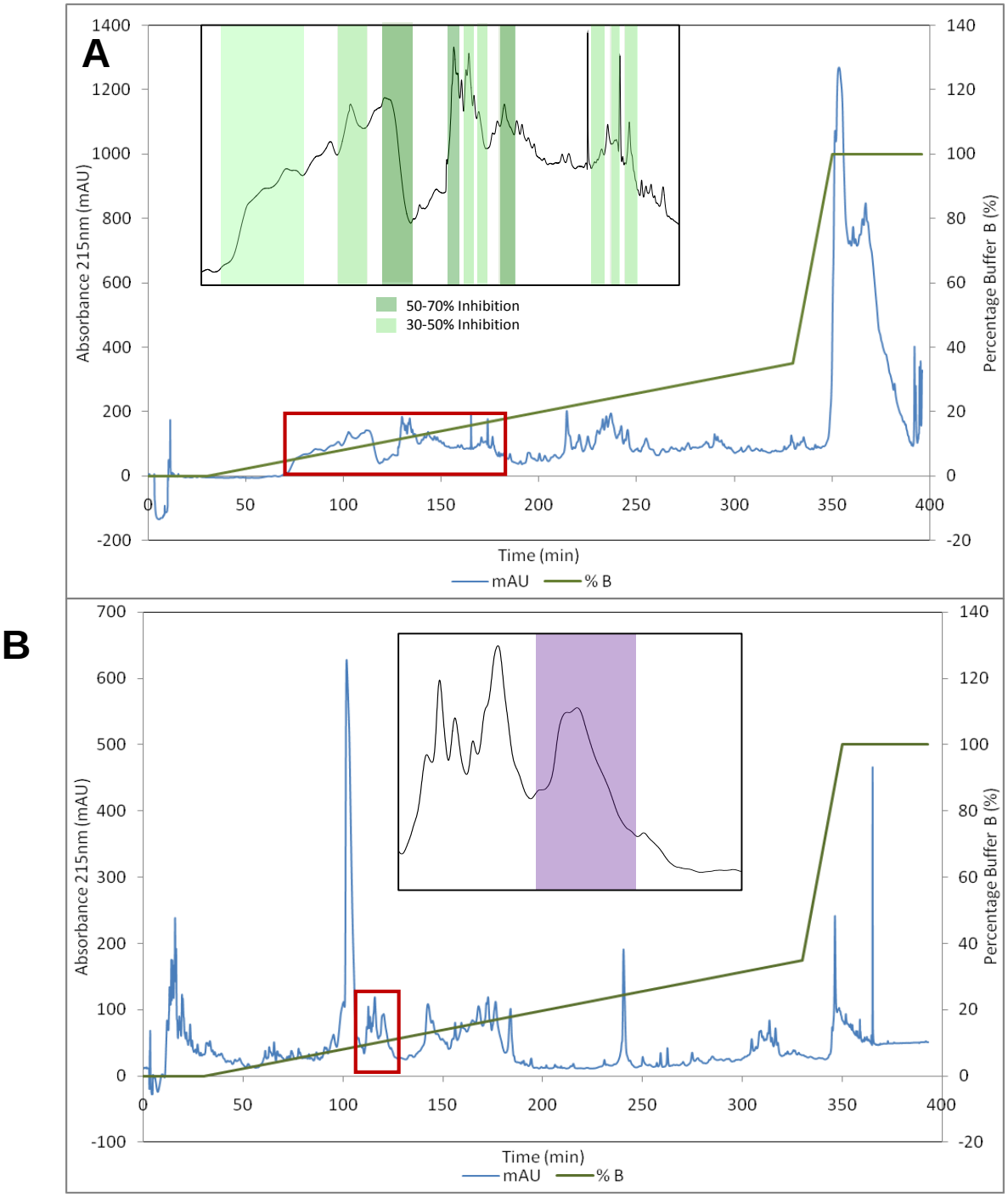

Figure 3.4. Reverse phase chromatography of FXa inhibitors from *D. reticulatus* on C18 column. Protein pool containing FXa inhibitory activity, DRFXaI-2 **(A)** and DRFXaI-3 **(B)** were loaded on a microbore C18 column. Proteins were eluted using a gradient of acetonitrile. Protein elution was monitored at 280 and 215 nm. Individual protein peaks were tested for FXa inhibitory activity. Inset shows the enlarged segment boxed in red. Fractions containing FXa inhibitory activity is highlighted in green (A) and purple (B) colour. Intense green colour indicates potent FXa inhibition.

An additional step of purification was carried out by generating a FXa-affinity column (Section 2.2.3.2) in order to pull down the FXa inhibitors. Fractions from the size-exclusion chromoatography step containing DRFXaI-3 were pooled. The mixture was then incubated in the column for 10 min before washing away the unbound. The bound FXa inhibitors were then eluted with high salt solution. At each step, a small sample was taken out to run the sample on a protein gel (Figure 3.5A), as well as to perform FXa inhibition assay (Figure 3.5B and 3.5C). Based on the FXa inhibition assay, comparing the activity between the samples before putting through the column and the unbound flowthrough, the the FXa inhibitors had successfully bound to the column. Most of the inhibitor had been eluted in the 1.0 M NaCl fraction.

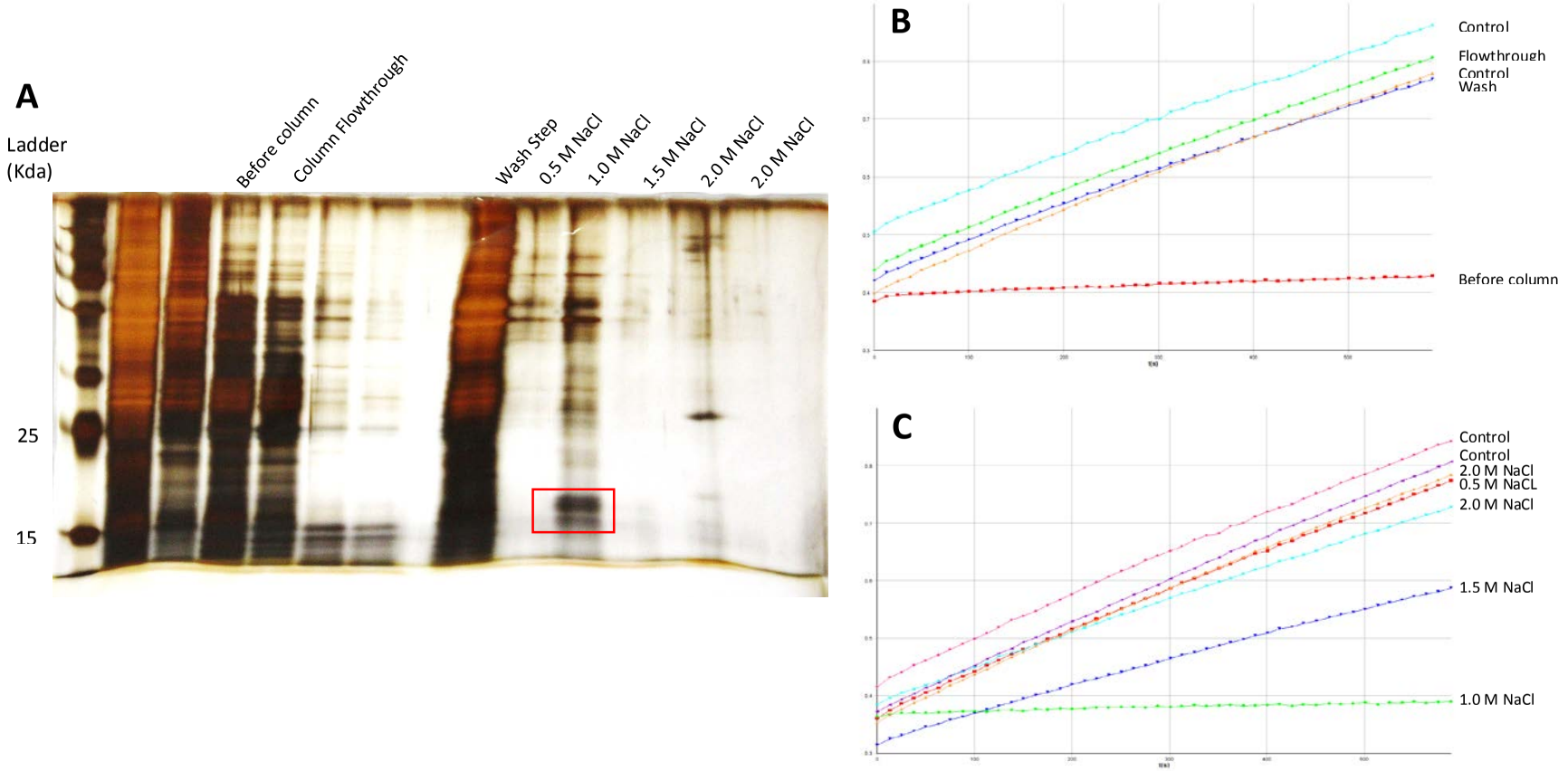

Figure 3.5. FXa-affinity column purification of DRFXaI-3. (A) Silver-stained protein electrophoresis gel of sample DRFXaI-3 before passing through the affinity column, unbound sample (column flowthrough), flowthrough during the wash step, and the elute with the various concentrations of NaCl. Protein band boxed up in red indicates the probable band from the FXa inhibitor. **(B)** and **(C)** FXa inhibition assay of the samples.

3.2 Sialome of *R. pulchellus*

3.2.1 Transcriptomes of *R. pulchellus*

As previously mentioned, anticoagulant proteins in the saliva of *R. pulchellus* are present in minute quantities. This makes their purification and identification challenging. Thus, another approach was used in which salivary anticoagulants could be examined. The sialotranscriptomes of the salivary glands of *R. pulchellus* were generated. In addition, to delineate the differences in salivary composition, male and female transcriptomes were generated independently. In order to represent the full repertoire of transcripts, ticks from different feeding time points were harvested for the generation of the library, and then pooled (refer to Section 2.1 for details). The mRNA sequences from the salivary glands of male and female ticks were determined using next generation Illumina sequencing (Section 2.3). A total of 50,460 coding sequences (CDS) were extracted from the combined assembly of 139,215,612 and 102,013,516 paired-end sequences (100 nt long) from male and female samples, respectively. The CDS were classified into four main categories – housekeeping proteins (H), putative secreted proteins (S), transposable elements (TE), and proteins of unknown function (U) (Table 6 and Figure 3.6).

Table 6: Functional classification of extracted coding sequences (CDS) from the sialotranscriptome of *R. pulchellus*.

Class	Number of CDS	Associated Reads	Percentage of Total Reads (%)
Housekeeping	11,499	18,001,992	25.56
Secreted	7,134	16,080,727	22.83
Transposable elements	2,195	2,989,302	4.24
Unknown	29,632	33,351,022	47.36
Total	**50,460**	**70,423,043**	**100**

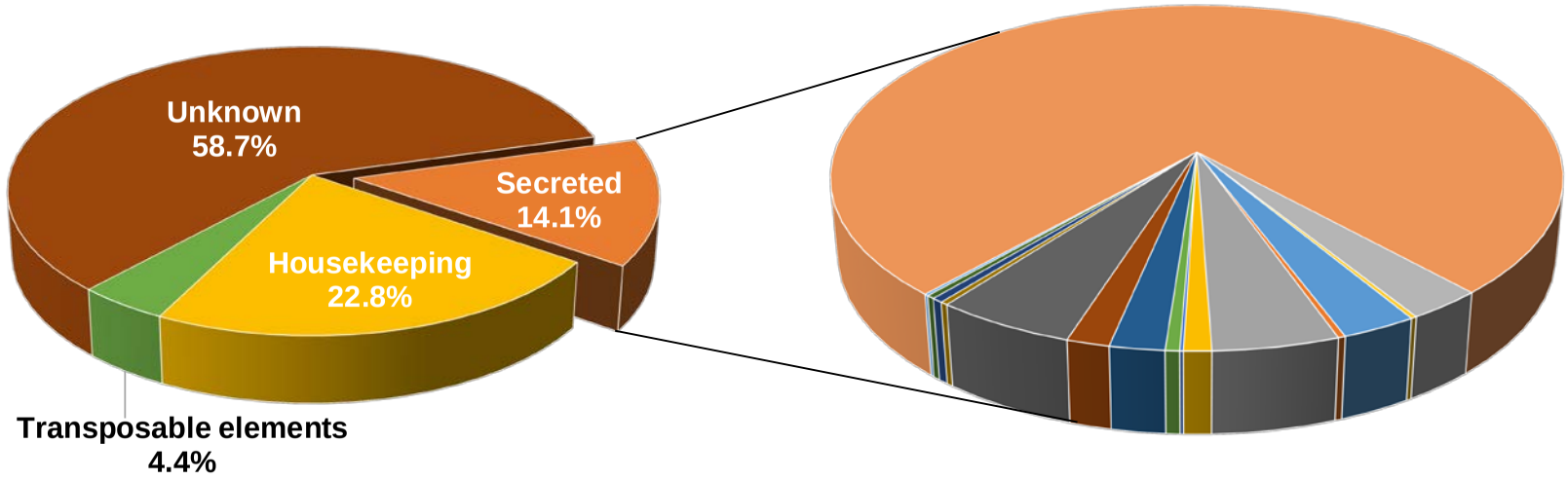

Figure 3.6. Components of *R. pulchellus* transcriptome. The transcriptome was classified into four categories – housekeeping, secreted, transposable elements and unknown. Secretory products are shown in more details, with further classification. The charts represent proportion of CDS in the transcriptome.

In summary, 11,499 CDS (25.6% reads) were associated with the H class, 7,134 CDS (22.8% reads) from the S class and 2,195 CDS (4.2% reads), including fragments, from the TE class. Finally, 47.4% of the reads were mapped to 29,632 putative CDS of unknown function and these CDS were not further analyzed. This class may contain secreted peptides. The proportion of S class transcripts was similar to a previously deep-sequenced sialotranscriptome of the *Amblyomma maculatum* tick (23.7%) (Karim et al., 2011). All CDS and their corresponding matches to several databases are available as a spreadsheet in the supplementary files DVD, with the file named "RP Sialotranscriptome". For each CDS, the numbers of reads derived from the male and female libraries, along with a chi-squared test for the significance of the differences in their read counts are tabulated.

3.2.1.1 Public sequence disclosure

Raw sequences were deposited on the Sequence Read Archive (SRA) from the NCBI under bioproject accession PRJNA170743. The individual reads received accession numbers SRR521835, SRR521944, SRR521951 and SRR521953. A total of 11,227 coding sequences and their translations were submitted to the Transcriptome Shotgun Assembly project deposited at DDBJ/EMBL/GenBank under the accessions GACK01000001-GACK01011227.

3.2.2 Proteomes of *R. pulchellus*

To confirm the presence of putative proteins from the transcriptome, we used shotgun proteomics to identify proteins in the SGEs of male and female adult *R. pulchellus*. The transcriptome was used as a database for the proteome. In total, 2,231 proteins were identified by one or more high confidence (>95%) peptides, of which 1,777 and 169 proteins were found exclusively in females and males respectively, and 285 proteins in both sexes. Out of the 2,231 proteins, 221 (9.9%) were of the S class, 1966 (88.1%) from the H class, 7 (0.3%) from the TE class and 37 (1.7%) from the U class (Table 7 and Figure 3.7). An excel sheet containing the list of peptides identified for each protein is with the DVD attachment, supplemental file “Peptide Summary”.

Table 7: Functional classification of proteins identified from the proteome of *R. pulchellus*.

Class	Number of Proteins	Percentage of Total (%)
Housekeeping	1966	88.1
Secreted	221	9.9
Transposable elements	7	0.3
Unknown	37	1.7
Total	**2231**	**100**

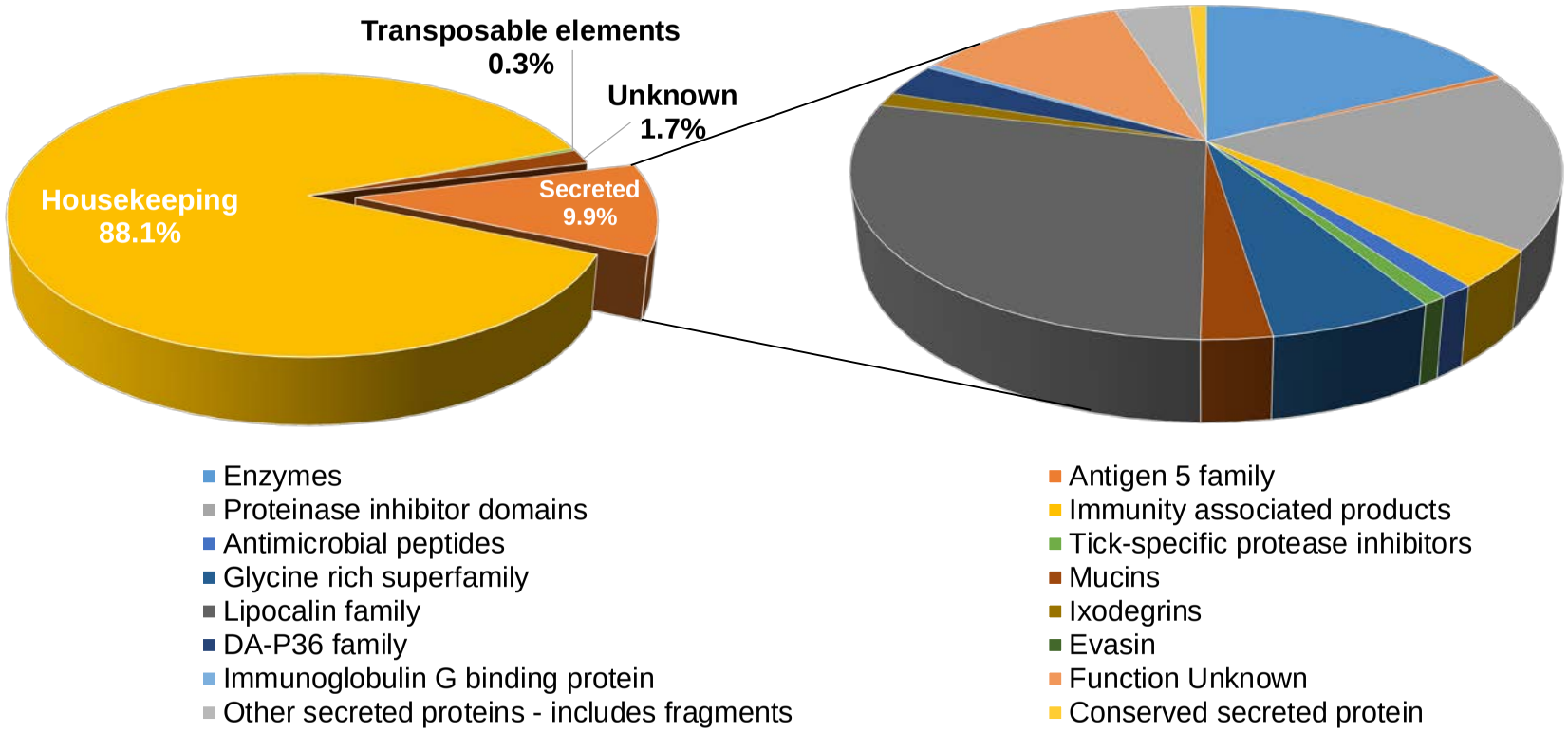

Figure 3.7. Components of *R. pulchellus* proteome. The proteome was classified into four categories – housekeeping, secreted, transposable elements and unknown. Secretory products are shown in more details, with further classification. The charts represent proportion of proteins in the proteome.

3.2.3 Housekeeping proteins

The 11,509 CDS associated with housekeeping function were further classified into 24 functional classes (Table 8 and 9). Protein synthesis machinery, transcription machinery and protein export machinery were among the top five classes, which were expected from an organ associated with a glandular function. The top two classes were signal transduction and hypothetical conserved proteins with unknown function (Galperin and Koonin, 2004). It is possible that some of these conserved proteins may belong to the other four classes, once we understand their function.

Several housekeeping proteins could have a role in the synthesis of small antagonists or agonists, or if secreted, they might play a direct role in saliva. Some enzymes involved in oxidative metabolism and detoxification, such as catalases, selenoproteins, peroxidases, thioredoxin peroxidases, could detoxify host oxidants associated with inflammation (Szabo et al., 2007). Cytochrome P450 enzymes could participate in the syntheses of prostaglandins known to be secreted in tick saliva (Dickinson et al., 1976; Higgs et al., 1976; Oliveira et al., 2011; Ribeiro et al., 1988; Ribeiro et al., 1992). Soluble epoxide hydrolases, if secreted, may affect host prostanoids (Marino, 2009; Spector et al., 2004). Some CDS of sphingomyelinases and deoxyribonucleases also show presence of signal peptides; such secreted enzymes could affect host immunity signaling (Utermohlen et al., 2008) and affect host neutrophil extracellular traps (NETs) (Wartha et al., 2007), respectively. Two of these deoxyribonucleases, RpIx75-684313 and RpIx75-682728, were found exclusively in females; 351 and 183 reads were found in

females as compared to only two and three reads in males, respectively. Lastly, sulfotransferases may inactivate dopamine, a secretagogue found in the salivary gland of ticks (Pichu et al., 2011).

Several transcripts encoding putative housekeeping proteins were found differentially expressed in males and females. Tyrosine aminotransferase and one 4-aminobutyrate transferase were overexpressed (>50-fold) in females. This was reflected in the proteome where tyrosine aminotransferase was only found in females, and 4-aminobutyrate transferase had 18 high confidence peptides in females but only two in males. This may be explained by the increased gluconeogenesis in females as they take larger blood meals than males. An α-fucosidase (RpIx75-667813) with signal peptide, indicative of secretion, could be a lysosomal enzyme, but may be actually secreted in saliva; it is 80-fold overexpressed in females. Males, on the other hand, have increased transcription (54-58-fold) for galactosyltransferases and α-1,4-N-acetylglucosaminyltransferase indicating higher amounts of glycoproteins or glycolipids which could be associated with the lubricating function of saliva during reproduction (Feldman-Muhsam et al., 1970). Twenty two transcripts coding for monocarboxylate transporters were overexpressed (52-404-fold) in females, while six ABC type transporters were overexpressed (169-382-fold) in males. Prolyl-4-hydroxylases, involved in post-translational modification of proline to hydroxyproline (Kivirikko and Pihlajaniemi, 1998), were found overexpressed (over 300-fold) in males. Males also have increased expression (100-fold) of the doublesex transcription factor (Baker and Wolfner, 1988), which is associated with male sex differentiation. Finally, males also have increased expression (270-450-fold) of dopamine β-monooxygenase, an

enzyme that catalyzes oxidation of dopamine to noradrenaline. This enzyme could be associated to the detoxification of dopamine, as in the case of the sulfotransferases mentioned above, or in modifying female products during fecundation, if secreted, and indeed it has a signal peptide indicative of secretion.

Table 8: Functional classification of extracted coding sequences (CDS) from the putative housekeeping class from the sialotranscriptome of *R. pulchellus*.

Class	No. of CDS	Associated Reads	% Total reads
Amino acid metabolism	140	223,972	1.24
Carbohydrate metabolism	304	403,332	2.24
Cytoskeletal proteins	316	413,798	2.30
Detoxification	77	59,983	0.33
Energy metabolism	212	1,026,836	5.70
Extracellular matrix	335	632,742	3.51
Intermediary metabolism	62	60,342	0.34
Lipid metabolism	446	502,702	2.79
Nuclear export	71	107,089	0.59
Nuclear regulation	424	415,586	2.31
Nucleotide Metabolism	172	122,319	0.68
Other immune related proteins	77	74,336	0.41
Oxidant metabolism/Detoxification	72	100,617	0.56
Proteasome machinery	442	585,484	3.25
Protein export	827	1,239,885	6.89
Protein modification	266	500,861	2.78
Protein modification/Protease	193	267,022	1.48
Protein synthesis machinery	467	1,637,732	9.10
Signal transduction	2,092	2,452,172	13.62
Storage	11	20,655	0.11
Transcription factor	210	252,611	1.40
Transcription machinery	1,319	1,749,731	9.72
Transporters and channels	721	1,068,327	5.93
Unknown conserved	2,243	4,083,858	22.69
Total	**11,499**	**18,001,992**	**100**

Table 9: Functional classification of proteins from the putative housekeeping class from the proteome of *R. pulchellus.*

Class	Number	% of total
Cytoskeletal	126	6.4
Extracellular matrix/cell adhesion	46	2.3
Immunity	12	0.6
Metabolism, amino acid	62	3.2
Metabolism, carbohydrate	89	4.5
Metabolism, energy	63	3.2
Metabolism, intermediate	29	1.5
Metabolism, lipid	81	4.1
Metabolism, nucleotide	45	2.3
Nuclear export	18	0.9
Nuclear regulation	89	4.5
Oxidant metabolism/detoxification	39	2.0
Proteasome machinery	74	3.8
Protein export machinery	134	6.8
Protein modification machinery	197	10.0
Protein synthesis machinery	131	6.7
Signal transduction	243	12.4
Signal transduction, apoptosis	42	2.1
Storage	5	0.3
Transcription factor	20	1.0
Transcription machinery	167	8.5
Transporters/storage	57	2.9
Unknown, conserved	197	10.0
Total	**1966**	**100**

3.2.4 Putative secreted proteins

The putative secreted proteins were classified into various families following a previous review (Francischetti et al., 2009a) and recent transcriptomes (Anatriello et al., 2010; Francischetti et al., 2011; Karim et al., 2011). Accordingly, 121 different families of coding sequences were identified (Table 10 and 11). They are divided into two general groups, ubiquitous protein families and tick-specific families. The ubiquitous protein families include, enzymes such as metalloproteases, serine proteases, carboxypeptidases, dipeptidyl peptidases, lipases, pyrophosphatases and apyrases, and non-enzymatic proteins such as members of the antigen 5 protein family, and proteinase inhibitors of the Kunitz, Kazal, serpin, cystatin, thyropin and trypsin inhibitor-like (TIL) domain families. They are found in most sialotranscriptomes as well as in the venoms of vespids and snakes. In addition, immunomodulatory proteins, such as lysozyme, defensins, MD-2-related lipid-recognition (ML) domain peptides, ixoderin, peptidoglycan recognition proteins and thioester proteins (TEP), are secreted by ticks to counter the host immune system. The tick-specific protein families include glycine-rich proteins (serve as tick cement), mucins (lubrication), evasins (bind host cytokines), DA-p36 (immunosuppressor), ixodegrins (potential anti-platelet agents) and tick-specific protease inhibitors such as carboxypeptidase inhibitors. In addition, members of lipocalin and the basic tail families were also found. Lipocalins bind to histamine, serotonin or inflammatory prostanoids. Seven additional groups of proteins, many representing multiple families, are described that have no known function, of which most were found uniquely in *R. pulchellus*.

