SALIVARY GLAND TRANSCRIPTOME OF *RHIPICEPHALUS (BOOPHILUS) MICROPLUS*

by

Siyamcela Genu

Submitted in accordance with the requirements for the degree

Master Degree in Life Science

in the

COLLEGE OF AGRICULTURE AND ENVIRONMENTAL SCIENCES

DEPARTMENT OF LIFE AND CONSUMER SCIENCE

at the

UNIVERSITY OF SOUTH AFRICA

FLORIDA CAMPUS

SUPERVISOR

PROF. B.J. MANS

November 2017

DEDICATION

To my dad and my children

This is my special dedication to my dad, **Sizwe Genu** (***Tat' uRhadebe***). You have been my pillar of strength throughout the years, since I was a little boy till this day. I never really expressed the undying love I have for you, and with this dissertation I hope to convey the message of the strong love I have for you. Your unconditional love is what have made me to be a man, the father, a husband and above all the humble and God fearing son. *Ndiyakuthanda tata wam kwaye ndicela uThixo akugcine khona ukuze uyibone impumelelo yam ngenxa yothando nemesebenzi yakho emihle*.

To my daughters, I know I spend most my time away from you because of work, but the love I have for you is unmeasurable. You always on my thoughts and prayers. I only see good in you and I always ask God to keep you under His wing and guidance. I want to be the best man in your lives and be the good parent whom you can ever wish for. To me you are a gift from God and I thank Him every day for giving me such wonderful people, whom I call my daughters. *Sinakekele nawe Azasakhe ndiyanithanda, kwaye ndifuna nihlale nikwazi oko imihla yonke. Nanga ningakhula niphumelele ebomini, kodwa kuko konke nibeke ithemba lenu ku Yehova.*

Proverbs 20:7

"The righteous man walks in his integrity; His children are blessed after him"

DECLARATION

'I declare that ***Salivary Gland Transcriptome of Rhipicephalus (Boophilus) microplus*** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

I further declare that I submitted the thesis to originality checking software. The result summary is attached.

I further declare that I have not previously submitted this work, or part of it, for examination at Unisa for another qualification or at any other higher education institution.'

Signed: _____ **Date:** _____

ACKNOWLEDGMENTS

- Almighty **God** for giving me the strength and courage to strive and persevere to achieve this project.

- My family: **Sizwe**, **Sinakekele**, **Azasakhe**, **Xhantilomzi**, **Thumeka**, **Thabisa**, and **Zusiphe**.

- My fiancé **Busisiwe Maseti**, I love you MamCirha. You have a special place in heart and I want to thank you for being a supportive, loving, caring and understanding partner. I thank the Lord for bringing you in my life.

- My supervisor Prof. **Ben J. Mans** for professional support and for giving me the opportunity to work with him and for ARC-OVI.

- My colleagues: **Philasande Gaven** and **Matsatsane Mahlaku** for support and motivation.

- Agricultural Research Council (**ARC**) and National Research Foundation (**NRF**) for funding.