Table 10. Functional classification of extracted coding sequences (CDS) from the putative secreted class from the sialotranscriptome of adult *Rhipicephalus pulchellus* ticks.

Class	CDS	Associated Reads	% Total reads
Enzymes			
Metalloproteases			
Typical tick ZnMc_salivary_gland_MPs CDD motif	98	180,020	1.120
Large ADAM's proteases - probably housekeeping	5	4,059	0.025
Other Zn metalloproteases, including collagenases	7	6,128	0.038
M13 family peptidase	14	25,327	0.158
Dipeptidyl peptidase	2	13,064	0.081
Serine protease	17	12,921	0.080
Male specific salivary serine protease	2	32,135	0.200
Suchi serine protease	2	3,421	0.021
Zinc carboxypeptidase	2	9,147	0.057
Serine carboxypeptidase - may be lysosomal	19	43,212	0.269
Legumain/asparaginyl peptidase	2	6,916	0.043
Chitinase	2	2,106	0.013
5'-nucleotidase/Apyrase	20	28,812	0.179
Ribonuclease	4	4,892	0.030
Ectonucleotide pyrophosphatase/phosphodiesterase family	3	5,489	0.034
Lipase	4	2,591	0.016
Inositol phosphatase	4	3,734	0.023
Antigen 5 family	20	65,252	0.406
Proteinase inhibitor domains			
Serpins	23	10,468	0.065
Monolaris	81	125,953	0.783
Bilaris	65	94,581	0.588
Trilaris	33	37,042	0.230

Class	CDS	Associated Reads	% Total reads
Tetralaris	6	32,425	0.202
Pentalaris	4	44,321	0.276
Hexalaris	1	368	0.002
Heptalaris	2	4,770	0.030
Similar to Kunitz domain	10	10,995	0.068
Cystatins	20	73,488	0.457
Thyropins with two domains	4	2,786	0.017
IGF/KAZAL/IG domain family	2	2,019	0.013
Chymotrypsin-elastase inhibitor ixodidin - TIL domain	33	636,225	3.956
MonoTil	13	26,055	0.162
BiTil	35	96,191	0.598
TriTil	12	50,388	0.313
PolyTil	2	12,311	0.077
Til-Like	1	31	0.000
Immunity associated products			
Lysozyme	4	2,397	0.015
Defensins	22	5,555	0.035
TEP proteins	23	50,188	0.312
ML domaini	9	1,533	0.010
Galectin	6	7,046	0.044
Peptidoglycan recognition protein	6	7,098	0.044
Ixoderin/Ficolin	6	17,076	0.106
Tick specific protein families, at least one family member has known function			
Antimicrobial peptides			
Microplusin	6	14,525	0.090
5.3 kDa antimicrobial family	4	317	0.002
Protease inhibitors			
Carboxypeptidase inhibitor	13	17,373	0.108
Basic Tail	12	18,408	0.115

Class	CDS	Associated Reads	% Total reads
18.3 kDa family	14	12,823	0.080
Glycine rich superfamily	151	7,640,865	47.516
Mucins	118	1,840,933	11.448
Lipocalin family			
Group I	209	309,678	1.926
Group II	20	34,862	0.217
Group III	13	45,186	0.281
Group IV	10	24,295	0.151
Group V	16	25,277	0.157
Group VI	13	176,753	1.099
Group VII	17	17,945	0.112
Group VIII - Deorphanized Dermacentor lipocalin family	19	20,816	0.129
New lipocalin family Rhipicephalus family XII	5	7,271	0.045
Other putative lipocalins	64	912,901	5.677
Ixodegrins	21	45,029	0.280
DA-P36 family	31	36,142	0.225
Evasin	22	19,746	0.123
Immunoglobulin G binding protein	12	1,382,686	8.598
Function Unknown			
Found in prostriates and metastriates			
Fibronectin domain containing protein family	2	157	0.001
Hematopoietic-stem cell progenitor	2	1,551	0.010
8.9 kDa family	60	114,278	0.711
23 kDa family	7	3,072	0.019
24 kDa family	9	3,193	0.020
One of each protein family - now Ixodidae	14	3,505	0.022
Previously thought as metastriate specific			
Ig domain containing secreted peptide	4	1,814	0.011
Deorphanized Ixodidae Rhip 45-937	4	1,901	0.012

Class	CDS	Associated Reads	% Total reads
Insulin growth factor binding protein - also found in Ixodes	6	2,065	0.013
Amblyomma maculatum family 40-33	9	2,599	0.016
Deorphanized metastriate-Rhipicephalus family XIV	4	912	0.006
Deorphanized metastriate family 40-80	1	626	0.004
Somatomedin domain protein family	4	599	0.004
Metastriate specific			
28 kDa Metastriate family	18	17,340	0.108
16 kDa family	2	8,950	0.056
8 kDa Amblyomma family	5	2,135	0.013
10 kDa acidic metastriate	1	6,626	0.041
Ixostatin-like	10	12,781	0.080
Deorphanized Dermacentor 9 kDa expansion	5	1,102	0.007
Rhip 45-304 - deorphanized on A. maculatum sialome	4	13,864	0.086
Rhipicephalus specific families			
Rhip 45-141	17	2,401	0.015
Rhip 45-151	15	3,199	0.020
Rhip 45-236	2	22,601	0.141
Short protein families possibly secreted			
Rhip 50-70	6	1,431	0.009
Rhip 50-10	107	16,133	0.100
Rhip 50-45	26	21,670	0.135
Rhip 50-18	79	10,396	0.065
Rhip 45-25	64	21,119	0.131
Rhip 45-85	16	1,445	0.009
Rhip 40-92	23	2,214	0.014
Rhip 45-2	623	179,776	1.118
Rhip 45-10	43	15,020	0.093
Rhip 45-11	162	36,474	0.227

Class	CDS	Associated Reads	% Total reads
Rhip 45-16	105	21,708	0.135
Rhip 45-22	56	13,453	0.084
Rhip 45-27	62	13,982	0.087
Rhip 45-28	53	41,876	0.260
Rhip 45-29	56	11,234	0.070
Rhip 45-37	43	2,067	0.013
Rhip 45-48	37	7,106	0.044
Rhip 45-49	35	4,196	0.026
Rhip 45-56	33	8,522	0.053
Rhip 45-61	30	52,808	0.328
Rhip 45-66	28	2,827	0.018
Rhip 45-69	27	11,930	0.074
Rhip 45-72	26	4,828	0.030
Rhip 45-73	26	2,433	0.015
Rhip 45-77	25	11,873	0.074
Rhip 45-82	24	2,064	0.013
Rhip 45-95	22	5,455	0.034
Rhip 45-130	18	2,533	0.016
Rhip 45-132	18	771	0.005
Rhip 45-133	18	1,524	0.010
Rhip 45-143	17	3,894	0.024
Rhip 45-144	17	3,781	0.024
Remaining peptides of large small family 40-1	3,340	619,033	3.850
Other secreted proteins - includes fragments	208	272,152	1.692
Conserved secreted protein	16	21,361	0.133
Total	**7,134**	**16,080,727**	**100**

Table 11. Functional classification of extracted coding sequences (CDS) from the putative secreted class from the proteome of adult *Rhipicephalus pulchellus* ticks.

Class	Number	% of total
Enzymes		
Metalloproteases	6	2.71
M13 family peptidase	2	0.90
Dipeptidyl peptidase	1	0.45
Serine protease	3	1.36
Male specific salivary serine protease	2	0.90
Zinc carboxypeptidase	1	0.45
Serine carboxypeptidase - may be lysosomal	10	4.52
Legumain/asparaginyl peptidase	2	0.90
5'-nucleotidase/Apyrase	8	3.62
Ribonuclease	2	0.90
Inositol phosphatase	2	0.90
Antigen 5 family	1	0.45
Proteinase inhibitor domains		
Serpins	7	3.17
Monolaris	2	0.90
Bilaris	5	2.26
Tetralaris	1	0.45
Similar to Kunitz domain	1	0.45
Cystatins	2	0.90
Chymotrypsin-elastase inhibitor ixodidin - TIL domain	1	0.45
BiTil	11	4.98
TriTil	8	3.62
Immunity associated products		
Defensins	1	0.45
TEP proteins	4	1.81
Galectin	3	1.36

Tick specific protein families, at least one family member has known function		
Antimicrobial peptides		
Microplusin	3	1.36
Protease inhibitors		
Carboxypeptidase inhibitor	1	0.45
Basic Tail	1	0.45
Glycine rich superfamily	14	6.33
Mucins	6	2.71
Lipocalin family		
Group I	34	15.38
Group II	3	1.36
Group III	4	1.81
Group IV	1	0.45
Group V	3	1.36
Group VI	2	0.90
Group VIII - Deorphanized Dermacentor lipocalin family	1	0.45
New lipocalin family Rhipicephalus family XII	2	0.90
Other putative lipocalins	13	5.88
Ixodegrins	3	1.36
DA-P36 family	7	3.17
Immunoglobulin G binding protein	1	0.45
Function Unknown		
Found in prostriates and metastriates		
Hematopoietic-stem cell progenitor	1	0.45
8.9 kDa family	7	3.17
24 kDa family	1	0.45
One of each protein family - now Ixodidae	4	1.81
Metastriate specific		
28 kDa Metastriate family	4	1.81
10 kDa acidic metastriate	1	0.45
Ixostatin-like	2	0.90

Rhip 45-304 - deorphanized on A. maculatum sialome	1	0.45
Rhipicephalus specific families		
Rhip 45-236	2	0.90
Short protein families possibly secreted	2	0.90
Other secreted proteins - includes fragments	9	4.07
Conserved secreted protein	2	0.90
Total	**221**	**100**

We have also extracted 5,265 CDS that represent putative small secreted peptides varying from 40-100 amino acid residues, grouped into 33 families according to their similarities. Except for matches to *R. sanguineus* peptides, no similarities are apparent to other known proteins. Although it is unclear whether they are artefacts (extracted from non-coding RNA) or real, they serve as targets for mass spectrometry experiments. Indeed a few peptides were identified by MS, such as RpSigp-496390 and RpSigp-663980, supporting existence of some of them in tick SGEs.

It is remarkable that 48% of S class reads map to glycine-rich proteins, which assist the tick to attach to their host (Table 10). Together with mucins (11% reads), lipocalin family (10% reads) and immunoglobulin G binding proteins (9% reads), they account for 78% of the S class reads. It was interesting to see the breakdown of these top classes, in terms of the proportion of reads from males and females. Figure 3.8 illustrates the contribution of each gender towards the total number of reads. Note that it was arranged in increasing order with relation to the total number of reads. The glycine-rich family, mucins, and immunoglobulin G binding proteins had more than two-fold the number of reads in males than in females. On the other hand, the lipocalin family had more than two-fold more reads in females as compared to males. As glycine-rich proteins and mucins do not produce suitable tryptic peptides, as mentioned above, few were identified in the proteome. The top classes of proteins that were identified in the proteomes were the lipocalin family (28%), proteases (12%) and the proteinase inhibitor domains (17%) (Table 11).

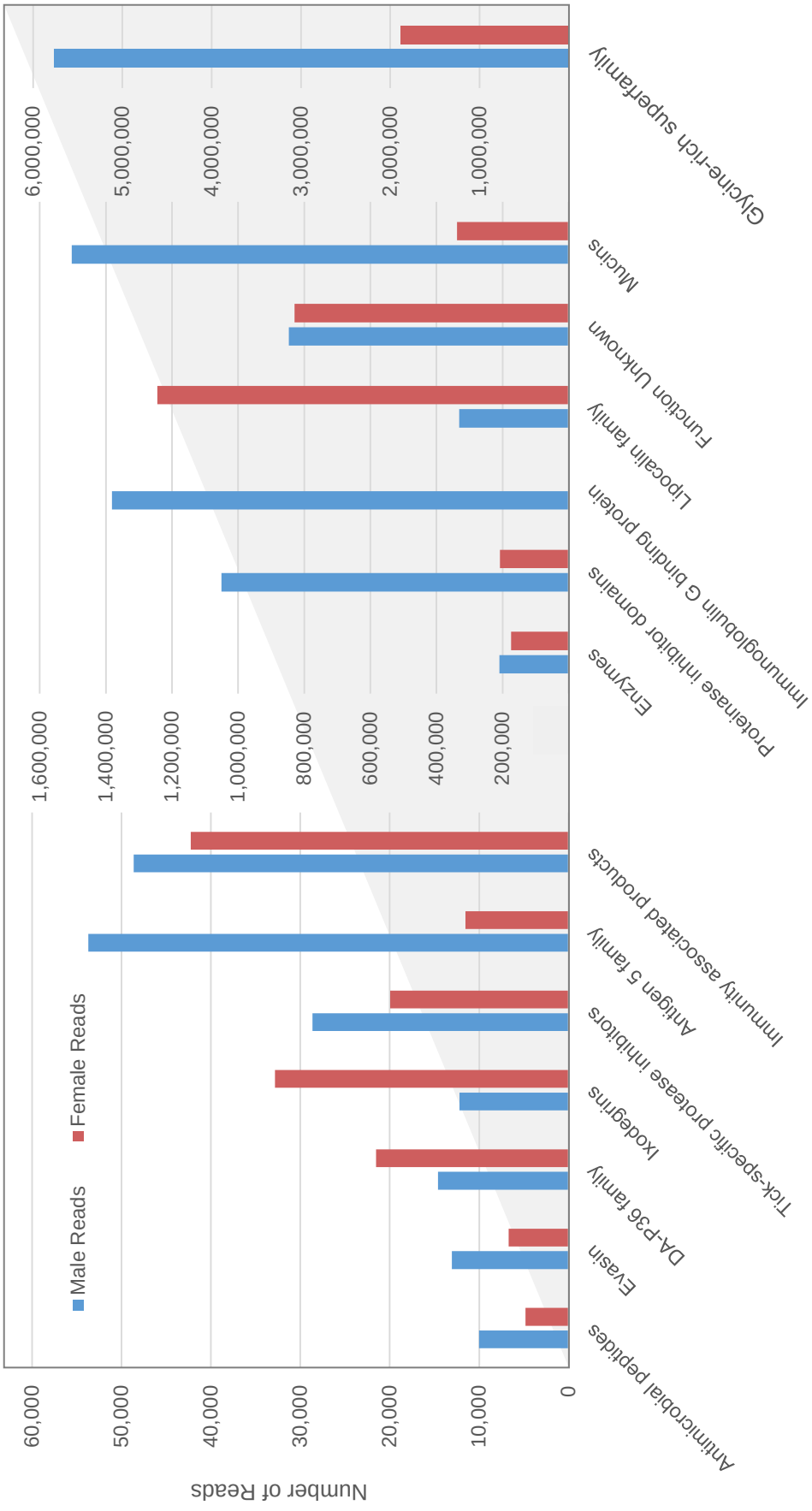

Figure 3.8 Differential expression of secretory proteins. The number of reads from male and female R. pulchellus transcriptome belonging to the secretory class in increasing order.

As noticed in previous sialomes of blood sucking arthropods, unique protein families were found at the genus or even at the species level (Francischetti et al., 2009a). This may be due to the host immune pressure which accelerates evolution of salivary gland genes. Previous families thought to be exclusively metastriate were found to have homologs in *I. scapularis*, albeit with a relatively low similarity, indicative of fast divergence of these salivary proteins (Francischetti et al., 2009a).

In this section, proteins from the various classifications will be reported. In addition, differences seen between the male and female sialome will be highlighted. Tables 12 and 13 reports the list of CDS and proteins found in the transcriptome and proteome, respectively, that were found to be gender-biased. For the transcriptome (Table 12), this is based on the normalized number of reads for a particular CDS, and gender-biasness is defined to be where the number of reads is 10 times more in one gender. Table 13 breaks down the results of the secretory proteins found in the proteome to report the number of proteins that were found in both gender and also the number that were found exclusively in either gender. To complement this, Figures 3.9 and 3.10 illustrates the various classes of secretory proteins that were gender-biased in each sex for both the transcriptome and proteome, respectively.

Table 12: Number of gender-biased CDS from the secretory class found in the transcriptome

Class	Females	Males
Enzymes		
Metalloproteases	18	5
Serine protease	0	6
Male specific salivary serine protease	0	2
Legumain/asparaginyl peptidase	0	1
5'-nucleotidase/Apyrase	0	2
Ribonuclease	2	0
Ectonucleotide pyrophosphatase/phosphodiesterase family	0	1
Lipase	1	0
Antigen 5 family	0	3
Proteinase inhibitor domains		
Serpins	1	4
Monolaris	17	3
Bilaris	7	1
Trilaris	1	0
Tetralaris	1	0
Similar to Kunitz domain	4	1
Cystatins	1	11
Chymotrypsin-elastase inhibitor ixodidin - TIL domain	9	11
MonoTil	0	1
BiTil	2	0
Immunity associated products		
Lysozyme	1	1
Defensins	8	1
TEP proteins	3	0
Ixoderin/Ficolin	1	0

Antimicrobial peptides		
Microplusin	2	0
5.3 kDa antimicrobial family	1	0
Protease inhibitors		
Carboxypeptidase inhibitor	1	0
Basic Tail	2	1
18.3 kDa family	2	0
Glycine rich superfamily	19	2
Mucins	8	5
Lipocalin family		
Group I	62	15
Group II	7	1
Group III	9	0
Group IV	7	0
Group V	5	2
Group VI	4	1
Group VII	3	0
Group VIII - Deorphanized Dermacentor lipocalin family	9	1
New lipocalin family Rhipicephalus family XII	0	0
Other putative lipocalins	20	6
Ixodegrins	8	1
DA-P36 family	15	0
Evasin	8	2
Immunoglobulin G binding protein	0	8
Function Unknown		
Found in prostriates and metastriates		
8.9 kDa family	22	6
24 kDa family	3	0
One of each protein family - now Ixodidae	11	0
Previously thought as metastriate specific		
Insulin growth factor binding protein - also found in Ixodes	3	0

Amblyomma maculatum family 40-33	2	0
Deorphanized metastriate-Rhipicephalus family XIV	0	1
Somatomedin domain protein family	2	0
Metastriate specific		
28 kDa Metastriate family	3	0
8 kDa Amblyomma family	1	0
Ixostatin-like	2	1
Deorphanized Dermacentor 9 kDa expansion	2	0
Rhip 45-304 - deorphanized on A. maculatum sialome	4	0
Short protein families possibly secreted	64	25
Other secreted proteins - includes fragments	25	10
Conserved secreted protein	3	0
Total	**416**	**142**

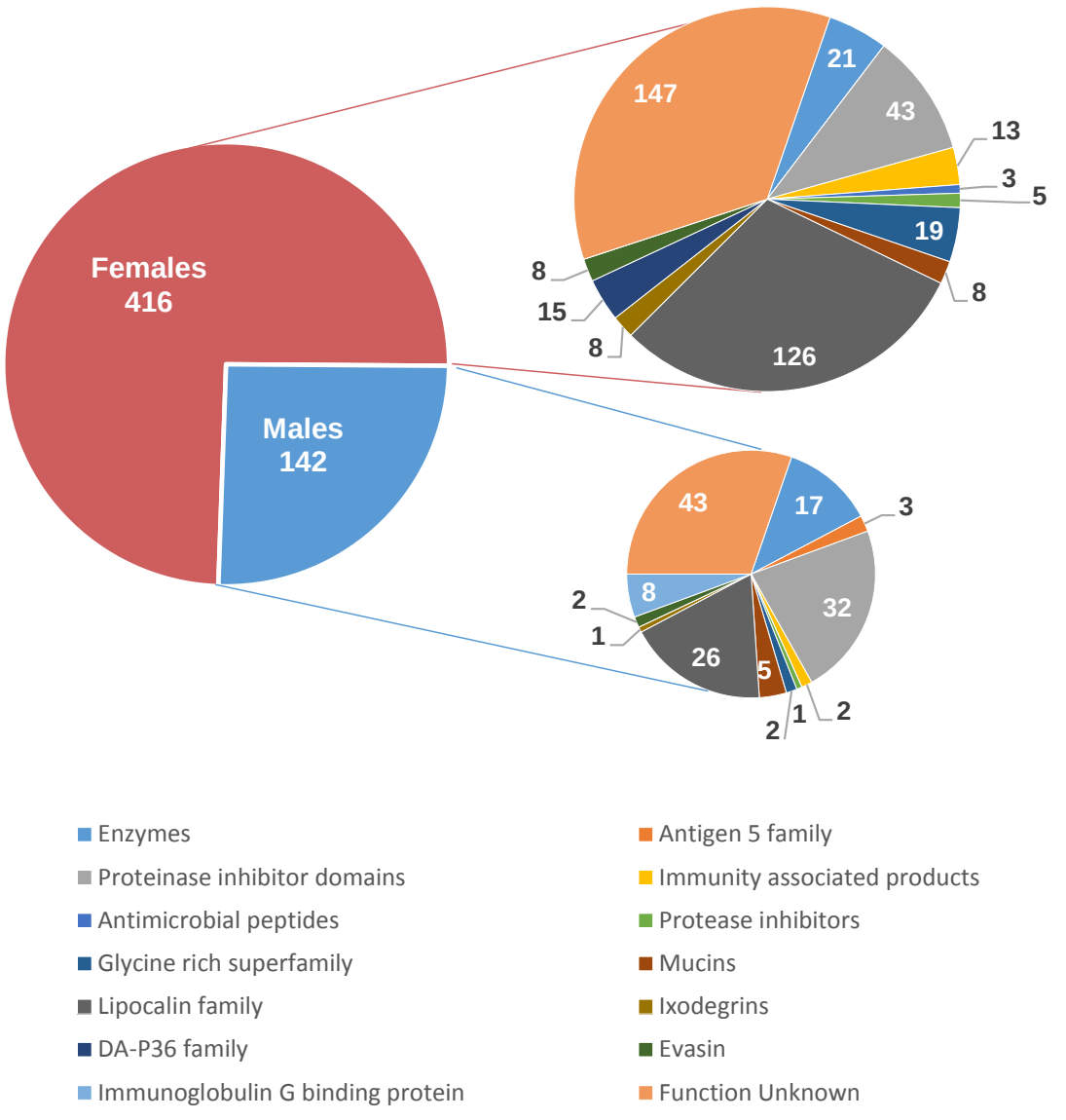

Figure 3.9. Classification of gender-biased CDS in *R. pulchellus* transcriptome. CDS which were 10 times more expressed in one gender was considered to be gender-biased. In total, 416 female CDS and 142 male CDS were gender-biased.

Table 13: Number of gender-biased proteins from the secretory class found in the proteome

Class	Found in both sexes		Found in males only		Found in females only	
	Number	% of total	Number	% of total	Number	% of total
Enzymes						
Metalloproteases	1	3.23	1	1.30	4	3.54
M13 family peptidase	2	6.45	0	0.00	0	0.00
Dipeptidyl peptidase	0	0.00	0	0.00	1	0.88
Serine protease	0	0.00	2	2.60	1	0.88
Male specific salivary serine protease	0	0.00	2	2.60	0	0.00
Zinc carboxypeptidase	0	0.00	0	0.00	1	0.88
Serine carboxypeptidase - may be lysosomal	0	0.00	9	11.69	1	0.88
Legumain/asparaginyl peptidase	0	0.00	1	1.30	1	0.88
5'-nucleotidase/Apyrase	2	6.45	2	2.60	4	3.54
Ribonuclease	0	0.00	0	0.00	2	1.77
Inositol phosphatase	0	0.00	0	0.00	2	1.77
Antigen 5 family	0	0.00	1	1.30	0	0.00
Proteinase inhibitor domains						

Serpins	1	3.23	5	6.49	1	0.88
Monolaris	0	0.00	1	1.30	1	0.88
Bilaris	0	0.00	0	0.00	5	4.42
Tetralaris	0	0.00	0	0.00	1	0.88
Similar to Kunitz domain	0	0.00	1	1.30	0	0.00
Cystatins	0	0.00	2	2.60	0	0.00
Chymotrypsin-elastase inhibitor ixodidin - TIL domain	0	0.00	1	1.30	0	0.00
BiTil	2	6.45	5	6.49	4	3.54
TriTil	0	0.00	8	10.39	0	0.00
Immunity associated products						
Defensins	0	0.00	1	1.30	0	0.00
TEP proteins	0	0.00	0	0.00	4	3.54
Galectin	0	0.00	0	0.00	3	2.65
Antimicrobial peptides						
Microplusin	0	0.00	1	1.30	2	1.77
Protease inhibitors						
Carboxypeptidase inhibitor	0	0.00	0	0.00	1	0.88
Basic Tail	0	0.00	0	0.00	1	0.88

Glycine rich superfamily	3	9.68	2	2.60	9	7.96
Mucins	0	0.00	4	5.19	2	1.77
Lipocalin family						
Group I	7	22.58	13	16.88	14	12.39
Group II	1	3.23	0	0.00	2	1.77
Group III	0	0.00	0	0.00	4	3.54
Group IV	0	0.00	0	0.00	1	0.88
Group V	0	0.00	0	0.00	3	2.65
Group VI	1	3.23	0	0.00	1	0.88
Group VIII - Deorphanized Dermacentor lipocalin family	0	0.00	0	0.00	1	0.88
New lipocalin family Rhipicephalus family XII	0	0.00	2	2.60	0	0.00
Other putative lipocalins	4	12.90	4	5.19	5	4.42
Ixodegrins	0	0.00	0	0.00	3	2.65
DA-P36 family	3	9.68	1	1.30	3	2.65
Immunoglobulin G binding protein	0	0.00	1	1.30	0	0.00
Function Unknown						
Found in prostriates and metastriates						
Hematopoietic-stem cell progenitor	0	0.00	0	0.00	1	0.88
8.9 kDa family	0	0.00	4	5.19	3	2.65

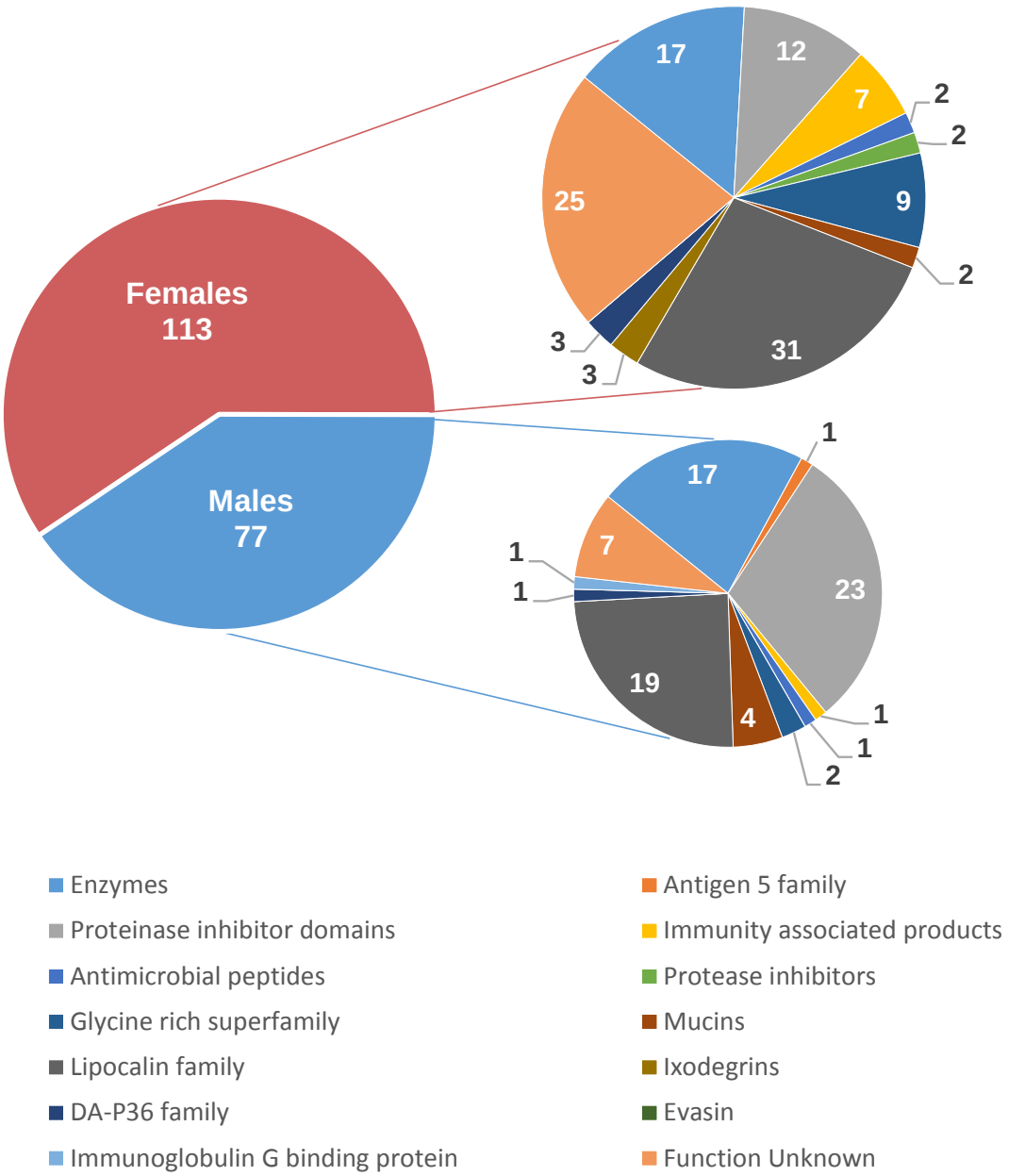

Figure 3.10. Classification of gender-biased proteins in *R. pulchellus* proteome. Proteins which were identified by MS only in one gender was considered to be gender-biased. In total, 113 proteins in females and 142 proteins in males were gender-biased.

3.2.4.1 Enzymes

Proteases were conspicuously represented in tick sialomes, including metalloproteases of the reprolysin family, M13 type peptidases, serine proteases, asparaginyl proteases and dipeptidyl peptidases.

The zinc metalloproteases, ubiquitous enzymes found in metazoans, are associated with embryonic development, inflammation, angiogenesis and connective tissue remodeling, among other functions (see (Klein and Bischoff, 2011; Wolfsberg et al., 1995)). Tick salivary metalloproteases distinctly have a Reprolysin_2 PFAM domain, typical of snake venom enzymes, as well as the ZnMc_salivary_gland motif from the Conserved Domain Database (CDD). These enzymes are important components of snake venoms where they cause hemorrhage in their prey (Hati et al., 1999). In tick saliva, they were found to be associated with disrupting the hemostatic system of their hosts via fibrinolysis (Francischetti et al., 2003) and angiogenesis inhibition (Francischetti et al., 2005a). Although the genome of *Ixodes scapularis* contains only six such proteins with more than 450 amino acid residues (and some more fragments), a previous deep transcriptome analysis of *Amblyomma maculatum* retrieved 47 near full-length sequences of this family (Karim et al., 2011). Similarly, the sialotranscriptome of *R. pulchellus* allowed the extraction of 98 sequences coding for typical tick salivary metalloproteases, 51 of which are full length or near full length. Several *R. pulchellus* metalloproteases appeared distinctly sex biased. For example, RpIx75-57164, RpSigp-931308, RpIx75-721814 and RpIx75-240814 were overexpressed in females (>100-fold), while RpIx75-919895 are overexpressed in males (57-fold). In accordance with its transcript

level bias among sexes, RpIx75-57164 was detected by proteomics in females with 29% coverage but was not detected in males. Proteomic analysis also revealed the product RpIx75-951936 to be highly expressed, with over 50% coverage in females and 21% coverage in males.

M13 domain/neprilysin proteases are involved in the inactivation of hormone peptides. In ticks, they could function by destroying inflammatory peptidic mediators such as cytokines, anaphylatoxins or bradykinin from the hosts. While most members of this family are extracellular, some are membrane bound via a transmembrane helix. Ten *R. pulchellus* M13 proteases were found without this transmembrane helix indicating they are secreted, similar to *A. maculatum* homologs, which are 50-70% identical to the *R. pulchellus* sequences. RpIx75-695314 and RpIx75-901875 were identified in both male and female proteomes.

Serine proteases are ubiquitous and commonly found in sialotranscriptomes of blood feeding ticks and insects where they may function as a fibrinolysin (Xu et al., 2008). The sialotranscriptome of *R. pulchellus* reveals 19 CDS for these proteins. Remarkably, several of these appear to be male-specific, five of which have 61 to 268 times more expression in males than females. All these proteins were identified solely in males by MS experiments. These enzymes could thus be associated with the reproductive biology of ticks, where male ticks reportedly use their mouthparts and their saliva during copulation (Kiszewski et al., 2001). Male-specific seminal fluid serine proteases are known to occur in *Drosophila* (LaFlamme et al., 2012) and the human prostate-specific antigen (PSA) is a serine protease (Bilhartz et al., 1991;

Veveris-Lowe et al., 2007). It is possible that some male salivary proteins are exerting a function homologous to seminal fluid proteins of other animals. We additionally report a legumain (RpIx75-689581) that is male specific in *R. pulchellus* (>600-fold more expression). This enzyme was also found only in male SGEs by MS. The only other salivary legumain (RpIx75-534217) was found to be equally expressed in both sexes, and was identified only in female glands via MS. The specific function of serine proteases and legumains in tick saliva remains unknown.

Dipeptidyl peptidases were reported as the kininase found in *I. scapularis* saliva (Ribeiro and Mather, 1998). These enzymes are responsible for the destruction of bradykinin, thus interfering with the host's tick rejection reactions. Two dipeptidyl peptidases were reported for *R. pulchellus*. The protein encoded by RpIx75-19165 was found by MS to be expressed in females, with 39% coverage, with its transcripts six times more abundant in females than in males.

Serine carboxypeptidases with unknown function in feeding or reproduction were found by MS and their transcript expression are enriched in male glands. Particularly, RpIx75-906737 and RpIx75-906740 were identified by MS with 21% coverage in male glands only. On the other hand, RpSigp-677431 and RpSigp-677433 (equivalently expressed in both sexes), were identified by MS in both sexes.

Apyrases are enzymes hydrolyzing ATP and ADP to AMP thus suppressing agonists of inflammation and platelet aggregation. Mosquito and tick apyrases belong to the 5' nucleotidase family and lack the carboxyterminal

membrane domain linking them to a phosphoglycoinositol membrane anchor (Champagne et al., 1995; Stutzer et al., 2009). Twenty coding sequences were found representing this protein family. RpIx75-419942 and RpIx75-419943 were found in only female SGEs, while RpIx75-907283 and RpIx75-907282 were found in males only reflecting the transcription bias.

Nucleotidases catalyzing the hydrolysis of dinucleotides (ectonucleotide hydrolase/phosphodiesterase) were found, in which some are overexpressed in male salivary glands. These enzymes may hydrolyze diadenosine polyphosphates released by platelets, attenuating their role in hemostasis and inflammation (Ogilvie et al., 1996; Schluter et al., 1996).

Ribonuclease coding transcripts were found overexpressed in female glands; the products for two enzymes (RpIx75-670367 and RpIx75-938568) were found by MS solely in females.

3.2.4.2 Proteinase inhibitor domains

The following families of protease inhibitor domains were found abundantly expressed in the sialotranscriptome of *R. pulchellus* (Table 10): Kunitz, serpin, TIL (associated with serine proteinase inhibition), and cystatin and thyropin (associated with cysteine proteinase inhibition and immunosuppression).

The Kunitz domain is 50-60 amino acid residues in length and their fold is conserved by three disulphide bonds. Most Kunitz domain containing proteins inhibit serine proteases, although some also block ion channels (Ascenzi et al.,

2003; Castaneda and Harvey, 2009; Francischetti et al., 2009a; Fry et al., 2009; Monteiro, 2005; Paesen et al., 2009). In some tick saliva proteins, Kunitz domains are found in tandem; depending on the number of Kunitz domains, they are classified as monolaris, bilaris, trilaris and so on. Some are well known inhibitors of various proteases in the coagulation cascade. For example, Ixolaris, a bilaris isolated from *I. scapularis*, is an anticoagulant that binds both factor VIIa and factor X thus inhibiting the extrinsic pathway (Carneiro-Lobo et al., 2009; Francischetti et al., 2002a; Monteiro et al., 2008a; Monteiro et al., 2005b; Nazareth et al., 2006). In the *R. pulchellus* transcriptome, 197 members containing Kunitz domains have been identified (Table 10). Six monolaris, two bilaris and one tetralaris proteins were overexpressed in females (>100-fold). In the proteome (Table 11), one monolaris, five bilaris and one tetralaris proteins were identified in females, as compared to only one monolaris protein in males. Further in Section 3.3, the bilaris protein family will be analyzed in detail.

Serpins are ubiquitous serine protease inhibitors and several in tick salivary glands act as clotting inhibitors or chymase inhibitor (Chalaire et al., 2011; Chmelar et al., 2011; Rodriguez-Valle et al., 2012). We have identified 23 serpins in the *R. pulchellus* transcriptome. Notably, several were expressed exclusively in males, such as RpIx75-772874 having 2,559 reads from males and only two from females. Not surprisingly, the protein was identified only in male SGEs by MS. It is possible that this serpin plays a role in tick reproduction as does serpins in the seminal fluid of *Drosophila* (Walker et al., 2006; Wolfner et al., 1997) and prostatic secretions of humans (Seregni et al., 1996).

Cystatins are ubiquitous inhibitors of cysteinyl proteases. Some tick salivary cystatins show immunosuppressive and anti-inflammatory activities (Horka et al., 2012; Kotsyfakis et al., 2008; Kotsyfakis et al., 2010; Kotsyfakis et al., 2007; Kotsyfakis et al., 2006; Salat et al., 2010; Schwarz et al., 2012; Yamaji et al., 2009). Several cystatin coding sequences were overexpressed in males, two of which are over 1,000-fold overexpressed (RpSigp-19398 and RpSigp-384543) and their protein products were detected by MS only in males. Cystatins were also detected in human seminal fluid (Moura et al., 2010; Yadav et al., 2013).

The TIL family in *R. pulchellus* contains 95 members, including peptides closely related to tick elastase inhibitors, which also have antimicrobial activity (Fogaca et al., 2005). Proteins with multiple TIL domains (biTIL, triTIL and polyTIL) were also recognized. Many of these transcripts were gender biased. RpSigp-666840 was 716 times overexpressed in males and its product was detected by MS only in male glands. Several others were found by MS in either sex reflecting their expression bias.

3.2.4.3 Immunity-related proteins

Sialotranscriptomes of hematophagous animals contain antimicrobial peptides and other proteins associated with immunity. The sialotranscriptome of *R. pulchellus* reveals four different lysozymes, one of which (RpIx75-909446) appears male specific and was found only in male SGEs by MS. In addition, 22 defensins were found, of which one (RpSigp-635550) was identified by MS only in males.

Full-length thioester containing proteins (TEP), involved in innate immunity (Blandin and Levashina, 2004; Kopacek et al., 2010) were identified. MS identification was achieved solely from female tissues. Pathogen-recognition proteins, such as ML-domain containing proteins (Horackova et al., 2010), peptidoglycan-binding proteins (Kim et al., 2000), galectins and ficolins (ixoderin) (Rego et al., 2005) were identified. Three out of six galectins were identified by MS in females only.

3.2.4.4 Antimicrobial peptides

Microplusin is a histidine-rich peptide from *R. microplus* (Fogaca et al., 2004), six homologs of which were found in *R. pulchellus.* Two microplusins (RpIx75-675437 and RpIx75-675436) are highly female biased (190- and 257-fold) and were found by MS only in female glands. The 5.3 kDa peptide family was initially described in *I. scapularis* sialotranscriptomes, some members were upregulated following infection with *Borrelia burgdorferi* (Ribeiro et al., 2006). Later, one of its members was shown to be an antimicrobial (Pichu et al., 2009). *R. pulchellus* has four contigs matching members of this peptide family, in which one (RpIx75-680620) was highly expressed in female SGEs (>50-fold).

3.2.4.5 Tick-specific protease inhibitors

The tick salivary carboxypeptidase inhibitor is a 97 amino acid residue peptide found in *R. bursa* that stimulates fibrinolysis by inhibiting the thrombin-activated fibrinolysis inhibitor (Arolas et al., 2005a; Arolas et al., 2005b).

Haemaphisalis and *A. maculatum* ticks also have this protein family (Gong et al., 2007). *R. pulchellus* sialotranscriptome reveals 13 CDS that are 32- 88% identical to the canonical *R. bursa* protein, indicating the diversity of this peptide family within a single genome. One of these CDS (RpSigp-946756) was more than 80 times overexpressed in female ticks, though it was not found in the proteome. Another protein (RpSigp-946756) was found in females only.

The basic tail protein family is found in abundance in *I. scapularis*, one member of which was shown to inhibit FXa (Narasimhan et al., 2002). In metastriates, this family is much smaller. Nine CDS were found in the *R. pulchellus* sialotranscriptome, one of which (RpIx75-966959) was 36-fold overexpressed in females and was found by MS only in female tissues. The 18.3 kDa family was shown by PSI-BLAST analysis to be part of the basic tail superfamily (Francischetti et al., 2009a). This family is abundant in both metastriate and prostriate ticks, and is well represented within the *R. pulchellus* sialotranscriptome. The diversity of this family is verified by their best matching metastriate tick proteins available in the NR database which varies from 33-65% identity. None of them have been functionally characterized.

3.2.4.6 Glycine-rich proteins and mucins

Ticks use a cement-like substance to attach themselves to their hosts (Sauer et al., 1995). These proteins are rich in glycine and are somewhat similar to spider silk proteins. Some of these proteins have been used as anti-tick vaccines (Bishop et al., 2002; Maruyama et al., 2010; Zhou et al., 2006). The sialotranscriptome of *R. pulchellus* presents 151 CDS, including 26 that have

over 100,000 reads mapped to them, indicating the abundance of these transcripts. Although this class of proteins constituted almost half of the total number of S class reads, only 14 were identified by MS. This low detection rate in MS can be explained by their sequence, which has fewer arginine and lysine residues. Thus, most tryptic fragments were too large for MS/MS sequence determination.

Mucins are proteins with a large number of *N*-acetyl-galactosamine residues linked to serine or threonine (Tian and Ten Hagen, 2009). However, their primary sequence is not conserved and they have large regions of repeats with low complexity. Similar to the glycine-rich proteins, some members are highly expressed and 15 have over 10,000 reads mapped to their CDS. Four of these highly expressed CDS are male specific and may have to do with reproduction (Feldman-Muhsam et al., 1970). For example, RpSigp-99535 is 890-fold overexpressed in males and was found by MS only in male SGEs. Although there were 118 members found in the transcriptome, only six (four exclusively in males and two exclusively in females) were detected in the proteomes. This could be due to their high levels of glycosylation and large glycosylated peptides are not amenable for MS/MS sequencing.

3.2.4.7 Lipocalins

Although lipocalins are ubiquitous, tick salivary lipocalins have no sequence similarity to other known proteins, but are recognized by their typical barrel-like structure (Flower et al., 2000) revealed in crystal structures (Mans et al., 2008c; Paesen et al., 1999). Literally lipocalin means ‘a cup of lipid’, as their

barrel structure contains normally a lipophilic interior where lipids bind and can be carried in the aqueous environment. Tick salivary lipocalins can function by scavenging biogenic amines (Mans et al., 2008c; Paesen et al., 1999; Paesen et al., 2000) or lipidic mediators of inflammation (Mans and Ribeiro, 2008a, b), normally using binding sites inside the barrel, but also by acting as anti-complement proteins using their side chains on the outer surface (Nunn et al., 2005). Supplemental file “RP Sialotranscriptome’ displays 350 coding sequences that are assigned to the lipocalin class. Some of these do not have any direct match by BLAST or RPS-BLAST to known lipocalins, but they cluster at 40% similarity to members of the lipocalin family and hence are grouped as a subcategory putative lipocalins. Several lipocalins show sex biased expression, 38 and 7 are overexpressed (>100-fold) in females and males, respectively. Mass spectrometry identified 59 of the 350 predicted lipocalins.

3.2.4.8 Ixodegrins

Ixodegrins are ~100-residue polypeptides that contain an RGD tripeptide flanked by cysteines (Francischetti et al., 2005b; Francischetti et al., 2009a) similar to disintegrins – platelet aggregation inhibitors found in snake venom (Calvete et al., 2003; Huang and Niewriarowski, 1994). *R. pulchellus* sialotranscriptome contains several peptides but in most, RGD is replaced by RED sequence. Three ixodegrins were identified by MS in female glands.

3.2.4.9 DA-p36 family

DA-p36 is a 36 kDa immunosuppressive protein isolated from the salivary glands of *Dermacentor andersoni* (Bergman et al., 2000). This family is widely found in metastriate ticks and 28 were identified in *R. pulchellus*. Fifteen were overexpressed in females (10-390-fold), while none were overexpressed in males (<3.8-fold). Seven members of this family were identified by MS – 3 exclusively in females, 1 in males, and 3 in both sexes.

3.2.4.10 Evasins

Evasins are cytokine-binding proteins isolated from the tick *R. sanguineus* (Deruaz et al., 2008; Frauenschuh et al., 2007). This family is found in all sialotranscriptomes of metastriate, but not in prostriate ticks. The family is extremely divergent; the best matches to *R. pulchellus* proteins provide only 26-56% identity at the primary sequence level. Four evasins were female specific (50-230-fold) while one evasin was male specific (54-fold).

3.2.4.11 Immunoglobulin G binding proteins

Tick immunoglobulin binding proteins help to scavenge host IgG that leaks into the hemolymph and salivary members of this family recycle host IgG from ticks back to hosts (Wang and Nuttall, 1995a, b, 1999; Wang et al., 1998). Male-only forms of this protein were identified in *R. appendiculatus* (Wang et al., 1998) and were shown to help their female counterparts feed. The sialomes

of *R. pulchellus* reveals several transcripts with expression virtually in males; for example, the transcript coding for RpSigp-668710 has over 480,000 reads mapped to it from the male library, but only 202 reads from the female library. Other members were also overexpressed in males (>1,000-fold). RpIx75-464069, which was 1,715-fold overexpressed in male glands, was found solely in male glands by MS.

3.2.4.12 Tick-specific, unknown function

Fifty eight families of proteins are tick specific and have no known function. This group includes 34 families of small proteins (<100 residues) that have no similarities to other known proteins and could be erroneously deduced from non-coding RNA. Only one peptide match was found for members of this group.

Prostriate/metastriate families

The members of “Fibronectin-domain protein” and “Hematopoietic-stem cell progenitor” families in prostriate and metastriate ticks are quite similar, suggesting they may have housekeeping function. Alternatively, they may be of low antigenicity to their hosts either because of their low expression or lack of strong epitopes. The transcriptome confirms their low levels of expression, making up less than 0.01% reads of the S class. However, RpIx75-854185 was identified by MS in female glands.

The 8.9 kDa family was found in all tick sialomes. Sixty CDS were found in the *R. pulchellus* sialotranscriptome (Supplemental file RP Siaotranscriptome). Some of these peptides have disintegrin domains. Both male and female specific expression were observed; RpIx75-940640, RpSigp-798933 and RpSigp-675391 were overtranscribed in females (12-26-fold) and were detected by MS only in female glands, while RpSigp-7927, RpSigp-972527, RpSigp-509737 and RpSigp-709817 were overtranscribed in male glands (4, 186, 650 and 994 -fold) and were detected by MS only in male glands.

The 23 kDa and 24 kDa protein families of ixodids were also represented in the *R. pulchellus* sialotranscriptome. RpIx75-677306 (24 kDa family) was found by MS in glands from both sexes.

The “one of each” family was initially thought to be only found in metatriates, where only one gene from the family was found in a tick. However, this was contradicted in the sialotranscriptome of *A. maculatum*, where several CDS were found. In addition, homologs were also found in *I. scapularis*. We here report 14 additional CDS for this diverse family, 4 of which were identified by MS in only female SGEs, in accordance with their female-biased transcription (39-172 times overexpressed).

Previously thought to be metastriate-specific families

Seven protein families that were identified previously as metastriate specific were found to have homologs in the *I. scapularis* proteome. Two members of “Insulin growth factor binding proteins” (IGFBP) family were found in *A. variegatum*, a short form with only the insulin growth factor-binding protein

homologs (IB) SMART domain, and a long form with additional Kazal and immunoglobulin (IG) domains (Ribeiro et al., 2011). The sialomes of *R. pulchellus* reveals six proteins of the short variety, which matches not only metastriate, but also *I. scapularis* proteins.

A deduced protein from *R. appendiculatus* expressed sequence tags (ESTs) (Francischetti et al., 2009a) was annotated as orphan as it had no similarities to other known proteins, but it matches *R. pulchellus* proteins having a weak Somatomedin SMART domain.

Metastriate specific families

There are seven additional protein families which are metastriate-specific. Five were previously described, while two, the 9 kDa protein family and the Rhip 45-304 family, were deorphanized with the R. pulchellus sialotranscriptome. The female-biased Rhip 45-304 family (~200-250 residues) was so named because the members cluster at 45% similarity within cluster 304 of the database protein clusterization algorithm. RpIx75-16871 protein was identified by MS only in female glands.

Rhipicephalus specific families

Three protein families from *R. pulchellus* have similarities exclusively to *R. appendiculatus* proteins and may represent genus-specific proteins. RpIx75-569044 and RpIx75-697446 from the 45-236 family were identified by MS in male SGEs.

3.2.5 Transposable elements

CDS encoding TEs are commonly found in sialotranscriptomes, and were also found in *R. pulchellus*. Both class I (Ty3/Gypsy, L1, Jockey, Bell, Outcast and Copia) and class II (hAT, Pogo, P, piggyback and tigger) elements (Wicker et al., 2007) were identified. Several of these elements, such as tigger sequence RpTe-704693, have intact transposases. The recently discovered element sola2 (Bao et al., 2009) was also represented. A new family of TEs matching the PFAM domain (DDE_4 superfamily endonuclease) was found, and is similar to *Ixodes* transposons (Repbase accession: IS4EU-1_BF) (Kapitonov and Jurka, 2008). Penelope-derived transcripts were also found abundantly.

Several transposons also show differential expression; RpSigp-1811, RpTe-11787 and RpTe-15723 are highly expressed in females and RpSigp-101516 and RpSigp-908814 in males. Some TE classes are abundantly transcribed, including several Gypsy elements with more than 20,000 reads mapped to their transcripts; RpTe-910969 was mapped from 65,688 reads and expressed in females 37-fold higher than in males. Most of the proteins associated with these Gypsy transcripts have frame shifts indicating these may function more as regulators of transposition than active transposition. It was recently proposed that TEs are associated with gene expression regulation (Kelley and Rinn, 2012). The observed differential expression of TEs between sexes may contribute to gender-dependent gene expression in *R. pulchellus* salivary glands. Seven TEs were identified by MS from female SGEs indicative of their translation.

3.3 Gender-dependent expression of Bilaris proteins

As shown in Section 3.1, the saliva of male ticks contained less FXa and thrombin inhibitors. Thus, in an attempt to search for male-specific anticoagulants, the bilaris proteins (with two tandem Kunitz domains) in the transcriptome were further analyzed, as this class of proteins has well-characterized anticoagulants (Soares et al., 2012; van de Locht et al., 1996b).

3.3.1 Subclasses of *R. pulchellus* Bilaris proteins

A total of 81 coding sequences for bilaris proteins in the *R. pulchellus* transcriptome has been identified. By systematic analysis, it was found that they could be further classified into five subclasses based on their number of cysteine residues, as well as their length (Figure 3.11). In a typical Kunitz domain, the fold is conserved by three conserved disulphide bonds: C1-C6, C2-C4 and C3-C5. In the first subclass, both Kunitz domains have all three disulphide bonds conserved. In all other subclasses, the first Kunitz domain contains all three conserved disulphide bridges. However, marked differences are seen in the second Kunitz domain. In the second subclass, cysteine residues C2 and C4 are substituted, resulting in the loss of one disulphide bridge in the second Kunitz domain. This is seen in ixolaris (Francischetti et al., 2002a). Similarly, proteins in subclass III has cysteine residues C3 and C5 substituted, thus it also has only two disulphide bridges in the second Kunitz domain. In subclass IV and V, an additional pair of cysteine residues, is present at the C-terminal of the second Kunitz domain. However, subclass V has a long proline-rich inter-domain segment (73 - 85 residues).

It is interesting to note that even within each subclass, similarity between the individual proteins, not taking into account isoforms, are low; the majority of the proteins within a subclass had less than 40% identity to each other. Proteins within each subclass were aligned and the consensus sequence for each subclass was generated and aligned with other subclasses (Figure 3.12). This alignment further highlights the characteristic differences across the five subclasses and the low conservation of sequences in the inter-cysteine regions within as well as out of each subclass.

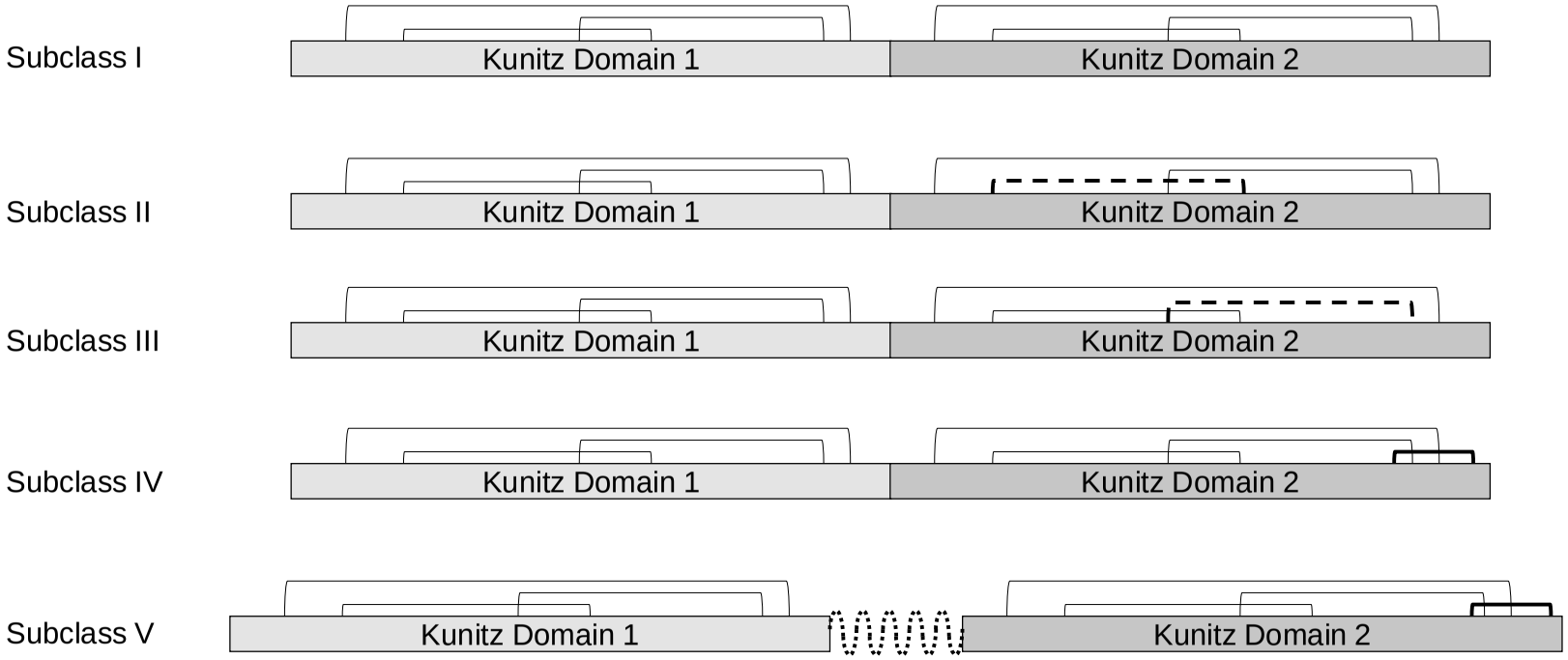

Figure 3.11. Five subclasses of bilaris proteins. Schematic representation of the five subclasses of bilaris proteins. Subclass I has two full tandem Kunitz domains, each three conserved disulphide bonds, represented by the solid lines (——). A typical Kunitz domain disulphide bond pairing is as follows: C1-C6, C2-C4 and C3-C5. Subclass II and III has two missing cysteine residues C2,C4 and C3,C5, respectively, in the second Kunitz domain, resulting in a loss of a disulphide bond. Missing disulphide bonds are represented by the bold dashed line (**– –**). Subclass IV and V has two extra cysteine residues in the second Kunitz domain, one located between the typical C4 and C5 cysteine residues, and the other after the C6 cysteine residue. This extra disulphide bond is represented by a bold solid line (**——**). However, subclass V contains a long inter-domain, proline-rich segment (~ 80 residues and represented by the bold dotted line (**·····**)).

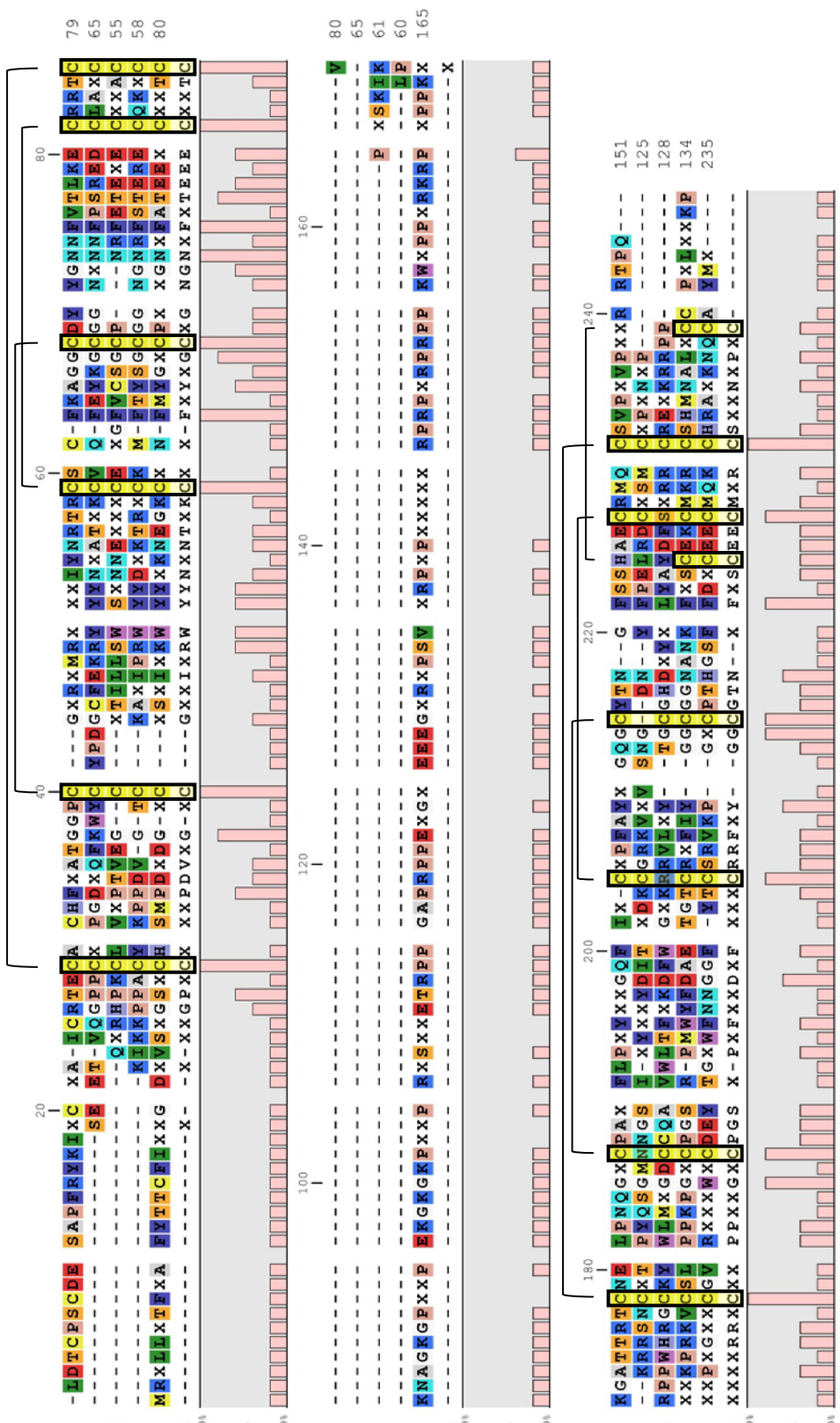

Figure 3.12. Sequence alignment of bilaris subclases. The consensus sequences of all five bilaris subclasses were aligned. Cysteine residues are highlighted in yellow and disulphide bridges are drawn to link them.

3.3.2 Differential expression of bilaris proteins

The sequences of all bilaris proteins were aligned using ClustalW and a phylogenetic tree was constructed (Section 2.3.3) (Figure 3.13). Proteins from the same subclass clustered together even though they had low sequence homology. Figure 3.13 also shows the abundance of reads from the transcriptome for each contig. Four proteins (RpSigp-921556, RpSigp-673073, RpIx75-909765 and RpIx75-477142) marked with the symbol ♀ were identified in the female proteome. This correlates to the transcriptome data; from the graph, it can be seen that these four proteins were overexpressed in females as compared to males.

In order to have a better comparison of the difference in number of reads of each contig from each sex, the normalized number of female reads was divided by the normalized number of male reads. Figure 3.14 illustrates the number of fold differences of each contig. There were altogether eight proteins that were more than 10 times expressed in either sex; seven were higher in females while one was higher in males. In addition, for each subclass, contigs were stratified according to the gender they were more highly expressed in. This data is represented in the table within the graph in Figure 3.14. It was interesting to see that the majority of subclass II bilaris proteins (10 of 12 contigs) were more highly expressed in females, while the majority of subclass I and IV bilaris proteins (16 of 19 and 28 of 34 contigs, respectively) were higher in males. This further suggests that male and female ticks employ different sets of proteins to target the host hemostatic system.

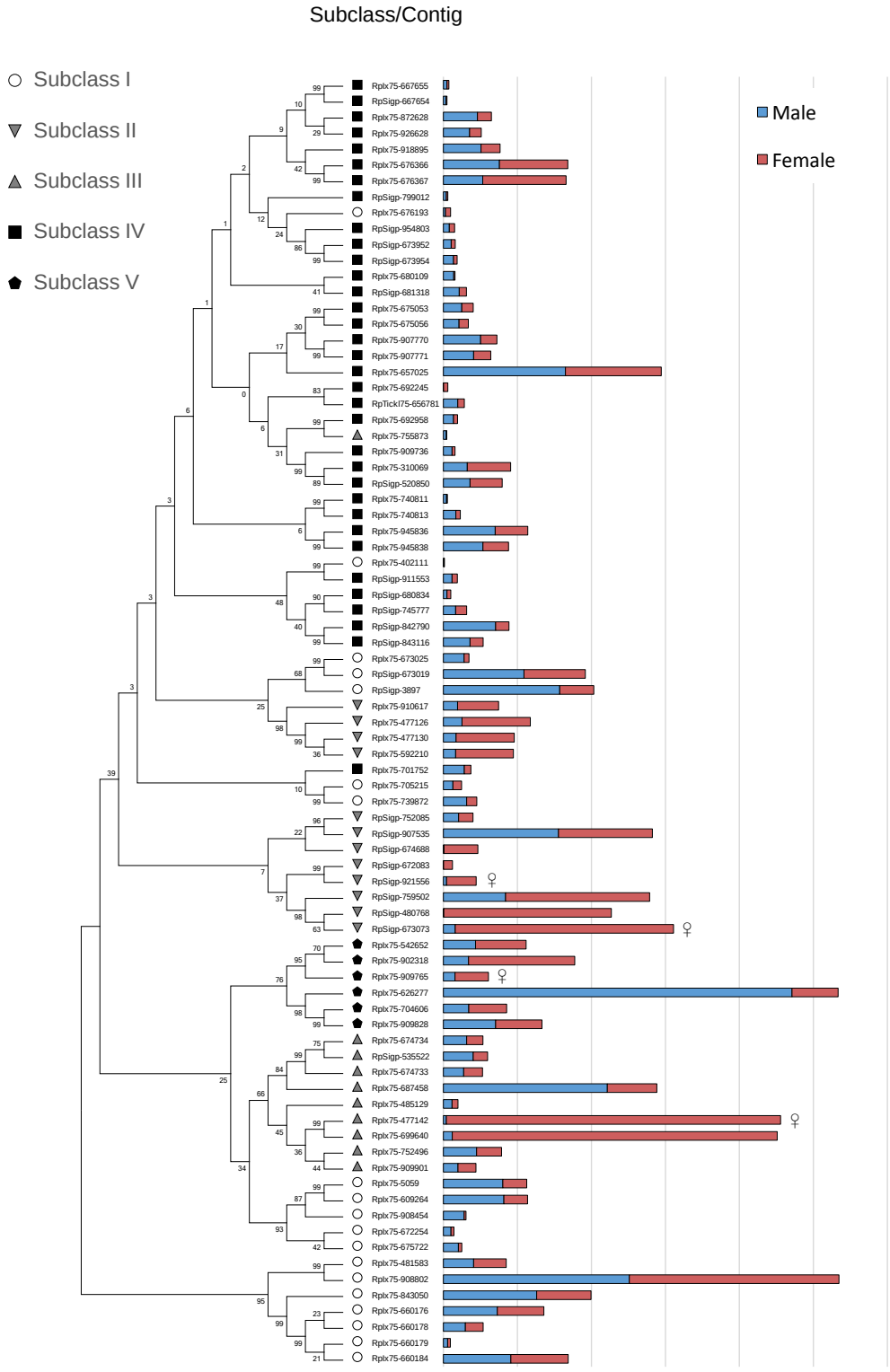

Figure 3.13. Phylogenetics and associated number of reads of bilaris CDS. 81 CDS belonging to the bilaris class were aligned with ClustalW and a phylogenetic tree was drawn using the Neighbor-Joining method in MEGA5. The subclass of each CDS is indicated by the symbol on the left of the accession number. The number of reads associated with each CDS are indicated by the bar chart, where red represents reads from the female library, and blue from the male library. Proteins that were identified in the female proteome are marked with a ♀ beside the bar chart.

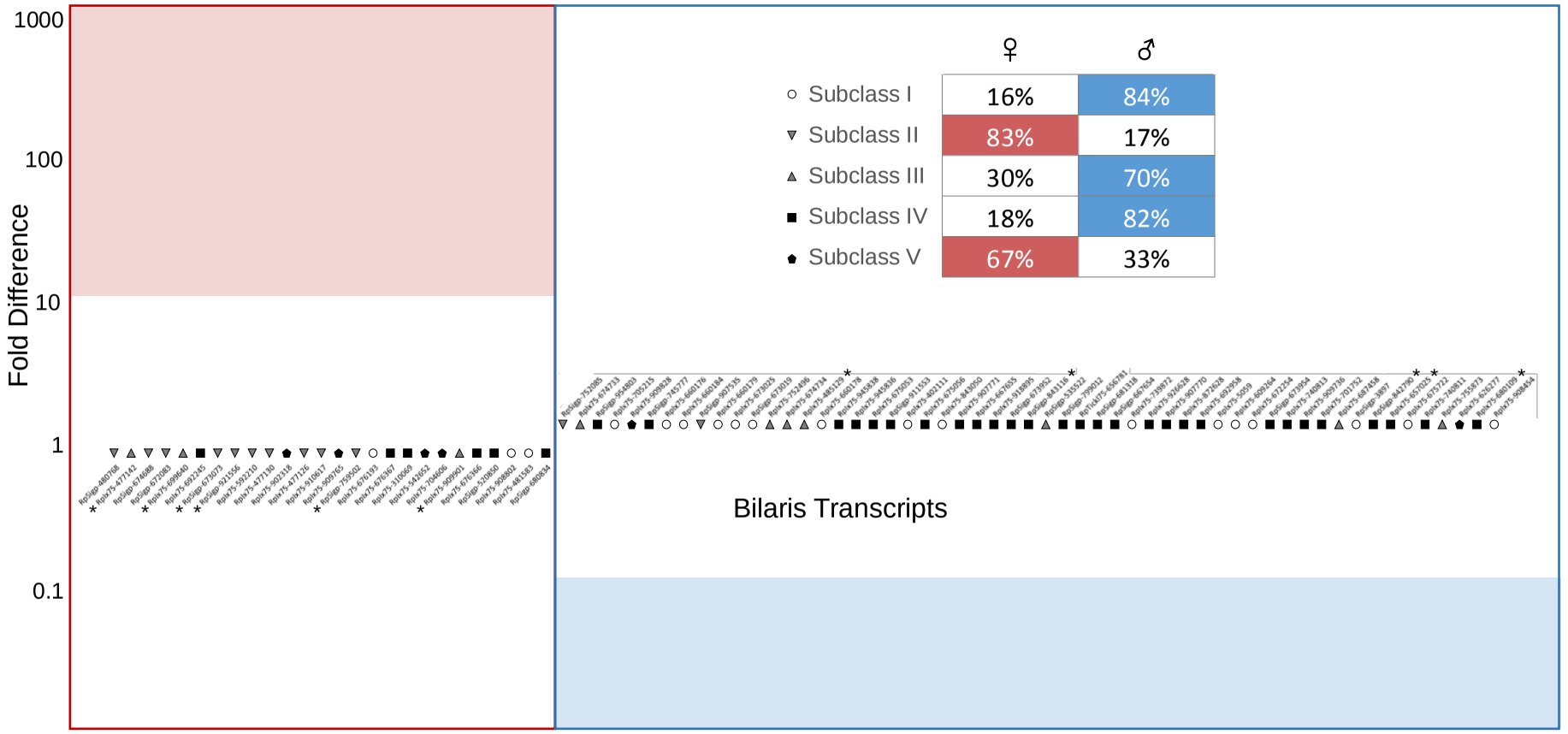

Figure 3.14. Expression difference of bilaris proteins between male and female *R. pulchellus*. Fold difference in the transcriptome is expressed over number of male reads. Seven transcripts were highly expressed in females while one was highly expressed in males. These sex biased transcripts which were more than 10 fold expressed in either sex are highlighted in grey. The subclass of the bilaris proteins are indicated beside the transcript name. It is noted that most of subclass I, III and IV were expressed by males, and subclass II and V in females, as tabulated in the inset. Transcripts marked with an asterisk (*) were used for RT-PCR studies.

3.3.3 Relative abundance of bilaris proteins

As data from the transcriptome only provides a rough estimate on the level of expression of these proteins, quantitative real time – polymerase chain reaction (qRT-PCR) experiments was performed on selected bilaris proteins. This allows the comparison of the relative abundance of mRNA of each gene in males and females, and thus the relative level of expression. Specifically, subclass II, III and IV, were of interest in this whole study due their unique structures previously mentioned. Thus, four genes were chosen from each subclass (full sequences of chosen genes are listed in the Appendix 1). The CDS chosen are also marked with an asterisk (*) in Figure 3.14.

RNA was first extracted from a new batch of salivary glands and the first-strand cDNA was synthesized (Section 2.5.1). Primers were designed based on the sequences from the transcriptome (Section 2.5.2), and also the primers for the β-actin internal control gene. They were checked for their specificity by first performing a PCR amplification (Section 2.5.3) with the cDNA as a template. Amplified products were subjected to gel electrophoresis on a DNA agarose gel (Section 2.5.3) to check for amplification and single products (Figure 3.15). However, one gene from subclass II (RpSigp-907535) was unable to be successfully amplified. This may be due to batch variation between the two batches of ticks as a new batch of ticks, different from the transcriptome, had been used. To verify the accuracy of the amplified products, the PCR products were purified (Section 2.5.5) and their sequences were determined (Section 2.5.6). Sequences obtained were clean, indicating that a primers used had resulted in a single PCR product. In most of the genes, point mutations

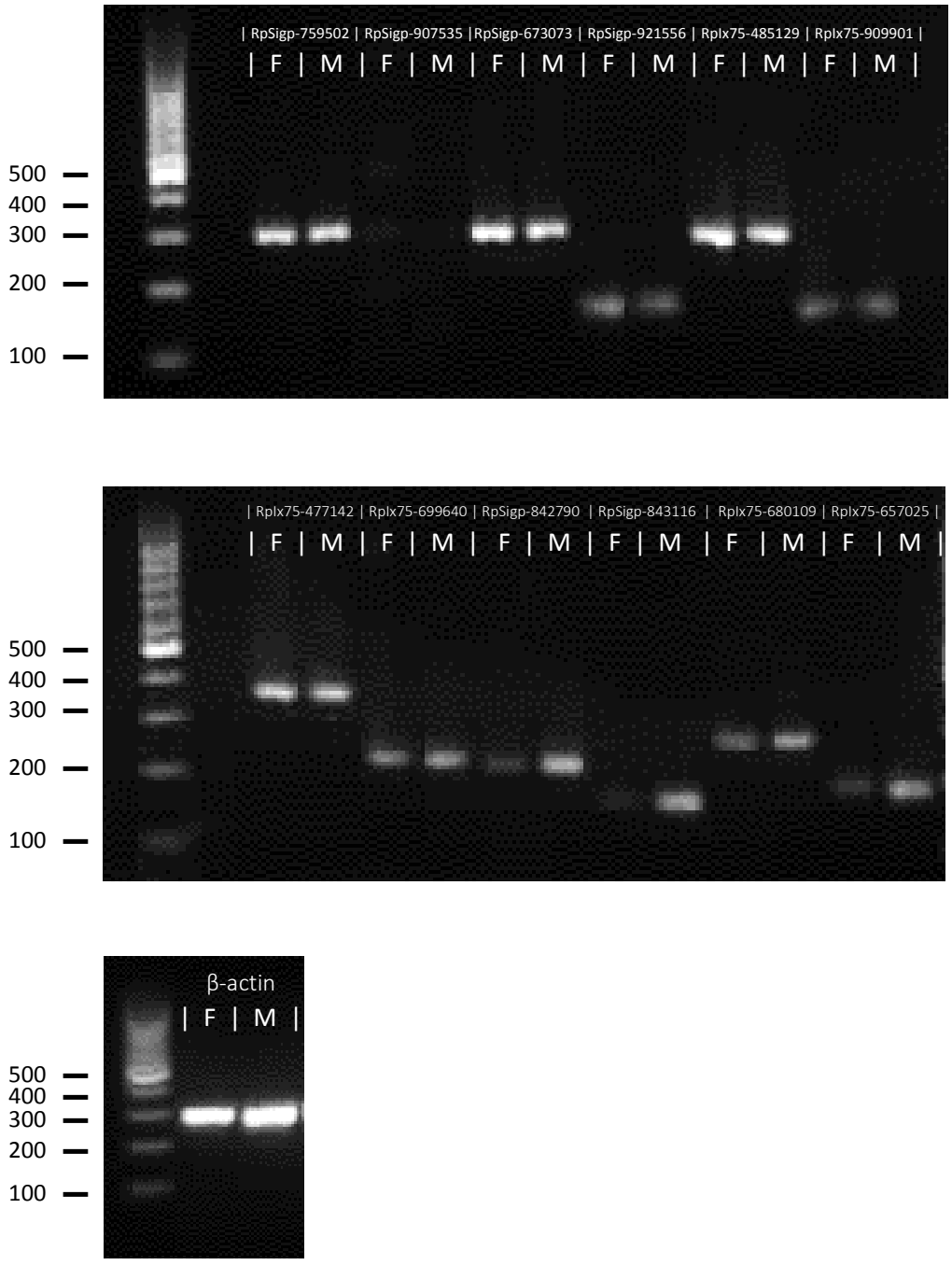

Figure 3.15. PCR screening of qPCR primers. Primers designed for the use of the qPCR experiments were used in a normal PCR to check for amplification and specificity. DNA gel electrophoresis was performed on the PCR products. In total, there were 12 bilaris genes and β-actin control gene. One gene was not able to be amplified (RpSigp-907535).

were observed throughout the amplified segment. This could be also explained by batch variation of the ticks used.

qRT-PCR was performed with β-actin as an internal control and the level of expression of males are shown in reference to the female genes (Figure 3.16). Two proteins RpIx75-699640 and RpIx75-477142 from subclass III had more than 30 times expression in females than in males, which correlates well to the transcriptome data. Proteins from subclass IV (RpSigp-842790, RpSigp-843116, RpIx75-680109 and PrIx75-657025) were highly expressed in males, with the former two being expressed more than 2000 times higher in males than females. This corresponds to the elevated number of male reads we see in the transcriptome for proteins belonging to subclass IV, and serves to validate the overall transcriptional differences observed between genders.

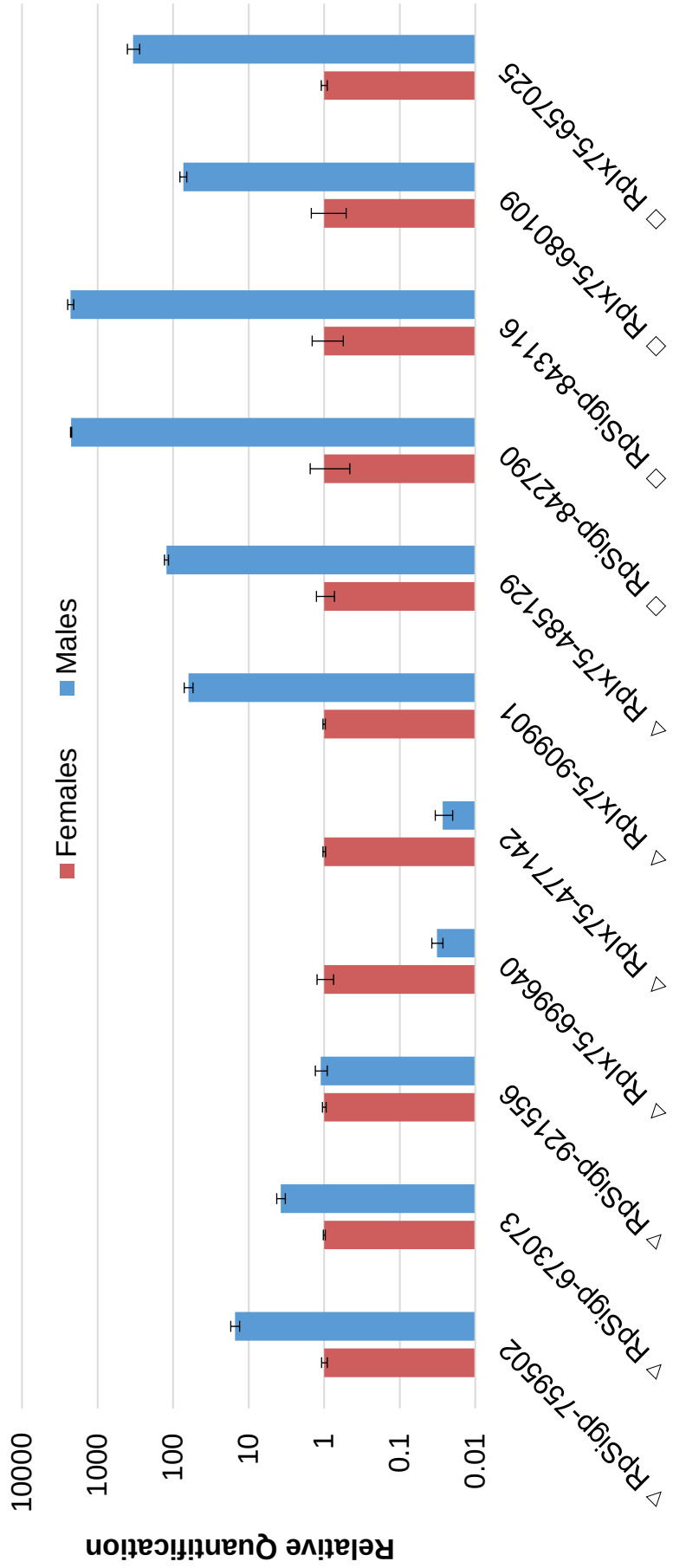

Figure 3.16. Quantitative PCR on selected bilaris proteins. RT-PCR was performed on selected bilaris proteins from subclass II, III and IV. Comparison of expression is in reference to females. Two independent biological samples were used to obtain the standard deviation.

3.4 Recombinant expression of Bilaris proteins

In order to study the differences of the various Bilaris proteins and how their functions relate to their abundance in each sex, functional assays would be required. This will allow us to investigate the region in the blood coagulation cascade in which they target. Hence, recombinant expression of these selected Bilaris proteins is essential. Briefly, they were amplified from the cDNA library and cloned into suitable expression vectors. They were then transfected into insect cell lines for expression. Subsequently, the proteins were screened for anticoagulant activity.

The genes of interest were amplified from the cDNA (Section 2.6.1) by a PCR step similar to Section 2.5.3. The extension time was increase to 45 s instead of 30 s. DNA gel electrophoresis was performed on the PCR products and DNA was extracted on the product bands. Figure 3.17 shows the DNA bands of six of the amplified genes RpSigp-759502, RpSigp-673073, RpIx75-485129, RpIx75-699640, RpSigp-843116 and RpIx75-680109. They were then checked for their sequence through DNA sequencing (Section 2.5.5). Although the PCR was performed on both males and females cDNA, certain genes were able to be amplified from both sexes (RpSigp-759502, RpSigp-673073, RpIx75-699640 and RpIx75-680109), while some in either one (RpIx75-485129 in females and RpSigp-843116 in males). In cases where they were amplified in both sexes, their sequences were verified to be identical before using the gene for subsequent steps.

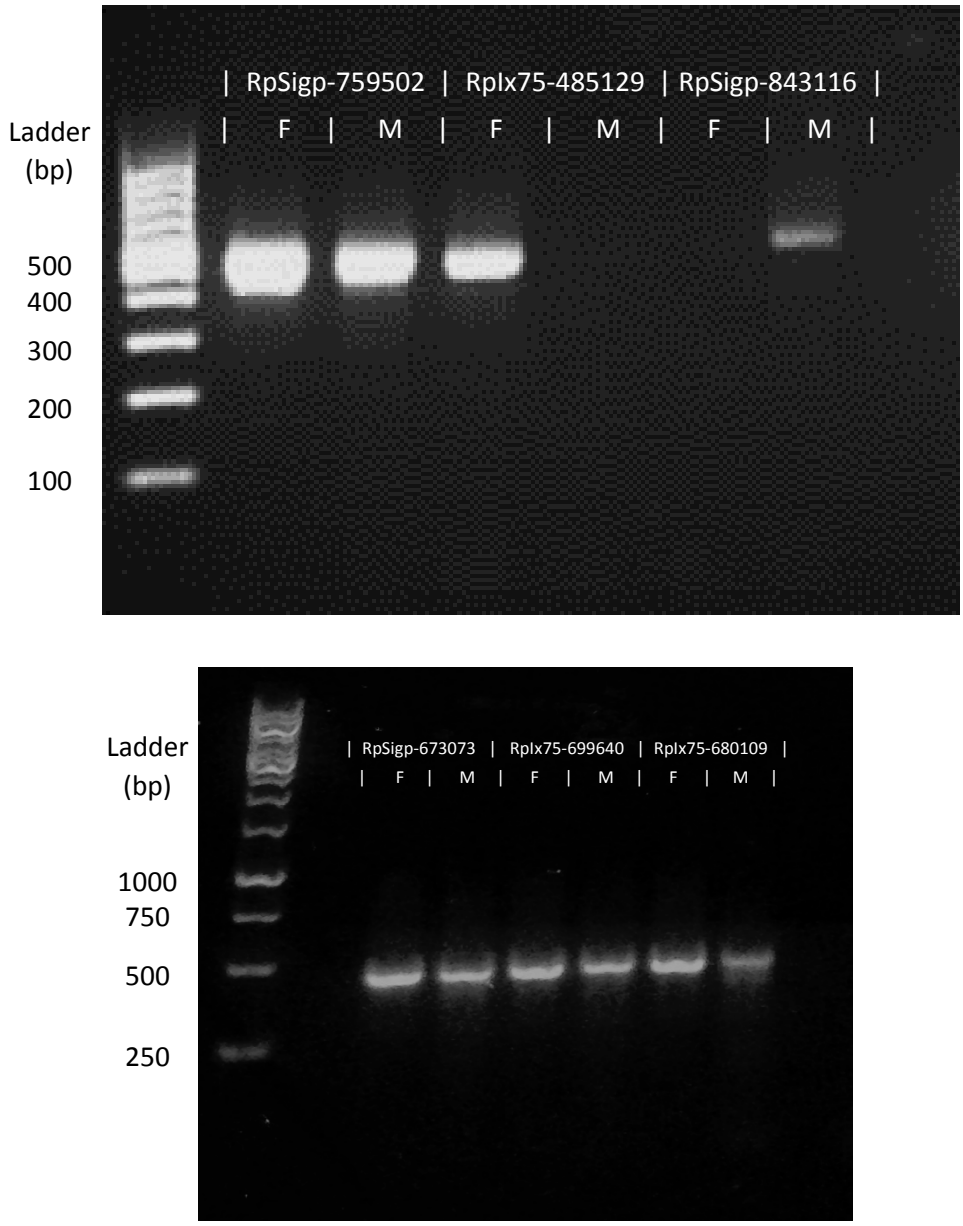

Figure 3.17. DNA gel electrophoresis of bilaris proteins. Six bilaris proteins were selected for amplification from male and female SG cDNA. DNA gel electrophoresis was performed on the PCR products to check for amplification and specificity.

Once verified, the genes were ligated into pIB/V5-His vectors, and transformed into competent One Shot TOP10 cells (Section 2.6.2). The plasmids were isolated and sequenced to check for success of ligation and orientation of the insert.

RpSigp-759502 was transfected into insect cell lines (Section 2.6.4) and after four days, the supernatant was harvested. The supernatant was concentrated and salts were removed by passing it through a 3K Amicon Ultra Centrifugal Filter and washing with PBS. The concentrated protein mixture was then screened for anticoagulant activity by measuring the APTT. Comparing to the controls with PBS, the APTT was not prolonged (Fig. 3.18).

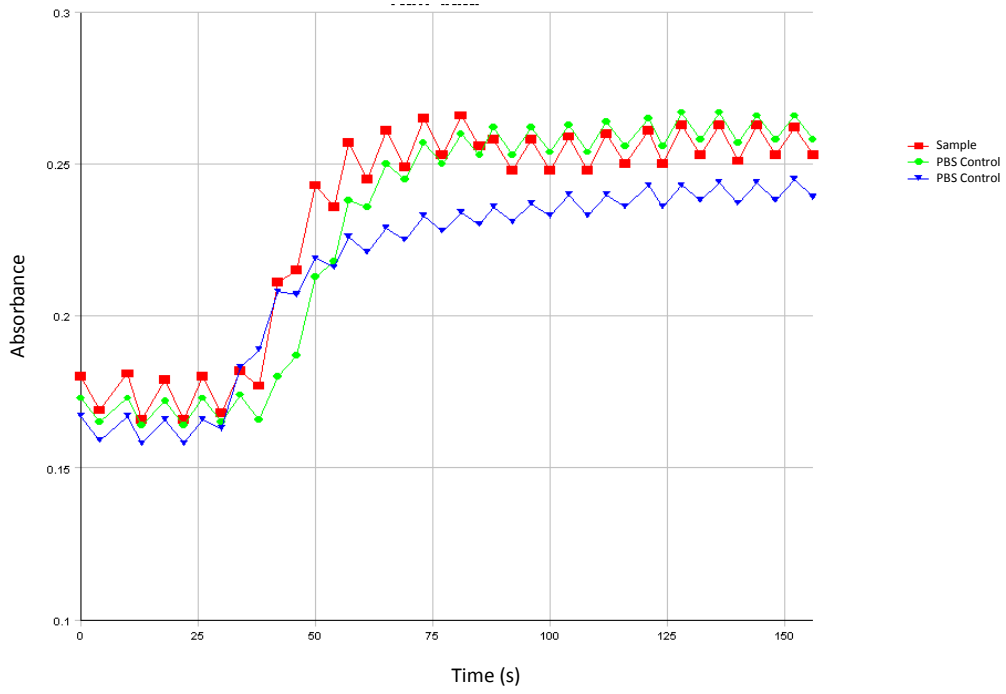

Figure 3.18. APTT assay of expressed RpSigp-759502. RpSigp-759502 was expressed in High Five insect cell lines. The concentrated supernatant was screened for anticoagulant properties via the APTT assay.

CHAPTER 4

Discussion

4.1 Anticoagulant activities of tick saliva

We have demonstrated that the saliva of male and female *R. pulchellus* do contain anticoagulant activities, in particular, anti-FXa and anti-thrombin capabilities. Additionally, these anticoagulants differ in composition and quantity between the male and female saliva. The female salivary glands contained about four times more proteins than males. This can be accounted by the amount of blood a female tick requires during feeding, which is vastly more than male ticks. Females imbibe on such a large amount of blood that they increase they body weight to more than 100-fold, while males usually do not increase more than 2-fold (Chen et al., 2012a; Chen et al., 2012b; Ma et al., 2013; Zheng et al., 2011). In order to do this successfully, it is essential that female ticks produce enough saliva to be infused into the host such that their duration of feeding will not be interrupted. All adult Ixodid ticks, with the exception of the *Ixodes* species, require a blood meal for the gonotrophic cycle to be initiated (Sonenshine, 1991). Female ticks require a much larger volume of blood than males because they digest the blood components, converts them into eggs (Kiszewski et al., 2001; Sonenshine, 1991) and oviposit after the blood meals. In fact, the volume of the blood-meal has been shown to correlate with the number of eggs deposited; the larger the blood meal, the higher the number of eggs laid (Chen et al., 2012a; Chen et al., 2012b; Faccini et al., 2010; Ma et al., 2013; Siroky et al., 2011; Zheng et al., 2011).

In addition to the quantity of proteins that are higher in the salivary glands of female *R. pulchellus*, we have also shown that there is a difference in salivary composition between the genders, in particular, their anticoagulant

content. Female SGEs were able to inhibit both key factors (FXa and thrombin) of the blood coagulation cascade very well, whereas males were not able to. This makes a significant contribution towards the large blood meal of the female ticks, ensuring that the blood from the host does not clot en route to the tick's gut.

Similar screening was also performed on another species of ticks, *D. reticulatus*. Like *R. pulchellus*, the protein elution profile of male and female *D. reticulatus* SGEs were clearly distinct. In addition, there were more FXa inhibitors present in the females as compared to males. However, unlike *R. pulchellus*, the SGEs of *D. reticulatus* consisted of FXa inhibitors but not thrombin inhibitors. It was evident that even after several rounds of purification to isolate the FXa inhibitors, there was still presence of multiple other contaminating proteins. The FXa inhibitors were able to be pulled-down via the affinity column. However, attempts to further purify and sequence the protein were unsuccessful.

In fact, efforts were made to isolate the anticoagulants that were present in the female salivary gland extracts in both species of ticks. However, the quantity of these anticoagulants in the saliva is very low, which made their purification a challenge. Several thousand ticks would be required for purification optimization and isolation. In addition, due to the limitation in supply of the ticks because of logistic problems and the dependency on the season of ticks, such large number of ticks were not feasible. Thus, this provided the impetus to use molecular techniques to study the anticoagulants in the saliva of *R. pulchellus.* This particular species of tick was chosen over *D. reticulatus*

due to the differences seen between the composition of male and female salivary gland extracts. Hence, we used this opportunity to perform the study separately on males and females. Their differential expression of secretory proteins would be an interesting topic, which has never been reported in this depth before.

4.2 Sialome of *R. pulchellus*

We have generated the sialome of male and female *R. pulchellus* independently. From the transcriptome, we reported a total of 50,460 CDS that was assembled from 241,229,128 paired-end reads. As for the proteome, 2,231 proteins were identified, of which 1,777 and 169 proteins were found exclusively in females and males respectively, and 285 proteins in both sexes.

The discrepancies seen between the transcriptome and proteome can be due to several factors. Firstly, the proteomics work was performed on six days fed adult ticks in comparison to a mixture of unfed, and various time points of fed adults ticks (refer to Chapter 2) in the transcriptome. Thus the proteome may have a bias towards proteins that are found after six-day feeding. Secondly, the inherent constrains of proteomics limits the number of proteins that can be detected. Due to lower sensitivity in detection limits, proteins existing in lower abundance may not be identified. The nature of proteins themselves also affects detection in MS. For example, in highly glycosylated proteins (mucins, some TIL family members and Rhip 45-10), glycosylated tryptic peptides are either too large (particularly in N-glycosyl moiety) or masses do not match the peptide sequences ‘generated’ from the transcriptome. In other cases (e.g.

glycine-rich proteins) tryptic peptides are too large for the set limits in MS/MS due to the lower frequency and unfavourable distribution of Lys and Arg residues. Lastly, transcriptomics also contributes to the discrepancies. The number of reads may not directly represent the level of expressed proteins as other parameters such as translation efficiency and protein turnover may also affect the exact quantity of proteins at a given time. In addition, the assembly of reads sometimes produces fragmented transcripts resulting in more than one CDS for a single protein. Further, some peptides spanning over such fragmented regions will not be identified.

As indicated in the above Section 3.2, many transcripts were found differentially expressed in male or female salivary glands, in many cases remarkably so, with hundreds to thousands fold more reads in one sex. It is evident that some classes were more highly expressed in one sex. It is interesting to note that the total number of upregulated transcripts (10-fold or more) in females was almost twice that of males (1193 vs. 621) (Table 14). This may be attributable to the difference in feeding habits between male and female ticks, as previously mentioned. Females secrete more proteins that allow a larger blood meal. This increased intake of blood, necessitates metabolic difference to process it.

Table 14. Total number of sex-biased CDS in the *R. pulchellus* transcriptome.

Class	Females	% total	Males	% total
Secreted	416	34.9	142	22.9
Housekeeping	213	17.9	107	17.2
Transposable element	43	3.6	18	2.9
Unknown	521	43.7	354	57.0
Total	**1193**	**100**	**621**	**100**

The proteome data, although not geared for detailed quantitative measurements, similarly indicated many differentially expressed proteins, mainly when we consider those proteins found in only one of the two sexes. Of 2,231 proteins identified with at least one high-confidence peptide match, 285 were detected in both sexes, while 169 and 1,777 were identified only in male or female tissues, respectively (Table 15). Although the number of proteins identified in females was about 10 times that of males, most of them are accounted for by housekeeping proteins (1,466 in females vs. 74 males). Among these proteins are notably those responsible for the transcription, synthesis, modification and export of proteins. This ties in with the transcriptome data where there were almost twice as many transcripts that were overexpressed in females as compared to males. In addition, a significant number of signal transduction proteins were found in females. A total of 226 secretory proteins were identified in the proteome – 31 found in both sexes, 77 exclusively in males, and 113 exclusively in females. The lipocalin family was found abundantly in both genders, which corresponds to the transcriptome where this family had the third most abundant reads.

The majority of the transcripts overexpressed in females (44%) and males (57%) were of the U class, but many have signal peptides indicative of secretion and could be novel peptides.

Table 15. Total number of gender-biased proteins from the *R. pulchellus* proteome.

Class	Found in female only		Found in male only	
	Number	% of total	Number	% of total
Secreted	113	6.4	77	45.6
Housekeeping	1632	91.8	83	49.1
Transposable element	7	0.4	0	0
Unknown	25	1.4	9	5.3
Total	1777	100	169	100

The H class transcripts that are upregulated in females (19%) included transporters and enzymes involved in gluconeogenesis (4-aminobutyrate aminotransferase and tyrosine transaminase). DNAses may actually be secreted and prevent NET formation (Wartha et al., 2007). Similarly, female lipases may be secreted (Bowman et al., 1997). On the contrary, the H class transcripts upregulated in males (17%) included sulfotransferases, sulfatases and multi-drug transporters for detoxification, and amino acid, ABC, and monoamine transporters. They also code for enzymes targeting aromatic amino acids, such as dopamine β-monooxygenase and aromatic-L-amino-acid/L-histidine decarboxylase, which may be associated with hormone or pheromone metabolism (Sonenshine, 1985). Lipases (with >400-fold more reads in males)

and long-chain acyl-CoA synthetase indicate exceptional lipid metabolism in male salivary glands. The TE class transcripts that are upregulated in males (3%) include transcription factor doublesex, which is associated with sex-specific gene expression (Portman, 2007; Shukla and Nagaraju, 2010).

Notable classes of secretory proteins that were in abundance in females in both the transcriptome and proteome belonged to the metalloprotease, Kunitz-type inhibitors, TEP proteins, glycine-rich superfamily, lipocalins, ixodegrins, and the DA-p36 family. On the other hand, those that were highly expressed in both the transcriptome and proteome of male ticks were the male specific serine proteases, antigen 5 family, serpins, cystatins, and the IgG binding proteins. The general trend of these differentially expressed proteins tends to be of immunosuppressive and anti-hemostatic properties and their specific functions have been outlined in detail in Section 3.2.4.

The way ticks feed has continuously been an intriguing topic. Earlier literature had previously reported a difference in salivary composition between male and female ticks (Gasperik J, 2000), though specific proteins were not been identified. However, other sources have identified certain classes of proteins to be male-specific. The immunoglobulin G binding proteins have been recognised as being produced in the male salivary glands to combat ingested host IgG. Ticks excrete ingested host IgGs through their salivary glands. IgG binding proteins that are secreted at the site of feeding, assist female ticks which are feeding in the vicinity to engorge at a faster rate (Wang and Nuttall, 1995b; Wang et al., 1998). In the sialome of *R. pulchellus*, we have shown

evidence that this family of proteins are highly expressed in males and in quantities that are high enough to be detected in the MS.

4.3 Reproduction of ticks

The mouthparts of male ticks have been reported to actively engage in reproduction. Male tick saliva assists reproduction following the observation that both soft and hard ticks salivate copiously on the surface of their spermatophores before pushing them with their mouthparts into the female's genital pore (Feldman-Muhsam et al., 1970). Male saliva was postulated to avoid stickiness of the spermatophore when it is being transferred from the male genitalia into the female's genital pore. Histological studies of tick salivary glands have described male-specific alveolus and cells (Chinery, 1965; Furquim et al., 2010; Till, 1961) that were postulated to assist tick reproduction (Feldman-Muhsam et al., 1970). Whether ticks salivate during their long copulation is unknown. The relationship between the salivary gland components and sexual differences in feeding or sexual behaviour of the ticks is not well understood.

Using a differential display approach, differential male salivary gene expression was observed when males were feeding together with and without female conspecific ticks (Anyomi et al., 2006; Bior et al., 2002). Another previously reported salivary gland transcriptome of male and female R. haemaphysaloides ticks have detected, by suppression subtractive hybridization protocols, 17 male-specific transcripts that include a serine

protease, cytochrome P450, an IgG binding protein and metalloproteases (Xiang et al., 2012), in concordance with our findings. In this thesis, we propose that some male-specific salivary gland proteins play a role in tick reproduction. Among these candidates are serpins, cystatins and serine proteases which are ubiquitously found in seminal fluids from Drosophila to humans, and also metalloproteases and legumains mentioned above. The differentially male expressed cytochrome P450 may be involved in the production of prostaglandins or other eicosanoids. Male housekeeping enzymes linked to lipid metabolism may relate to additional secreted lipidic products. Male-specific mucins may also fall under this class of seminal-fluid like proteins and the male upregulated galactosyl-transferase may assist the post-translational modifications abundant in mucins. The increased expression of mucins reflected in the male transcriptome may point towards the male saliva being a lubricant as mentioned above.

4.4 Bilaris proteins

While male differentially-expressed transcripts can be hypothesized to play a role in reproduction, it is also possible that they result from males expressing a different anti-hemostatic cocktail compare to females, due to their different feeding requirements and/or to diversify their antigenic cocktail. Kunitz-type proteinase inhibitors are potentially important in anticoagulant function and may target various parts of the coagulation cascade (Corral-Rodriguez et al., 2009). We have specifically chosen the bilaris proteins from this family of inhibitors to investigate and illustrate the differential expression of

anticoagulants between male and female ticks. In this class of inhibitors, we have identified five distinct subclasses of bilaris proteins by their unique pattern of disulphide bonds present. The variations in the inter-cysteine regions suggest diverse functions for these proteins despite similar structures. They might target different proteases and their complexes, and inhibit them through different mechanisms. Through the evolution of this protein family in ticks through gene duplication, it is not unusual for molecules of the same protein scaffold to exert diverse functions (Louw et al., 2013; Mans et al., 2008a; Mans and Neitz, 2004). Likewise, proteins with different scaffold may also exhibit similar functions. This expansion of the Kunitz-type inhibitor family is likely to facilitate the long-term feeding habit of hard ticks (Dai et al., 2012).

In addition, our study have shown that there appears to be distinct pattern of expression of this group of proteins. It was evident that male ticks preferentially expressed subclasses I and IV while female expressed subclass II of the bilaris proteins. This points towards a different anti-hemostatic cocktail within the saliva that is employed by each gender. Once again, this observation may be explained by the contrast in feeding behaviour displayed by both gender.

In order to fully understand the significance of this difference on feeding between males and females, it would be ideal to study the functional aspects of these bilaris proteins. We have attempted the recombinant expression of selected bilaris proteins. However, functional tests have been unsuccessful in detecting the presence of anticoagulant activity. This may be due to several reasons including too little amounts of proteins expressed. Further optimization

would be required to increase transfection efficiency and stable expression may be more efficient than transient expression of the gene. If required, expression could be done in animal cell lines to increase yield.

Even without the functional tests of the bilaris proteins at this points, we can still do minor speculations on the activity of these proteins based on their protein sequences. A classic Kunitz-type serine protease inhibitor is the bovine pancreatic trypsin inhibitor. It inhibits trypsin through its reactive site loop Lys15(P_1)-Ala16(P_1') which is located right after the second cysteine residue Cys14. Looking at the sequences of our 12 bilaris proteins, none of the proteins has both residues conserved (Fig 3.19). However, both RpSigp-842790 and RpSigp-843116 has the conserved Lys at the P_1 position and Gly at the P_1' position of the first Kunitz domain. In addition, RpIx75-680109 has Arg and Ala at the P_1 and P_1' position of the first Kunitz domain. This may suggest the high possibility of these three proteins inhibiting the blood coagulation factors in a mechanism similar to that of BPTI.

Besides BPTI, tick-derived Kunitz-type inhibitors which have their structures elucidated include boophilin, ornithodorin and TAP. In all three, studies have shown that the N-terminal residues play a large role in anchoring and inhibiting the blood coagulation factors (see section 1.2.1). In fact, these N-terminal residues are more important than the reactive site loop for inhibition. The key residues of TAP's Tyr-Asn-Arg, ornithodorin's Ser-Leu-Asn-Val, and boophilin's Arg-Asn-Gly were not found in any of our 12 bilaris proteins. However, RpIx75-485129 has Ser-Arg-Asn as its N-terminal residues which may indicate that it might be able to be inserted into the active site of the serine

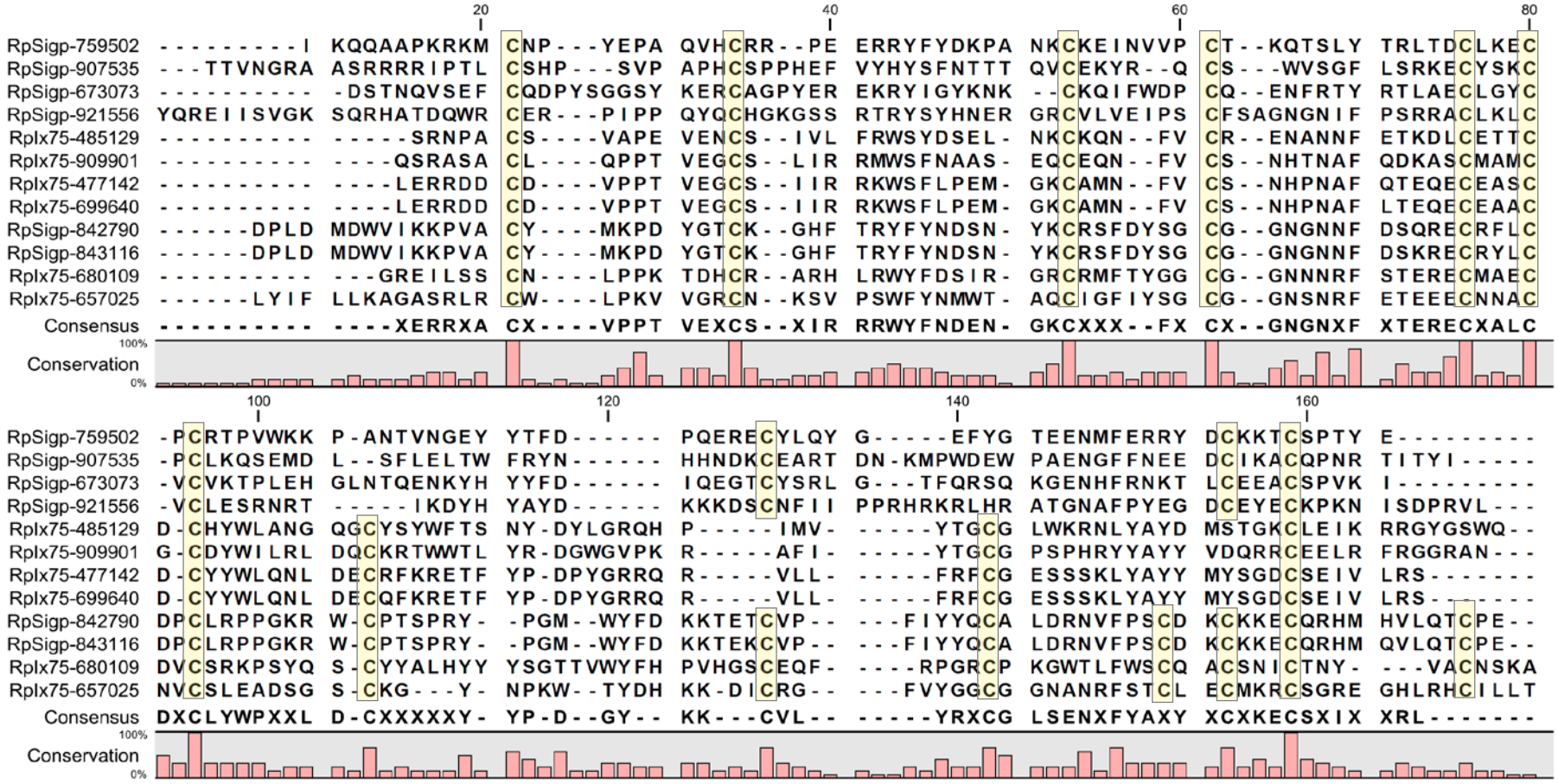

Figure 3.19. Sequence alignment of bilaris proteins. The 12 selected bilaris proteins are aligned with ClustalW. Conserved cysteine residues are highlighted in yellow. The reactive site loop for a Kunitz domain (P_1-P_1') is located after the 2nd cysteine residue on each domain.

proteases. Even so, looking at the variance of the key residues of these three known inhibitors reveals that ticks indeed have a wide diversity of peptides that have evolved to inhibit the blood coagulation factors in their own unique ways. Hence, there is a very high possibility that these bilaris proteins reported here are able to present a new mechanism of inhibiting the blood coagulation factors.

CHAPTER 5

Conclusion and Future Perspectives

5.1 Conclusion

Ticks are ectoparasites and vectors of viruses, bacteria, protozoa and nematodes. They rely exclusively on vertebrate blood for their survival. During feeding, ticks inject into their hosts, a sophisticated salivary potion that induces vasodilation, and impedes platelet aggregation, blood clotting and host immunity, thus overcoming host responses. In this study, we have shown for the first time that salivary gland extracts from female *R. pulchellus* inhibit blood coagulation factor Xa and thrombin four fold higher than that of male ticks suggesting the possibility of male *R. pulchellus* using different strategies from females to obtain their blood meals.

In an attempt to relate to their physiological disparity, we have unraveled the transcriptome and proteome of the salivary glands of male and female *R. pulchellus* separately. This thesis presents the first deep-sequencing transcriptome of ticks aimed at distinguishing salivary gender-specific transcripts, and also the sialome of *R. pulchellus* for the very first time. We have reported numerous proteins that are overexpressed in a gender-specific manner, most of which were of anti-hemostatic and immunomodulatory functions. In females, we note the abundance in expression of Kunitz-type inhibitors and lipocalins which allows the females to feed on large amounts of blood, leading to engorgement. With more blood fed, it directly translates into more eggs laid after the meal. On the other hand in males, large quantity of IgG binding proteins secreted assist the females in their blood meals. Several other classes of proteins such as cystatins, mucins, and serpins which are also highly expressed in males are hypothesized to play a role in reproduction. This also

shows how males are concerned with the perpetuation of his own genes. Hence, we can see that both the males and females are directed to the ultimate aim of the continuation of species.

We have also analyzed a single class of Kunitz-type serine protease inhibitors in detail, and further classified them into five subclasses. Quantitative PCR data indicate that male and female *R. pulchellus* selectively express certain subclasses of these proteins. Although attempts at recombinant expression of these bilaris proteins were unsuccessful, sequence analysis hints at a high possibility of these proteins' ability to inhibit serine proteases, and perhaps with a novel mechanism.

This approach of examining the male and female sialomes of ticks independently has opened up opportunities to discover new salivary proteins. It also allows us to have an initial look into various strategies deployed by each sex enabling them to feed successfully off their hosts, and determine candidate male salivary proteins that may assist reproduction. This annotated database also represents a mining field for potential drug lead compounds for pharmaceutical uses as it reports numerous novel putative anti-hemostatic compounds.

5.2 Future perspectives

With this newly reported sialome of *R. pulchellus*, many avenues have been opened up for future work on various classes of proteins. We first outline the immediate work that could be performed on the existing results that we report, followed by a big picture with other experiments that could be performed based on this sialome.

5.2.1 Functional studies on bilaris proteins

In this study, we have reported new subclasses of bilaris proteins. Although the loss of the C2-C4 disulphide bond in the second Kunitz domain (subclass II) have been previously reported in other inhibitors (e.g. ixolaris), the loss of the C3-C5 bond have never been reported to date. Although most Kunitz-type proteins have been reported to display anti-hemostatic properties, most are known to inhibit their target proteases with novel and non-canonical methods (Corral-Rodriguez et al., 2009; Louw et al., 2013). It has been suggested that the loss of the C2-C4 disulphide bond probably confers its unconventional binding mechanism (Louw et al., 2013). Hence, it would be interest to firstly, do functional studies on all five subclasses of bilaris proteins, and investigate their structure-function relationships with their respective targets in the blood coagulation cascade.

As previously mentioned, as the source of these ticks and their salivary gland extracts are limited, molecular methods can be applied to amplify the corresponding gene from the cDNA libraries of the ticks. These genes encoding

for the bilaris proteins can be cloned into expression vectors, and transfected into insect cell lines for expression. Once the bilaris proteins have been expressed and purified, screening assays for the inhibition of various blood coagulation factors can be performed. The functional studies of these proteins may help us delineate additional aspects of differential feeding habits between male and female ticks.

5.2.2 Functional studies on monolaris protein

The bovine pancreatic trypsin inhibitor (BPTI) is a classic single Kunitz domain (monolaris) inhibitor that was isolated more than four decades ago (Ascenzi et al., 2003; Burck et al., 1967; Cerwinsky et al., 1967), and now developed into a drug aprotinin. BPTI is a small globular protein, 58 amino acid residues in length. In the course of analysis of the Kunitz-type proteins, we found a unique monolaris protein in the transcriptome and an isoform of it (RpSigp-799633 and RpIx75-799634). The sequence alignment of these two proteins with BPTI is shown below:

```
>BPTI                    RPDFCLEPPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCGGA
>RpIx75-799634 KRPKEDLLAYPKQGNCKAAFVRYYCMESG-----------CKQFTG--CYRGGFRSRKE--CTKICPNSGKKRPRWPHSQRERSTGLQE
>RpSigp-799633 KHHKDDLLAYPKQGKCKAAFVRHYCMESG-----------CKQFTG--CYRGGFRSRKE--CTKTCPNFRKKRPRWPHNERERSTGLQE
```

Even though this monolaris protein is longer in length than BPTI, it is unique as the regions between C2 and C3, C3 and C4, and C4 and C5, has deletions, resulting in it being significantly shorter than that of BPTI. Thus, it would be interesting to express this particular protein to see how it functions as

an inhibitor. Such short Kunitz domain region would perhaps improve the binding to the target protease.

5.2.3 Time-dependent expression of salivary proteins

As ticks feed on their host, their salivary glands are stimulated, resulting in the increased expression of salivary proteins over the period of a few days. Studies have shown that the protein content in the salivary glands before and after feeding increased up to 10 times, with only selected proteins being preferentially synthesised during this time (Slovak et al., 2000). However, it is not known exactly which proteins are the ones that are upregulated during feeding.

With the transcriptome of *R. pulchellus* now elucidated, we can perform proteomic studies on the relative increased expression of individual proteins in the salivary glands. The isobaric tags for relative and absolute quantitation (iTRAQ) technology could be applied for these experiments, where up to eight biological samples could be studied simultaneously. Briefly, the experiment could be designed such that there will be four samples of salivary gland extracts from each sex – unfed, 2-day fed, 4-day fed, and 6-day fed. These eight samples will be processed and individually labelled with their unique isobaric tags. The samples are then pooled and analyzed by tandem mass spectrometry (MS/MS). Upon fragmentation of each peptide, the isobaric tag will be fragmented to generate a reporter ion, which is different for each sample, such that the relative quantity of the particular peptide can be calculated. With this,

the upregulation of specific proteins in the salivary glands will be identified. Comparisons can also be made between the upregulated proteins in male and female ticks.

References

Anatriello, E., Ribeiro, J.M., de Miranda-Santos, I.K., Brandao, L.G., Anderson, J.M., Valenzuela, J.G., Maruyama, S.R., Silva, J.S., and Ferreira, B.R. (2010). An insight into the sialotranscriptome of the brown dog tick, *Rhipicephalus sanguineus*. BMC genomics *11*, 450.
Andersen, J.F., and Montfort, W.R. (2000). The crystal structure of nitrophorin 2. A trifunctional antihemostatic protein from the saliva of Rhodnius prolixus. JBiolChem *275*, 30496-30503.
Arca, B., Lombardo, F., Valenzuela, J.G., Francischetti, I.M., Marinotti, O., Coluzzi, M., and Ribeiro, J.M. (2005). An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. J Exp Biol *208*, 3971-3986.
Ariens, R.A., Lai, T.S., Weisel, J.W., Greenberg, C.S., and Grant, P.J. (2002). Role of factor XIII in fibrin clot formation and effects of genetic polymorphisms. Blood *100*, 743-754.
Arolas, J.L., Lorenzo, J., Rovira, A., Castella, J., Aviles, F.X., and Sommerhoff, C.P. (2005a). A carboxypeptidase inhibitor from the tick *Rhipicephalus bursa*: isolation, cDNA cloning, recombinant expression, and characterization. The Journal of biological chemistry *280*, 3441-3448.
Arolas, J.L., Popowicz, G.M., Lorenzo, J., Sommerhoff, C.P., Huber, R., Aviles, F.X., and Holak, T.A. (2005b). The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode. Journal of molecular biology *350*, 489-498.
Ascenzi, P., Bocedi, A., Bolognesi, M., Spallarossa, A., Coletta, M., De Cristofaro, R., and Menegatti, E. (2003). The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): a milestone protein. Current protein & peptide science *4*, 231-251.
Baker, B.S., and Wolfner, M.F. (1988). A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. Genes Dev *2*, 477-489.
Bao, W., Jurka, M.G., Kapitonov, V.V., and Jurka, J. (2009). New superfamilies of eukaryotic DNA transposons and their internal divisions. Molecular biology and evolution *26*, 983-993.
Becker, B.F., Heindl, B., Kupatt, C., and Zahler, S. (2000). Endothelial function and hemostasis. Zeitschrift fur Kardiologie *89*, 160-167.
Bergman, D.K., Palmer, M.J., Caimano, M.J., Radolf, J.D., and Wikel, S.K. (2000). Isolation and molecular cloning of a secreted immunosuppressant protein from *Dermacentor andersoni* salivary gland. J Parasitol *86*, 516-525.
Bilhartz, D.L., Tindall, D.J., and Oesterling, J.E. (1991). Prostate-specific antigen and prostatic acid phosphatase: biomolecular and physiologic characteristics. Urology *38*, 95-102.
Birol, I., Jackman, S.D., Nielsen, C.B., Qian, J.Q., Varhol, R., Stazyk, G., Morin, R.D., Zhao, Y., Hirst, M., Schein, J.E*., et al.* (2009). De novo transcriptome assembly with ABySS. Bioinformatics *25*, 2872-2877.
Bishop, R., Lambson, B., Wells, C., Pandit, P., Osaso, J., Nkonge, C., Morzaria, S., Musoke, A., and Nene, V. (2002). A cement protein of the tick *Rhipicephalus appendiculatus*, located in the secretory e cell granules of the type III salivary gland acini, induces strong antibody responses in cattle. International journal for parasitology *32*, 833-842.
Blandin, S., and Levashina, E.A. (2004). Thioester-containing proteins and insect immunity. Molecular immunology *40*, 903-908.
Bowman, A.S., Gengler, C.L., Surdick, M.R., Zhu, K., Essenberg, R.C., Sauer, J.R., and Dillwith, J.W. (1997). A novel phospholipase A2 activity in saliva of the lone star tick, *Amblyomma americanum* (L.). Experimental parasitology *87*, 121-132.

Brankamp, R.G., Blankenship, D.T., Sunkara, P.S., and Cardin, A.D. (1990). Ghilantens: anticoagulant-antimetastatic proteins from the South American leech, Haementeria ghilianii. JLab ClinMed *115*, 89-97.
Brass, L.F. (2003). Thrombin and platelet activation. Chest *124*, 18s-25s.
Burck, P.J., Hamill, R.L., Cerwinsky, E.W., and Grinnan, E.L. (1967). Acidic bovine pancreatic trypsin inhibitor. II. Specificity of inhibition. Biochemistry *6*, 3180-3184.
Calvete, J.J., Moreno-Murciano, M.P., Theakston, R.D., Kisiel, D.G., and Marcinkiewicz, C. (2003). Snake venom disintegrins: novel dimeric disintegrins and structural diversification by disulphide bond engineering. The Biochemical journal *372*, 725-734.
Calvo, E., Pham, V.M., Lombardo, F., Arca, B., and Ribeiro, J.M. (2006). The sialotranscriptome of adult male *Anopheles gambiae* mosquitoes. Insect Biochem Mol Biol *36*, 570-575.
Campos, I.T., Amino, R., Sampaio, C.A., Auerswald, E.A., Friedrich, T., Lemaire, H.G., Schenkman, S., and Tanaka, A.S. (2002). Infestin, a thrombin inhibitor presents in Triatoma infestans midgut, a Chagas' disease vector: gene cloning, expression and characterization of the inhibitor. Insect BiochemMolBiol *32*, 991-997.
Cappello, M., Vlasuk, G.P., Bergum, P.W., Huang, S., and Hotez, P.J. (1995). Ancylostoma caninum anticoagulant peptide: a hookworm-derived inhibitor of human coagulation factor Xa. ProcNatlAcadSciUSA *92*, 6152-6156.
Carneiro-Lobo, T.C., Konig, S., Machado, D.E., Nasciutti, L.E., Forni, M.F., Francischetti, I.M., Sogayar, M.C., and Monteiro, R.Q. (2009). Ixolaris, a tissue factor inhibitor, blocks primary tumor growth and angiogenesis in a glioblastoma model. Journal of thrombosis and haemostasis : JTH *7*, 1855-1864.
Castaneda, O., and Harvey, A.L. (2009). Discovery and characterization of cnidarian peptide toxins that affect neuronal potassium ion channels. Toxicon : official journal of the International Society on Toxinology *54*, 1119-1124.
Cerwinsky, E.W., Burck, P.J., and Grinnan, E.L. (1967). Acidic bovine pancreatic trypsin inhibitor. I. Purification and physical characterization. Biochemistry *6*, 3175-3179.
Chalaire, K.C., Kim, T.K., Garcia-Rodriguez, H., and Mulenga, A. (2011). *Amblyomma americanum* (L.) (Acari: Ixodidae) tick salivary gland serine protease inhibitor (serpin) 6 is secreted into tick saliva during tick feeding. The Journal of experimental biology *214*, 665-673.
Champagne, D.E. (2005). Antihemostatic molecules from saliva of blood-feeding arthropods. Pathophysiology of haemostasis and thrombosis *34*, 221-227.
Champagne, D.E., Smartt, C.T., Ribeiro, J.M., and James, A.A. (1995). The salivary gland-specific apyrase of the mosquito *Aedes aegypti* is a member of the 5'-nucleotidase family. Proc Natl Acad Sci U S A *92*, 694-698.
Chen, X., Yu, Z., Guo, L., Li, L., Meng, H., Wang, D., Liu, R., and Liu, J. (2012a). Life cycle of Haemaphysalis doenitzi (Acari: Ixodidae) under laboratory conditions and its phylogeny based on mitochondrial 16S rDNA. Experimental & applied acarology *56*, 143-150.
Chen, Z., Li, Y., Liu, Z., Yang, J., and Yin, H. (2012b). The life cycle of Hyalomma rufipes (Acari: Ixodidae) under laboratory conditions. Experimental & applied acarology *56*, 85-92.
Chmelar, J., Oliveira, C.J., Rezacova, P., Francischetti, I.M., Kovarova, Z., Pejler, G., Kopacek, P., Ribeiro, J.M., Mares, M., Kopecky, J.*, et al.* (2011). A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. Blood *117*, 736-744.
Chopin, V., Salzet, M., Baert, J., Vandenbulcke, F., Sautiere, P.E., Kerckaert, J.P., and Malecha, J. (2000). Therostasin, a novel clotting factor Xa inhibitor from the rhynchobdellid leech, Theromyzon tessulatum. JBiolChem *275*, 32701-32707.
Clemetson, K.J., and Clemetson, J.M. (2001). Platelet collagen receptors. Thrombosis and haemostasis *86*, 189-197.

Corral-Rodriguez, M.A., Macedo-Ribeiro, S., Barbosa Pereira, P.J., and Fuentes-Prior, P. (2009). Tick-derived Kunitz-type inhibitors as antihemostatic factors. Insect Biochem Mol Biol *39*, 579-595.
Dai, S.X., Zhang, A.D., and Huang, J.F. (2012). Evolution, expansion and expression of the Kunitz/BPTI gene family associated with long-term blood feeding in Ixodes Scapularis. BMC evolutionary biology *12*, 4.
Davie, E.W., Fujikawa, K., and Kisiel, W. (1991). The coagulation cascade: initiation, maintenance, and regulation. Biochemistry *30*, 10363-10370.
Deruaz, M., Frauenschuh, A., Alessandri, A.L., Dias, J.M., Coelho, F.M., Russo, R.C., Ferreira, B.R., Graham, G.J., Shaw, J.P., Wells, T.N*., et al.* (2008). Ticks produce highly selective chemokine binding proteins with antiinflammatory activity. The Journal of experimental medicine *205*, 2019-2031.
Dickinson, R.G., O'Hagan, J.E., Shotz, M., Binnington, K.C., and Hegarty, M.P. (1976). Prostaglandin in the saliva of the cattle tick *Boophilus microplus*. Aust J Exp Biol Med Sci *54*, 475-486.
Dobrovolsky, A.B., and Titaeva, E.V. (2002). The fibrinolysis system: regulation of activity and physiologic functions of its main components. Biochemistry Biokhimiia *67*, 99-108.
Duckert, P., Brunak, S., and Blom, N. (2004). Prediction of proprotein convertase cleavage sites. Protein Eng Des Sel *17*, 107-112.
Dunwiddie, C., Thornberry, N.A., Bull, H.G., Sardana, M., Friedman, P.A., Jacobs, J.W., and Simpson, E. (1989). Antistasin, a leech-derived inhibitor of factor Xa. Kinetic analysis of enzyme inhibition and identification of the reactive site. JBiolChem *264*, 16694-16699.
Faccini, J.L., Cardoso, A.C., Onofrio, V.C., Labruna, M.B., and Barros-Battesti, D.M. (2010). The life cycle of Amblyomma auricularium (Acari: Ixodidae) using rabbits (Oryctolagus cuniculus) as experimental host. Experimental & applied acarology *50*, 71-77.
Feldman-Muhsam, B., Borut, S., and Saliternik-Givant, S. (1970). Salivary secretion of the male tick during copulation. Journal of insect physiology *16*, 1945-1949.
Flower, D.R., North, A.C., and Sansom, C.E. (2000). The lipocalin protein family: structural and sequence overview. Biochim Biophys Acta *1482*, 9-24.
Fogaca, A.C., Almeida, I.C., Eberlin, M.N., Tanaka, A.S., Bulet, P., and Daffre, S. (2005). Ixodidin, a novel antimicrobial peptide from the hemocytes of the cattle tick *Boophilus microplus* with inhibitory activity against serine proteinases. Peptides *27*, 667-674.
Fogaca, A.C., Lorenzini, D.M., Kaku, L.M., Esteves, E., Bulet, P., and Daffre, S. (2004). Cysteine-rich antimicrobial peptides of the cattle tick *Boophilus microplus*: isolation, structural characterization and tissue expression profile. Dev Comp Immunol *28*, 191-200.
Francischetti, I.M. (2010). Platelet aggregation inhibitors from hematophagous animals. Toxicon : official journal of the International Society on Toxinology *56*, 1130-1144.
Francischetti, I.M., Anderson, J.M., Manoukis, N., Pham, V.M., and Ribeiro, J.M. (2011). An insight into the sialotranscriptome and proteome of the coarse bontlegged tick, *Hyalomma marginatum rufipes*. Journal of proteomics *74*, 2892-2908.
Francischetti, I.M., Mather, T.N., and Ribeiro, J.M. (2003). Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrin(ogen)lytic activities in the saliva of the Lyme disease tick vector *Ixodes scapularis*. Biochem Biophys Res Commun *305*, 869-875.
Francischetti, I.M., Mather, T.N., and Ribeiro, J.M. (2004). Penthalaris, a novel recombinant five-Kunitz tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick vector of Lyme disease, Ixodes scapularis. ThrombHaemost *91*, 886-898.
Francischetti, I.M., Mather, T.N., and Ribeiro, J.M. (2005a). Tick saliva is a potent inhibitor of endothelial cell proliferation and angiogenesis. Thrombosis and haemostasis *94*, 167-174.

Francischetti, I.M., My Pham, V., Mans, B.J., Andersen, J.F., Mather, T.N., Lane, R.S., and Ribeiro, J.M. (2005b). The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). Insect Biochem Mol Biol *35*, 1142-1161.
Francischetti, I.M., Sa-Nunes, A., Mans, B.J., Santos, I.M., and Ribeiro, J.M. (2009a). The role of saliva in tick feeding. Front Biosci *14*, 2051-2088.
Francischetti, I.M., Valenzuela, J.G., Andersen, J.F., Mather, T.N., and Ribeiro, J.M. (2002a). Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, Ixodes scapularis: identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex. Blood *99*, 3602-3612.
Francischetti, I.M., Valenzuela, J.G., Andersen, J.F., Mather, T.N., and Ribeiro, J.M. (2002b). Ixolaris, a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick, Ixodes scapularis: identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa/tissue factor complex. Blood *99*, 3602-3612.
Francischetti, I.M.B., Sá-Nunes, A., Mans, B.J., Santos, I.M., and Ribeiro, J.M.C. (2009b). The role of saliva in tick feeding. Frontiers in Biosciences *14*, 2051-2088.
Frauenschuh, A., Power, C.A., Deruaz, M., Ferreira, B.R., Silva, J.S., Teixeira, M.M., Dias, J.M., Martin, T., Wells, T.N., and Proudfoot, A.E. (2007). Molecular cloning and characterization of a highly selective chemokine-binding protein from the tick *Rhipicephalus sanguineus*. The Journal of biological chemistry *282*, 27250-27258.
Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H.G., Hoffken, H.W., Reuschenbach, P., Otte, M., and Dodt, J. (1993). A Kazal-type inhibitor with thrombin specificity from Rhodnius prolixus. JBiolChem *268*, 16216-16222.
Fry, B.G., Roelants, K., Champagne, D.E., Scheib, H., Tyndall, J.D., King, G.F., Nevalainen, T.J., Norman, J.A., Lewis, R.J., Norton, R.S*., et al.* (2009). The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual review of genomics and human genetics *10*, 483-511.
Furie, B., and Furie, B.C. (2008). Mechanisms of thrombus formation. The New England journal of medicine *359*, 938-949.
Galperin, M.Y., and Koonin, E.V. (2004). 'Conserved hypothetical' proteins: prioritization of targets for experimental study. Nucleic acids research *32*, 5452-5463.
Gaspar, A.R., Joubert, A.M., Crause, J.C., and Neitz, A.W. (1996). Isolation and characterization of an anticoagulant from the salivary glands of the tick, Ornithodoros savignyi (Acari: Argasidae). ExpApplAcarol *20*, 583-598.
Gasperik J, K.P., Slovák M, Hajnická V,Fuchsberger N, Nuttall PA (2000). Differences in salivary gland extracts derived from male and/or female ticks. Paper presented at: Ticks and Tick-borne Pathogens: Into the 21st Century (Institute of Zoology, Slovak Academy of Sciences, Bratislava, Slovakia).
Gong, H., Zhou, J., Liao, M., Hatta, T., Harnnoi, T., Umemiya, R., Inoue, N., Xuan, X., and Fujisaki, K. (2007). Characterization of a carboxypeptidase inhibitor from the tick *Haemaphysalis longicornis*. Journal of insect physiology *53*, 1079-1087.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q*., et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology *29*, 644-652.
Grutter, M.G. (1994). Proteinase inhibitors: another new fold. Structure *2*, 575-576.
Grutter, M.G., Priestle, J.P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J., and Stone, S.R. (1990). Crystal structure of the thrombin-hirudin complex: a novel mode of serine protease inhibition. EMBO J *9*, 2361-2365.
Harrison, L.M., Nerlinger, A., Bungiro, R.D., Cordova, J.L., Kuzmic, P., and Cappello, M. (2002). Molecular characterization of Ancylostoma inhibitors of coagulation factor Xa. Hookworm anticoagulant activity in vitro predicts parasite bloodfeeding in vivo. JBiolChem *277*, 6223-6229.

Hati, R., Mitra, P., Sarker, S., and Bhattacharyya, K.K. (1999). Snake venom hemorrhagins. Crit Rev Toxicol *29*, 1-19.
Higgs, G.A., Vane, J.R., Hart, R.J., Porter, C., and Wilson, R.G. (1976). Prostaglandins in the saliva of the cattle tick, *Boophilus microplus*(Canestrini) (Acarina, Ixodidae). Bull Ent Res *66*, 665-670.
Horackova, J., Rudenko, N., Golovchenko, M., Havlikova, S., and Grubhoffer, L. (2010). IrML - a gene encoding a new member of the ML protein family from the hard tick, *Ixodes ricinus*. J Vector Ecol *35*, 410-418.
Horka, H., Staudt, V., Klein, M., Taube, C., Reuter, S., Dehzad, N., Andersen, J.F., Kopecky, J., Schild, H., Kotsyfakis, M*., et al.* (2012). The tick salivary protein sialostatin L inhibits the Th9-derived production of the asthma-promoting cytokine IL-9 and is effective in the prevention of experimental asthma. Journal of immunology *188*, 2669-2676.
Huang, T.F., and Niewriarowski, S. (1994). Disintegrins: The naturally-occurring antagonists of platelet fibrinogen receptors. J Toxicol Toxin Rev *13*, 253-273.
Isawa, H., Yuda, M., Yoneda, K., and Chinzei, Y. (2000). The insect salivary protein, prolixin-S, inhibits factor IXa generation and Xase complex formation in the blood coagulation pathway. JBiolChem *275*, 6636-6641.
Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. Trends Biochem Sci *23*, 403-405.
Jennings, L.K. (2009). Mechanisms of platelet activation: need for new strategies to protect against platelet-mediated atherothrombosis. Thrombosis and haemostasis *102*, 248-257.
Julenius, K., Molgaard, A., Gupta, R., and Brunak, S. (2005). Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology *15*, 153-164.
Kapitonov, V.V., and Jurka, J. (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. Nature reviews Genetics *9*, 411-412; author reply 414.
Karim, S., Singh, P., and Ribeiro, J.M. (2011). A deep insight into the sialotranscriptome of the gulf coast tick, Amblyomma maculatum. PloS one *6*, e28525.
Kariuki, E.K., Penzhorn, B.L., and Horak, I.G. (2012). Ticks (Acari: Ixodidae) infesting cattle and African buffaloes in the Tsavo conservation area, Kenya. The Onderstepoort journal of veterinary research *79*, E1-4.
Kaufman, W.R. (2004). Assuring paternity in a promiscuous world: are there lessons for ticks among the insects ? Parasitology *129*, S145-S160.
Kelley, D.R., and Rinn, J.L. (2012). Transposable elements reveal a stem cell specific class of long noncoding RNAs. Genome biology *13*, R107.
Kim, Y.S., Ryu, J.H., Han, S.J., Choi, K.H., Nam, K.B., Jang, I.H., Lemaitre, B., Brey, P.T., and Lee, W.J. (2000). Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. J Biol Chem *275*, 32721-32727.
Kiszewski, A.E., Matuschka, F.R., and Spielman, A. (2001). Mating strategies and spermiogenesis in ixodid ticks. Annual review of entomology *46*, 167-182.
Kivirikko, K.I., and Pihlajaniemi, T. (1998). Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. Adv Enzymol Relat Areas Mol Biol *72*, 325-398.
Klein, T., and Bischoff, R. (2011). Active metalloproteases of the A Disintegrin and Metalloprotease (ADAM) family: biological function and structure. Journal of proteome research *10*, 17-33.
Koh, C.Y., Kazimirova, M., Trimnell, A., Takac, P., Labuda, M., Nuttall, P.A., and Kini, R.M. (2007). Variegin, a novel fast and tight binding thrombin inhibitor from the tropical bont tick. JBiolChem *282*, 29101-29113.
Kopacek, P., Hajdusek, O., Buresova, V., and Daffre, S. (2010). Tick innate immunity. Adv Exp Med Biol *708*, 137-162.

Kotsyfakis, M., Anderson, J.M., Andersen, J.F., Calvo, E., Francischetti, I.M., Mather, T.N., Valenzuela, J.G., and Ribeiro, J.M. (2008). Cutting edge: Immunity against a "silent" salivary antigen of the Lyme vector *Ixodes scapularis* impairs its ability to feed. Journal of immunology *181*, 5209-5212.
Kotsyfakis, M., Horka, H., Salat, J., and Andersen, J.F. (2010). The crystal structures of two salivary cystatins from the tick *Ixodes scapularis* and the effect of these inhibitors on the establishment of *Borrelia burgdorferi* infection in a murine model. Molecular microbiology *77*, 456-470.
Kotsyfakis, M., Karim, S., Andersen, J.F., Mather, T.N., and Ribeiro, J.M. (2007). Selective cysteine protease inhibition contributes to blood-feeding success of the tick *Ixodes scapularis*. The Journal of biological chemistry *282*, 29256-29263.
Kotsyfakis, M., Sa-Nunes, A., Francischetti, I.M., Mather, T.N., Andersen, J.F., and Ribeiro, J.M. (2006). Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. J Biol Chem *281*, 26298-26307.
LaFlamme, B.A., Ram, K.R., and Wolfner, M.F. (2012). The *Drosophila melanogaster* seminal fluid protease "seminase" regulates proteolytic and post-mating reproductive processes. PLoS genetics *8*, e1002435.
Lai, R., Takeuchi, H., Jonczy, J., Rees, H.H., and Turner, P.C. (2004). A thrombin inhibitor from the ixodid tick, Amblyomma hebraeum. Gene *342*, 243-249.
Laskowski, M., Jr., and Kato, I. (1980). Protein inhibitors of proteinases. AnnuRevBiochem *49*, 593-626.
Lehane, M.J. (2005). The biology of blood-sucking in insects, 2 edn (Cambridge: Cambridge University Press).
Liao, M., Zhou, J., Gong, H., Boldbaatar, D., Shirafuji, R., Battur, B., Nishikawa, Y., and Fujisaki, K. (2009). Hemalin, a thrombin inhibitor isolated from a midgut cDNA library from the hard tick Haemaphysalis longicornis. JInsect Physiol *55*, 164-173.
Liu, C.C., Brustad, E., Liu, W., and Schultz, P.G. (2007). Crystal structure of a biosynthetic sulfo-hirudin complexed to thrombin. JAmChemSoc *129*, 10648-10649.
Louw, E., van der Merwe, N.A., Neitz, A.W., and Maritz-Olivier, C. (2013). Evolution of the tissue factor pathway inhibitor-like Kunitz domain-containing protein family in Rhipicephalus microplus. International journal for parasitology *43*, 81-94.
Ma, M., Guan, G., Chen, Z., Liu, Z., Liu, A., Gou, H., Ren, Q., Li, Y., Niu, Q., Yang, J.*, et al.* (2013). The life cycle of Haemaphysalis qinghaiensis (Acari: Ixodidae) ticks under laboratory conditions. Experimental & applied acarology *59*, 493-500.
Macedo-Ribeiro, S., Almeida, C., Calisto, B.M., Friedrich, T., Mentele, R., Sturzebecher, J., Fuentes-Prior, P., and Pereira, P.J. (2008). Isolation, cloning and structural characterisation of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. PLoSOne *3*, e1624.
Mans, B.J., Andersen, J.F., Francischetti, I.M., Valenzuela, J.G., Schwan, T.G., Pham, V.M., Garfield, M.K., Hammer, C.H., and Ribeiro, J.M. (2008a). Comparative sialomics between hard and soft ticks: implications for the evolution of blood-feeding behavior. Insect Biochem Mol Biol *38*, 42-58.
Mans, B.J., Andersen, J.F., Schwan, T.G., and Ribeiro, J.M. (2008b). Characterization of anti-hemostatic factors in the argasid, Argas monolakensis: implications for the evolution of blood-feeding in the soft tick family. Insect BiochemMolBiol *38*, 22-41.
Mans, B.J., and Francischetti, I.M.B. (2010). Sialomic perspectives on the evolution of blood-feeding behavior in Arthropods: Future therapeutics by natural design. In Toxins and Hemostasis: From Bench to Bed-side, R.M. Kini, K.J. Clemetson, F.S. Markland, M.A. McLane, and T. Morita, eds. (Springer).

Mans, B.J., Louw, A.I., and Neitz, A.W. (2002). Amino acid sequence and structure modeling of savignin, a thrombin inhibitor from the tick, Ornithodoros savignyi. Insect BiochemMolBiol *32*, 821-828.
Mans, B.J., and Neitz, A.W.H. (2004). Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochemistry and Molecular Biology *34*, 1-17.
Mans, B.J., and Ribeiro, J.M. (2008a). Function, mechanism and evolution of the moubatin-clade of soft tick lipocalins. Insect Biochem Mol Biol *38*, 841-852.
Mans, B.J., and Ribeiro, J.M. (2008b). A novel clade of cysteinyl leukotriene scavengers in soft ticks. Insect Biochem Mol Biol *38*, 862-870.
Mans, B.J., Ribeiro, J.M., and Andersen, J.F. (2008c). Structure, function, and evolution of biogenic amine-binding proteins in soft ticks. The Journal of biological chemistry *283*, 18721-18733.
Marino, J.P., Jr. (2009). Soluble epoxide hydrolase, a target with multiple opportunities for cardiovascular drug discovery. Current topics in medicinal chemistry *9*, 452-463.
Markwardt, F. (1994). The development of hirudin as an antithrombotic drug. ThrombRes *74*, 1-23.
Maruyama, S.R., Anatriello, E., Anderson, J.M., Ribeiro, J.M., Brandao, L.G., Valenzuela, J.G., Ferreira, B.R., Garcia, G.R., Szabo, M.P., Patel, S*., et al.* (2010). The expression of genes coding for distinct types of glycine-rich proteins varies according to the biology of three metastriate ticks, *Rhipicephalus (Boophilus) microplus*, *Rhipicephalus sanguineus* and *Amblyomma cajennense*. BMC genomics *11*, 363.
Mende, K., Petoukhova, O., Koulitchkova, V., Schaub, G.A., Lange, U., Kaufmann, R., and Nowak, G. (1999). Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect. Dipetalogaster maximus cDNA cloning, expression and characterization. EurJBiochem *266*, 583-590.
Mieszczanek, J., Harrison, L.M., and Cappello, M. (2004a). Ancylostoma ceylanicum anticoagulant peptide-1: role of the predicted reactive site amino acid in mediating inhibition of coagulation factors Xa and VIIa. MolBiochemParasitol *137*, 151-159.
Mieszczanek, J., Harrison, L.M., Vlasuk, G.P., and Cappello, M. (2004b). Anticoagulant peptides from Ancylostoma caninum are immunologically distinct and localize to separate structures within the adult hookworm. MolBiochemParasitol *133*, 319-323.
Ministry of Health, S. (2012). Principle causes of death. In Singapore health facts (Ministry of Health, Singapore).
Monroe, D.M., and Hoffman, M. (2006). What does it take to make the perfect clot? Arteriosclerosis, thrombosis, and vascular biology *26*, 41-48.
Monteiro, R.Q. (2005). Targeting exosites on blood coagulation proteases. Anais da Academia Brasileira de Ciencias *77*, 275-280.
Monteiro, R.Q., Rezaie, A.R., Bae, J.S., Calvo, E., Andersen, J.F., and Francischetti, I.M. (2008a). Ixolaris binding to factor X reveals a precursor state of factor Xa heparin-binding exosite. Protein science : a publication of the Protein Society *17*, 146-153.
Monteiro, R.Q., Rezaie, A.R., Bae, J.S., Calvo, E., Andersen, J.F., and Francischetti, I.M. (2008b). Ixolaris binding to factor X reveals a precursor state of factor Xa heparin-binding exosite. Protein Sci *17*, 146-153.
Monteiro, R.Q., Rezaie, A.R., Ribeiro, J.M., and Francischetti, I.M. (2005a). Ixolaris: a factor Xa heparin-binding exosite inhibitor. BiochemJ *387*, 871-877.
Monteiro, R.Q., Rezaie, A.R., Ribeiro, J.M., and Francischetti, I.M. (2005b). Ixolaris: a factor Xa heparin-binding exosite inhibitor. The Biochemical journal *387*, 871-877.
Moura, A.A., Souza, C.E., Stanley, B.A., Chapman, D.A., and Killian, G.J. (2010). Proteomics of cauda epididymal fluid from mature Holstein bulls. Journal of proteomics *73*, 2006-2020.
Murakami, M.T., Rios-Steiner, J., Weaver, S.E., Tulinsky, A., Geiger, J.H., and Arni, R.K. (2007). Intermolecular interactions and characterization of the novel factor Xa exosite involved in

macromolecular recognition and inhibition: crystal structure of human Gla-domainless factor Xa complexed with the anticoagulant protein NAPc2 from the hematophagous nematode Ancylostoma caninum. JMolBiol *366*, 602-610.
Narasimhan, S., Koski, R.A., Beaulieu, B., Anderson, J.F., Ramamoorthi, N., Kantor, F., Cappello, M., and Fikrig, E. (2002). A novel family of anticoagulants from the saliva of *Ixodes scapularis*. Insect Mol Biol *11*, 641-650.
Nazareth, R.A., Tomaz, L.S., Ortiz-Costa, S., Atella, G.C., Ribeiro, J.M., Francischetti, I.M., and Monteiro, R.Q. (2006). Antithrombotic properties of Ixolaris, a potent inhibitor of the extrinsic pathway of the coagulation cascade. Thrombosis and haemostasis *96*, 7-13.
Neeper, M.P., Waxman, L., Smith, D.E., Schulman, C.A., Sardana, M., Ellis, R.W., Schaffer, L.W., Siegl, P.K., and Vlasuk, G.P. (1990). Characterization of recombinant tick anticoagulant peptide. A highly selective inhibitor of blood coagulation factor Xa. JBiolChem *265*, 17746-17752.
Nielsen, H., Brunak, S., and von Heijne, G. (1999). Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein engineering *12*, 3-9.
Nienaber, J., Gaspar, A.R., and Neitz, A.W. (1999). Savignin, a potent thrombin inhibitor isolated from the salivary glands of the tick Ornithodoros savignyi (Acari: Argasidae). ExpParasitol *93*, 82-91.
Nieswandt, B., Pleines, I., and Bender, M. (2011). Platelet adhesion and activation mechanisms in arterial thrombosis and ischaemic stroke. Journal of thrombosis and haemostasis : JTH *9 Suppl 1*, 92-104.
Nunn, M.A., Sharma, A., Paesen, G.C., Adamson, S., Lissina, O., Willis, A.C., and Nuttall, P.A. (2005). Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. Journal of immunology *174*, 2084-2091.
Nutt, E., Gasic, T., Rodkey, J., Gasic, G.J., Jacobs, J.W., Friedman, P.A., and Simpson, E. (1988). The amino acid sequence of antistasin. A potent inhibitor of factor Xa reveals a repeated internal structure. JBiolChem *263*, 10162-10167.
Ogilvie, A., Blasius, R., Schulze-Lohoff, E., and Sterzel, R.B. (1996). Adenine dinucleotides: a novel class of signalling molecules. Journal of autonomic pharmacology *16*, 325-328.
Oliveira, C.J., Sa-Nunes, A., Francischetti, I.M., Carregaro, V., Anatriello, E., Silva, J.S., de Miranda Santos, I.K., Ribeiro, J.M., and Ferreira, B.R. (2011). Deconstructing Tick Saliva: Non protein molecules with potent immunomodullatory properties. The Journal of biological chemistry *286*, 10960-10969.
Otlewski, J., Jelen, F., Zakrzewska, M., and Oleksy, A. (2005). The many faces of protease-protein inhibitor interaction. EMBO J *24*, 1303-1310.
Paesen, G.C., Adams, P.L., Harlos, K., Nuttall, P.A., and Stuart, D.I. (1999). Tick histamine-binding proteins: isolation, cloning, and three- dimensional structure. Mol Cell *3*, 661-671.
Paesen, G.C., Adams, P.L., Nuttall, P.A., and Stuart, D.L. (2000). Tick histamine-binding proteins: lipocalins with a second binding cavity. Biochim Biophys Acta *1482*, 92-101.
Paesen, G.C., Siebold, C., Dallas, M.L., Peers, C., Harlos, K., Nuttall, P.A., Nunn, M.A., Stuart, D.I., and Esnouf, R.M. (2009). An ion-channel modulator from the saliva of the brown ear tick has a highly modified Kunitz/BPTI structure. Journal of molecular biology *389*, 734-747.
Pichu, S., Ribeiro, J.M., and Mather, T.N. (2009). Purification and characterization of a novel salivary antimicrobial peptide from the tick, *Ixodes scapularis*. Biochem Biophys Res Commun *390*, 511-515.
Pichu, S., Yalcin, E.B., Ribeiro, J.M., King, R.S., and Mather, T.N. (2011). Molecular characterization of novel sulfotransferases from the tick, *Ixodes scapularis*. BMC Biochem *12*, 32.
Portman, D.S. (2007). Genetic control of sex differences in *C. elegans* neurobiology and behavior. Advances in genetics *59*, 1-37.

Rego, R.O., Hajdusek, O., Kovar, V., Kopacek, P., Grubhoffer, L., and Hypsa, V. (2005). Molecular cloning and comparative analysis of fibrinogen-related proteins from the soft tick *Ornithodoros moubata* and the hard tick *Ixodes ricinus*. Insect Biochem Mol Biol *35*, 991-1004.
Ribeiro, J.M., Alarcon-Chaidez, F., Francischetti, I.M., Mans, B.J., Mather, T.N., Valenzuela, J.G., and Wikel, S.K. (2006). An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. Insect Biochem Mol Biol *36*, 111-129.
Ribeiro, J.M., Anderson, J.M., Manoukis, N.C., Meng, Z., and Francischetti, I.M. (2011). A further insight into the sialome of the tropical bont tick, Amblyomma variegatum. BMC genomics *12*, 136.
Ribeiro, J.M., and Francischetti, I.M. (2003). Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annual review of entomology *48*, 73-88.
Ribeiro, J.M., Makoul, G.T., and Robinson, D.R. (1988). *Ixodes dammini*: evidence for salivary prostacyclin secretion. The Journal of parasitology *74*, 1068-1069.
Ribeiro, J.M., and Mather, T.N. (1998). *Ixodes scapularis*: salivary kininase activity is a metallo dipeptidyl carboxypeptidase. Experimental parasitology *89*, 213-221.
Ribeiro, J.M., Schneider, M., and Guimaraes, J.A. (1995). Purification and characterization of prolixin S (nitrophorin 2), the salivary anticoagulant of the blood-sucking bug Rhodnius prolixus. BiochemJ *308 (Pt 1)*, 243-249.
Ribeiro, J.M.C., Evans, P.M., MacSwain, J.L., and Sauer, J. (1992). *Amblyomma americanum*: Characterization of salivary prostaglandins E2 and F2alpha by RP-HPLC/bioassay and gas chromatography-mass spectrometry. Exp Parasitol *74*, 112-116.
Richardson, J.L., Fuentes-Prior, P., Sadler, J.E., Huber, R., and Bode, W. (2002). Characterization of the residues involved in the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. Biochemistry *41*, 2535-2542.
Richardson, J.L., Kroger, B., Hoeffken, W., Sadler, J.E., Pereira, P., Huber, R., Bode, W., and Fuentes-Prior, P. (2000). Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. EMBO J *19*, 5650-5660.
Rios-Steiner, J.L., Murakami, M.T., Tulinsky, A., and Arni, R.K. (2007). Active and exo-site inhibition of human factor Xa: structure of des-Gla factor Xa inhibited by NAP5, a potent nematode anticoagulant protein from Ancylostoma caninum. JMolBiol *371*, 774-786.
Rodriguez-Valle, M., Vance, M., Moolhuijzen, P.M., Tao, X., and Lew-Tabor, A.E. (2012). Differential recognition by tick-resistant cattle of the recombinantly expressed *Rhipicephalus microplus* serine protease inhibitor-3 (RMS-3). Ticks and tick-borne diseases *3*, 159-169.
Rydel, T.J., Ravichandran, K.G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., and Fenton, J.W. (1990). The structure of a complex of recombinant hirudin and human alpha-thrombin. Science *249*, 277-280.
Rydel, T.J., Tulinsky, A., Bode, W., and Huber, R. (1991). Refined structure of the hirudin-thrombin complex. JMolBiol *221*, 583-601.
Salat, J., Paesen, G.C., Rezacova, P., Kotsyfakis, M., Kovarova, Z., Sanda, M., Majtan, J., Grunclova, L., Horka, H., Andersen, J.F*., et al.* (2010). Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*. The Biochemical journal *429*, 103-112.
Sauer, J.R., McSwain, J.L., Bowman, A.S., and Essenberg, R.C. (1995). Tick salivary gland physiology. Annual review of entomology *40*, 245-267.
Scacheri, E., Nitti, G., Valsasina, B., Orsini, G., Visco, C., Ferrera, M., Sawyer, R.T., and Sarmientos, P. (1993). Novel hirudin variants from the leech Hirudinaria manillensis. Amino acid sequence, cDNA cloning and genomic organization. EurJBiochem *214*, 295-304.
Schluter, H., Tepel, M., and Zidek, W. (1996). Vascular actions of diadenosine phosphates. Journal of autonomic pharmacology *16*, 357-362.

Schwarz, A., Valdes, J.J., and Kotsyfakis, M. (2012). The role of cystatins in tick physiology and blood feeding. Ticks and tick-borne diseases *3*, 117-127.
Seregni, E., Botti, C., Ballabio, G., and Bombardieri, E. (1996). Biochemical characteristics and recent biological knowledge on prostate-specific antigen. Tumori *82*, 72-77.
Shukla, J.N., and Nagaraju, J. (2010). Doublesex: a conserved downstream gene controlled by diverse upstream regulators. Journal of genetics *89*, 341-356.
Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J., and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. Genome research *19*, 1117-1123.
Siroky, P., Erhart, J., Petrzelkova, K.J., and Kamler, M. (2011). Life cycle of tortoise tick Hyalomma aegyptium under laboratory conditions. Experimental & applied acarology *54*, 277-284.
Slovak, M., Hajnicka, V., Labuda, M., and Fuchsberger, N. (2000). Comparison of the protein profiles of salivary gland extracts derived from three species of unfed and partially fed ixodid ticks analysed by SDS-PAGE. Folia parasitologica *47*, 67-71.
Slovak, M., Labuda, M., and Marley, S.E. (2002). Mass laboratory rearing of *Dermacentor reticulatus* ticks (Acarina, Ixodidae). Biologia (Bratisl) *57*, 261-266.
Soares, T.S., Watanabe, R.M., Tanaka-Azevedo, A.M., Torquato, R.J., Lu, S., Figueiredo, A.C., Pereira, P.J., and Tanaka, A.S. (2012). Expression and functional characterization of boophilin, a thrombin inhibitor from *Rhipicephalus (Boophilus) microplus* midgut. Veterinary parasitology *187*, 521-528.
Sonenshine, D.E. (1985). Pheromones and other semiochemicals of the acari. Annual review of entomology *30*, 1-28.
Sonenshine, D.E. (1991). Biology of ticks (New York: : Oxford University Press).
Sonnhammer, E.L., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol *6*, 175-182.
Spector, A.A., Fang, X., Snyder, G.D., and Weintraub, N.L. (2004). Epoxyeicosatrienoic acids (EETs): metabolism and biochemical function. Prog Lipid Res *43*, 55-90.
Stark, K.R., and James, A.A. (1995). A factor Xa-directed anticoagulant from the salivary glands of the yellow fever mosquito Aedes aegypti. ExpParasitol *81*, 321-331.
Stassens, P., Bergum, P.W., Gansemans, Y., Jespers, L., Laroche, Y., Huang, S., Maki, S., Messens, J., Lauwereys, M., Cappello, M*., et al.* (1996). Anticoagulant repertoire of the hookworm Ancylostoma caninum. ProcNatlAcadSciUSA *93*, 2149-2154.
Steiner, V., Knecht, R., Bornsen, K.O., Gassmann, E., Stone, S.R., Raschdorf, F., Schlaeppi, J.M., and Maschler, R. (1992). Primary structure and function of novel O-glycosylated hirudins from the leech Hirudinaria manillensis. Biochemistry *31*, 2294-2298.
Stone, S.R., and Hofsteenge, J. (1986). Kinetics of the inhibition of thrombin by hirudin. Biochemistry *25*, 4622-4628.
Strube, K.H., Kroger, B., Bialojan, S., Otte, M., and Dodt, J. (1993). Isolation, sequence analysis, and cloning of haemadin. An anticoagulant peptide from the Indian leech. JBiolChem *268*, 8590-8595.
Stutzer, C., Mans, B.J., Gaspar, A.R., Neitz, A.W., and Maritz-Olivier, C. (2009). *Ornithodoros savignyi:* soft tick apyrase belongs to the 5'-nucleotidase family. Experimental parasitology *122*, 318-327.
Szabo, C., Ischiropoulos, H., and Radi, R. (2007). Peroxynitrite: biochemistry, pathophysiology and development of therapeutics. Nature reviews Drug discovery *6*, 662-680.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution *28*, 2731-2739.

Tian, E., and Ten Hagen, K.G. (2009). Recent insights into the biological roles of mucin-type O-glycosylation. Glycoconj J *26*, 325-334.
Tuszynski, G.P., Gasic, T.B., and Gasic, G.J. (1987). Isolation and characterization of antistasin. An inhibitor of metastasis and coagulation. JBiolChem *262*, 9718-9723.
Utermohlen, O., Herz, J., Schramm, M., and Kronke, M. (2008). Fusogenicity of membranes: the impact of acid sphingomyelinase on innate immune responses. Immunobiology *213*, 307-314.
Valenzuela, J.G. (2004). Exploring tick saliva: from biochemistry to 'sialomes' and functional genomics. Parasitology *129*, S83-S94.
van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kroger, B., Hoffken, W., and Bode, W. (1995). Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. EMBO J *14*, 5149-5157.
van de Locht, A., Stubbs, M.T., Bode, W., Friedrich, T., Bollschweiler, C., Hoffken, W., and Huber, R. (1996a). The ornithodorin-thrombin crystal structure, a key to the TAP enigma? EMBO J *15*, 6011-6017.
van de Locht, A., Stubbs, M.T., Bode, W., Friedrich, T., Bollschweiler, C., Hoffken, W., and Huber, R. (1996b). The ornithodorin-thrombin crystal structure, a key to the TAP enigma? The EMBO journal *15*, 6011-6017.
Versteeg, H.H., Heemskerk, J.W., Levi, M., and Reitsma, P.H. (2013). New fundamentals in hemostasis. Physiological reviews *93*, 327-358.
Veveris-Lowe, T.L., Kruger, S.J., Walsh, T., Gardiner, R.A., and Clements, J.A. (2007). Seminal fluid characterization for male fertility and prostate cancer: kallikrein-related serine proteases and whole proteome approaches. Seminars in thrombosis and hemostasis *33*, 87-99.
Vitali, J., Martin, P.D., Malkowski, M.G., Robertson, W.D., Lazar, J.B., Winant, R.C., Johnson, P.H., and Edwards, B.F. (1992). The structure of a complex of bovine alpha-thrombin and recombinant hirudin at 2.8-A resolution. JBiolChem *267*, 17670-17678.
Walker, J.B. (1955). Rhipicephalus pulchellus Gerstacker 1873: a description of the larva and nymph with notes on the adults and on its biology. Parasitology *45*, 95-98.
Walker, M.J., Rylett, C.M., Keen, J.N., Audsley, N., Sajid, M., Shirras, A.D., and Isaac, R.E. (2006). Proteomic identification of *Drosophila melanogaster* male accessory gland proteins, including a pro-cathepsin and a soluble gamma-glutamyl transpeptidase. Proteome science *4*, 9.
Wang, H., and Nuttall, P.A. (1995a). Immunoglobulin-G binding proteins in the ixodid ticks, *Rhipicephalus appendiculatus, Amblyomma variegatum* and *Ixodes hexagonus*. Parasitology *111*, 161-165.
Wang, H., and Nuttall, P.A. (1995b). Immunoglobulin G binding proteins in male *Rhipicephalus appendiculatus* ticks. Parasite Immunol *17*, 517-524.
Wang, H., and Nuttall, P.A. (1999). Immunoglobulin-binding proteins in ticks: new target for vaccine development against a blood-feeding parasite. Cell Mol Life Sci *56*, 286-295.
Wang, H., Paesen, G.C., Nuttall, P.A., and Barbour, A.G. (1998). Male ticks help their mates to feed. Nature *391*, 753-754.
Warkentin, T.E., Greinacher, A., and Koster, A. (2008). Bivalirudin. ThrombHaemost *99*, 830-839.
Wartha, F., Beiter, K., Normark, S., and Henriques-Normark, B. (2007). Neutrophil extracellular traps: casting the NET over pathogenesis. Curr Opin Microbiol *10*, 52-56.
Waxman, L., Smith, D.E., Arcuri, K.E., and Vlasuk, G.P. (1990). Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. Science *248*, 593-596.
Wei, A., Alexander, R.S., Duke, J., Ross, H., Rosenfeld, S.A., and Chang, C.H. (1998). Unexpected binding mode of tick anticoagulant peptide complexed to bovine factor Xa. JMolBiol *283*, 147-154.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O*., et al.* (2007). A unified classification system for eukaryotic transposable elements. Nature reviews Genetics *8*, 973-982.
Wolfner, M.F., Harada, H.A., Bertram, M.J., Stelick, T.J., Kraus, K.W., Kalb, J.M., Lung, Y.O., Neubaum, D.M., Park, M., and Tram, U. (1997). New genes for male accessory gland proteins in *Drosophila melanogaster*. Insect Biochem Mol Biol *27*, 825-834.
Wolfsberg, T.G., Primakoff, P., Myles, D.G., and White, J.M. (1995). ADAM, a novel family of membrane proteins containing A Disintegrin And Metalloprotease domain: multipotential functions in cell-cell and cell-matrix interactions. The Journal of cell biology *131*, 275-278.
Xu, X., Yang, H., Ma, D., Wu, J., Wang, Y., Song, Y., Wang, X., Lu, Y., Yang, J., and Lai, R. (2008). Toward an understanding of the molecular mechanism for successful blood feeding by coupling proteomics analysis with pharmacological testing of horsefly salivary glands. Mol Cell Proteomics *7*, 582-590.
Yadav, V.K., Chhikara, N., Gill, K., Dey, S., Singh, S., and Yadav, S. (2013). Three low molecular weight cysteine proteinase inhibitors of human seminal fluid: purification and enzyme kinetic properties. Biochimie *95*, 1552-1559.
Yamaji, K., Tsuji, N., Miyoshi, T., Islam, M.K., Hatta, T., Alim, M.A., Anisuzzaman, M., Kushibiki, S., and Fujisaki, K. (2009). A salivary cystatin, HlSC-1, from the ixodid tick *Haemaphysalis longicornis* play roles in the blood-feeding processes. Parasitology research *106*, 61-68.
Zhao, Q.Y., Wang, Y., Kong, Y.M., Luo, D., Li, X., and Hao, P. (2011). Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC bioinformatics *12 Suppl 14*, S2.
Zheng, H., Yu, Z., Chen, Z., Zhou, L., Zheng, B., Ma, H., and Liu, J. (2011). Development and biological characteristics of Haemaphysalis longicornis (Acari: Ixodidae) under field conditions. Experimental & applied acarology *53*, 377-388.
Zhou, J., Gong, H., Zhou, Y., Xuan, X., and Fujisaki, K. (2006). Identification of a glycine-rich protein from the tick *Rhipicephalus haemaphysaloides* and evaluation of its vaccine potential against tick feeding. Parasitology research *100*, 77-84.

Appendices

Appendix 1. Sequence of genes chosen from the bilaris subclasses

>RpSigp-759502
ATGTTTCGCATGCCTTTTCCGTTCCTTGCGTGCCTATTGGGAATAATGACGGTACTT
TGGCTCAGCGTCACCGGTATCAAACAACAAGCGGCACCTAAGAGGAAAATGTGCAAC
CCTTATGAACCCGCACAAGTACACTGTCGTCGTCCTGAAGAGAGACGCTATTTTTAC
GACAAACCCGCCAATAAATGCAAGGAAATTAATGTTGTTCCTTGTACTAAACAAACC
AGCCTTTATACAAGGCTCACAGATTGCTTGAAGGAATGCGCACCAAAGTCCCCGTGT
AGAACTCCTGTATGGAAGAAACCTGCAAATACAGTAAACGGGGAATACTATACTTTC
GATCCTCAAGAACGTGAATGCTATTTGCAATACGGAGAATTTTACGGCACAGAAGAA
AACATGTTTGAGCGCCGCTACGATTGCAAAAAAACTTGCTCTCCAACTTATGAATAA

>RpSigp-907535
ATGATGGCGTCTCGTCTTTTCATTTGGCTCGTTTGCTTCATTGAACTGGGAGTCTAT
GGTCCTGCCGAGACGACTACTGTGAACGGACGAGCCGCATCACGAAGAAGACGAATT
CCAACCTTGTGCAGTCATCCTTCTGTACCGGCACCACACTGCAGTCCACCTCACGAA
TTTGTCTACCACTACAGTTTTAATACAACAACTCAAGTATGCGAGAAATACAGGCAA
TGCAGTTGGGTTTCTGGATTTCTTTCTCGGAAGGAATGCTACTCTAAATGCGGACAT
AATTCGCCATGCTTGAAGCAGTCAGAGATGGACTTGAGTTTTCTAGAGCTAACATGG
TTTCGTTATAATCATCATAATGACAAATGTGAAGCTCGTACTGATAACAAAATGCCA
TGGGATGAGTGGCCAGCGGAGAACGGTTTCTTTAACGAAGAGGACTGCATTAAGGCA
TGTCAACCGAATCGAACAATAACTTACATCTGA

>RpSigp-673073
ATGTTTCGCATGCCTTTTCCTTTCCTTGTGTGCCTAATTGGAAAAATGGCCGTAGTT
TCCCTCGGCGTCGCCGATAGCACAAATCAAGTGTCTGAATTTTGCCAAGATCCTTAT
TCTGGTGGCTCATACAAGGAACGATGTGCTGGCCCGTATGAACGAGAGAAACGATAT
ATAGGATATAAAAATAAATGCAAACAAATTTTTTGGGATCCATGCCAAGAAAACTTT
AGAACTTACAGAACACTCGCAGAATGTCTAGGATATTGTGCAAAGGAAAGCGTGTGC
GTTAAAACACCCCTGGAGCATGGCCTAAATACTCAAGAGAATAAATATCACTACTAT
TTTGACATTCAAGAAGGAACATGTTATTCAAGATTAGGAACTTTTCAACGCTCACAA
AAAGGGGAGAATCATTTTCGAAACAAGACCTTATGTGAGGAGGCTTGCTCACCGGTG
AAAATTTGA

>RpSigp-921556
ATGCATCACTATTTCTTCATTTTGTTCCTCTGCGTGATGGGTACAAATGAAGCATAC
CAGCGTGAAATAATATCAGTGGGCAAATCCCAACGCCATGCTACTGACCAATGGAGG
TGTGAGCGTCCTATACCTCCTCAGTATCAATGCCATGGCAAAGGGAGTTCCCGAACA
CGATATAGTTACCATAATGAAAGAGGACGTTGCGTCTTGGTGGAGATTCCGTCATGT
TTTTCAGCAGGAAACGGAAACATATTTCCTTCAAGAAGAGCTTGCCTGAAATTGTGC
AATCCTGAATCCGTGTGCCTGGAGAGTAGAAATCGCACCATCAAAGATTATCATTAC
GCGTATGACAAAAAGAAAGACAGTTGCAACTTCATTATACCACCGCGACATAGGAAA
AGACTTCATAGGGCGACAGGAAATGCATTTCCTTATGAAGGTGACTGCGAATATGAA
TGTAAACCGAAGAATATTTCGGACCCAAGGGTCTTATAG

>RpIx75-485129
ATGGATGCAACATTTGTAACAGTGAGTCTTCTGGCCATATATGTCACTTTCTGGAAG
CACAGCAGCGCCACTGCTTCAAGGAACCCTGCCTGCAGTGTAGCCCCTGAGGTTGAA
AACTGCTCTATAGTTCTCTTTCGATGGAGCTACGATTCGGAATTGAACAAGTGTAAG
CAGAATTTCGTTTGCCGTGAAAATGCGAATAACTTTGAGACAAAAGACCTTTGCGAA
ACTACATGCCCACCAGTTTCAGGAGTCACGCCTGCGCCGAAACCGAAGGATTGCCAC
TATTGGCTCGCTAATGGACAAGGATGTTACTCATATTGGTTTACATCTAATTATGAT
TACTTGGGAAGACAACACCCTATAATGGTATACACTGGATGTGGACTGTGGAAAAGA
AATCTTTATGCTTACGACATGAGTACGGGAAAATGCTTAGAAATAAAAAGAAGAGGT
TATGGTTCATGGCAATGA

>RpIx75-909901
ATGTCTAATTCATACATATATAAAATTTGTCTGGTCCTTCCAGTCCTACTTCAGCAG
TACTCGGCAAACGCCCAATCAAGAGCGAGTGCTTGTCTACAGCCGCCTACAGTGGAA
GGGTGCAGTCTCATTCGCCGGATGTGGAGTTTCAACGCTGCCTCAGAACAATGCGAA
CAAAACTTTGTCTGCTCCAACCACACGAATGCTTTTCAAGATAAAGCGTCATGCATG
GCGATGTGTCCTTCAGGTCCAAGTCCCCAACCTCCAGAGAGCCATAGAGGTTGCGAC
TACTGGATTCTTCGATTGGATCAGTGCAAACGGACGTGGTGGACATTATACCGGGAT
GGTTGGGGAGTACCAAAGAGGGCATTCATATACACTGGATGCGGGCCTTCACCACAT
AGGTACTACGCCTACTACGTGGACCAACGCCGATGTGAAGAACTTCGCTTTCGTGGC
GGACGAGCAAACTAA

>RpIx75-477142
ATGTGGACATTCAATGTGTTCATATTTATACTGGCTAATAATGCACATCTGTTCAGG
TATGCTGGTGCTCTTGAAAGAAGGGATGATTGTGACGTACCACCAACCGTGGAAGGG
TGCAGCATTATACGACGAAAGTGGAGCTTTCTGCCCGAAATGGGGAAGTGCGCAATG
AACTTCGTATGCTCGAATCACCCCAATGCATTCCAAACAGAACAGGAGTGCGAAGCT
TCTTGCCCGCCAGATACAAGCAACAAACGTACCCCAAGAGATGACTGCTATTATTGG
CTTCAAAATTTGGATGAGTGCCGATTCAAACGTGAAACGTTCTACCCAGACCCGTAC
GGCCGAAGACAGCGAGTTTTGCTGTTCAGATTCTGTGGCGAATCAAGCTCGAAGTTA
TATGCGTACTACATGTACAGTGGGGACTGCAGCGAAATTGTGTTGCGAAGCTGA

>RpIx75-699640
ATGTGGACGCTCAATGTGTTCATATTTATACTGGCTATTAATGCATATCTGTTCAGG
TATGCTGGTGCTCTTGAAAGAAGGGATGATTGTGACGTACCACCAACCGTGGAAGGG
TGCAGCATTATACGGCGAAAGTGGAGCTTTCTGCCCGAAATGGGGAAGTGCGCAATG
AACTTCGTATGCTCGAATCACCCCAATGCATTCCTAACTGAACAGGAGTGCGAAGCT
GCTTGCCCACCAGATACAGGCCACAAACCTACCCCAAGAGATGACTGCTATTATTGG
CTTCAAAATTTGGATGAGTGCCAATTCAAACGTGAAACGTTCTACCCAGACCCGTAC
GGCCGAAGACAGCGAGTTTTGCTGTTCAGATTCTGTGGCGAATCAAGCTCGAAGTTA
TATGCGTACTACATGTACAGTGGCGATTGCAGCGAAATTGTCTTGCGAAGCTGA

>RpSigp-842790
ATGCTAATTAGACAATTTATGGCTGTGTTTGAGCTCTTTAGTGGCCTACGGTACGTG
TCAGGTGACCCTTTGGATATGGACTGGGTCATCAAAAAGCCAGTGGCCTGCTACATG
AAACCAGATTACGGTACCTGCAAAGGACATTTCACGCGTTATTTCTACAACGACTCA
AACTACAAATGCAGAAGTTTCGACTATAGCGGCTGTGGAGGAAACGGCAACAATTTC
GACTCACAACGAGAATGCAGATTCTTATGTGGTGTGAAATACGACCCAGACAAGGAT
CCGTGTCTGCGTCCTCCTGGCAAGAGATGGTGTCCCACCTCGCCAAGATATCCAGGA

ATGTGGTATTTTGACAAGAAAACCGAAACATGTGTCCCGTTCATTTATTACCAATGT
GCCTTGGATCGGAATGTGTTTCCGTCTTGCGACAAGTGCAAGAAAGAATGTCAGCGT
CACATGCATGTCTTACAAACATGCCCTGAATAG

>RpSigp-843116
ATGCTAATGAAACAATTCATGGCTGTCTTTCAGCTCTTTTGTGGCCTGTGGTACGTG
TCAGGTGACCCTTTGGACATGGACTGGGTCATCAAAAAGCCAGTGGCCTGCTACATG
AAGCCAGATTACGGCACCTGCAAAGGACATTTCACGCGTTATTTCTACAATGACTCA
AACTACAAGTGCAGAAGTTTCGACTATAGCGGATGTGGAGGAAACGGAAACAATTTC
GACTCAAAACGAGAGTGCAGATATTTATGTGGAGTGAAATACGACCCAGACAAGGAT
CCGTGTCTGCGTCCTCCTGGCAAAAGATGGTGTCCCACCTCACCAAGATATCCAGGA
ATGTGGTATTTTGACAAGAAAACCGAAAAATGTGTCCCGTTCATTTATTACCAATGT
GCCTTGGATCGGAATGTGTTCCCGTCTTGCGACAAGTGCAAGAAAGAATGTCAGCGA
CACATGCAAGTGTTACAAACGTGCCCTGAATAG

>RpIx75-680109
GAGACCATGAATCTCGCAGGCCTGCTTCTTTGCTTGTTTTGCCTGGCATTTACGCAT
GTTTCTATATTAGCTGGAAGAGAAATACTATCAAGCTGCAACCTGCCACCAAAAACC
GACCACTGCCGAGCACGACATCTAAGATGGTATTTCGATTCCATACGCGGCCGCTGC
AGAATGTTCACTTATGGTGGATGTGGAGGAAACAACAACCGCTTTTCAACAGAAAGA
GAATGTATGGCCGAATGCGCGCCCACTTCTCCATATCCGGACGTATGCAGCCGGAAG
CCAAGCTATCAGTCCTGCTATTACGCTTTGCACTACTATTATTCGGGTACTACAGTG
TGGTACTTTCATCCCGTACATGGCAGCTGTGAGCAGTTTCGACCTGGCCGGTGTCCC
AAGGGTTGGACCTTGTTCTGGAGCTGCCAGGCATGCTCAAACATATGCACTAATTAT
GTAGCCTGCAACAGCAAAGCGGAGTAA

>RpIx75-657025
TTATATATTTTTCTTCTCAAAGCTGGTGCATCACGCCTTCGCTGCTGGCTACCTAAA
GTTGTTGGACGATGTAACAAGTCGGTACCATCGTGGTTCTACAACATGTGGACGGCC
CAATGCATAGGTTTCATCTACAGTGGCTGTGGAGGAAATTCGAACAGATTCGAAACG
GAGGAAGAGTGCAACAATGCCTGCTTGCCCCCAAATAAGGCAAAGAAGAACGTCTGC
AGTTTGGAAGCGGATTCGGGCTCCTGCAAAGGATATAATCCAAAGTGGACGTACGAC
CACAAGAAGGACATATGCCGGGGATTCGTGTACGGTGGTTGTGGAGGCAACGCCAAC
AGATTCAGTACATGCTTGGAGTGCATGAAAAGGTGCAGCGGAAGAGAGGGTCACTTG
CGTCATTGCATTTTACTGACACCGAAATTCAACGACAAGTTTTACATGGCATGGGAG
CCCGAATAA

Appendix 2. List of primers for cloning

Gene	Forward Primer Sequence (5'-3')	Length (bp)	Reverse Primer Sequence (5'-3')	Length (bp)	Product Size (bp)
RpSigp-759502	AAAATGGTTCGCATGCCTTTTCCGTTCCTTGCG	33	GGGCGATCATTATTCATAAGTTGGAGAGC	29	465
RpSigp-673073	AAA ATG GTT CGC ATG CCT TTT CCT TTC CTT GTG TGC C	37	GGG TCA AAT TTT CAC CGG TGA GCA AGC	27	471
RpIx75-485129	AAA ATG GAT GCA ACA TTT GTA ACA GTG AGT CTT CTG GCC	39	GGG CTA TTC AGG GCA TGT TTG TAA GAC ATG	30	495
RpIx75-699640	AAAATGGGGACGCTCAATGTGTTCATATTTATACTGGC	38	GGGCGATCAGCTTCGCAAGACAATTTCGCTGCAATCG	37	462
RpSigp-843116	AAAATGGTAATGAAACAATTCATGGCTGTCTTTCAGCTC	39	GGGCGATCACTATTCAGGGCACGTTTGTAACACTTGCATG	40	498
RpIx75-680109	GGG ATG GAG ACC ATG AAT CTC GCA GGC CT	29	GGG TTA CTC CGC TTT GCT GTT GCA GGC T	28	492