Table of Contents

Title page

Dedication……..……..……**I**

Declaration……..……..……..**II**

Acknowledgements……..……**III**

Table of contents……....**IV-VIII**

List of tables……..……**IX**

List of figures……..……..……..**X-XI**

List of acrynomies and abbreviations……..…..**XII-XIII**

Summary……..……....**XIV**

Chapter 1: Introduction and literature review……..……**1**

1.1. Ticks as a global problem……..……..……**1**

1.2. Tick taxonomy……..……..**1-2**

1.3. *Rhipicephalus (Boophilus) microplus*……..……**3-4**

1.4. Tick control and vaccines need……..……..……..**4**

1.4.1. Importance of tick control……..……..……**5-7**

1.5. Vaccine types: Exposed VS hidden vaccines……..……**7-8**

1.5.1. Exposed vaccines……..……..…**8-10**

1.5.2. Hidden (concealed) vaccines……..……..…**11-13**

1.6. Tick salivary glands molecules: good targets for vaccine development...**14-15**

1.7. Tick feeding and host interactions……..……**15-17**

1.8. Tick feeding and survival: anti-hemostatics and immunomodulators……..…**17**

1.9. Transcriptome studies……..……**19-22**

1.10. Proteomics studies……..……..**24-25**

1.11. Overall aim and objectives……..**26**

1.11.1. Overall aim……..……**26**

1.11.2. Objectives……..……**26**

Chapter 2: Experimental approach and methodology……..……..…..**27**

2.1. Introduction…….……**27**

2.1.1. Transcriptomics and next-generation sequencing…….…..**27**

(a) The traditional sequencing methods/Sanger sequencing method.....**28-29**

(b) Next-generation sequencing (NGS)……**29-33**

(i) 454 sequencing……..…….……..**34-35**

(ii) ABI/SOLiD sequencing……..……....**36-38**

(iii) Illumina sequencing……..……..**38-40**

2.1.2. Proteomics approach on tick antigens identification……..……**41-42**

2.2. Transcriptomics…….……**43**

2.2.1. Tick feeding, collection and ethics clearance…….……..**43**

2.2.2. Tick salivary glands isolation/preparation of salivary gland extracts…..…..**43**

2.2.3. RNA isolation from tick salivary glands……..……..**43-44**

2.2.4. Determination of RNA quantification by NanoDrop……..……..**44**

2.2.5. NGS: salivary gland cDNAs for transcriptome analysis……..……..……..**44**

(a) Messenger RNA purification and fragmentation……..…..**44-46**

(b) First strand cDNA synthesis…….……..**46**

(c) Second strand cDNA synthesis……..…...…..**46-47**

(d) Perform end repair……...……...……**47**

(e) 3’ ends adenylation…….……**48**

(f) Adapter ligation……..…..**48-49**

(g) DNA fragments enrichment……..……..…..**49-50**

2.2.6. Complementary DNA (cDNA) library validation……...……**50**

(a) Qubit® HS quantification assay……...……**50**

(b) Agarose gel electrophoresis analysis……..…….**51**

2.2.7. Gel clean-up……..…..**51-52**

2.2.8. Qubit® HS quantification assay…….……**52**

2.2.9. Qubit® HS quantification assay…….…**52**

2.2.10. Agarose gel electrophoresis analysis……..……..……**52-53**

2.2.11. Transcriptome sequencing……..…….**53**

(a) Sequencing with Illumina MiSeq sequencing…….……...**53**

2.2.12. *De novo* assembly of transcriptome data…….…….**53**

(a) Data filtering…….……...…….**53-54**

(b) Transcriptome *de novo* assembly and annotation…….……...**54**

(c) Creating a non-redundant open-reading frame (ORF) set and selection based on RPKM cut off……..…...……..…**54-55**

(d) Identification of correct open reading frames (ORFs)……..……**55**

(e) Identification of putative secretory and house-keeping proteins……..**55**

(f) Quality assessment of the transcriptome using BUSCO…….…..**55**

2.3. Proteomics…….…..**56**

2.3.1. Salivary gland extracts (SGEs) preparation…….……...**56**

2.3.2. Sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis…..**56**

2.3.3. Gel electrophoresis……..…..**56-57**

2.3.4. Proteomics studies…….……..**57**

2.3.5. Proteome data analysis…….…..…**57**

(a) Database searching……..……..**57**

(b) Criteria for protein identification……..**57-58**

Chapter 3: Transcriptome sequencing and data analysis, and discussion…….59

3.1. Introduction……..……**59-61**

3.2. Results……..…….…….**62**

3.2.1. RNA isolation and purity……..**62-63**

3.2.2. cDNA library validation……..……**63-64**

3.2.3. Illumina sequencing and quality trimming……..……**64-65**

3.2.4. *De novo* assembly……**66**

3.2.5. BUSCO analysis to determine the quality of the transcriptome…..…**66-67**

3.2.6. Functional annotation……..……**67**

3.2.6.1. Housekeeping proteins……..……..……**68-69**

3.2.6.2. Putative secreted proteins……...……**70-72**

3.3. Discussion……..……..……**73**

3.3.1. Housekeeping proteins…….……...**74**

3.3.2. Putative secreted proteins……..……..…**75**

(a) Lipocalins…….……..…**75-76**

(b) Metalloproteases……..……..…….**76-77**

(c) Basic pancreatic trypsin inhibitors/Kunitz……..**77-78**

(d) Cement proteins…….**79-80**

3.3.3. Improved *R. (B.) microplus* salivary gland transcriptome coverage…....**80**

3.4. Conclusion…….……..…**81**

Chapter 4: Proteome analysis: data analysis and discussion…….……**82**

4.1. Introduction……..……..…..**82-83**

4.2. Results……..…**84-87**

4.2.1. Deep *R. (B.) microplus* salivary gland transcriptome coverage…….…..**88**

4.2.1.1. Housekeeping proteins……..……..……**89**

4.2.1.2. Putative secreted proteins…….……..**91**

4.3. Discussion……..……..……..**94**

4.3.1. Housekeeping proteins……..…….**95**

4.3.2. Putative secreted proteins……**95-96**

(a) Cement proteins……..…...**96**

(b) Metalloproteases…….……..…...**97**

(c) Lipocalins……..……..**97-98**

(d) Basic protease trypsin inhibitor (BPTI/Kunitz)……..…….**98-99**

(e) Basic tail secreted proteins (BTSP)……...…….**99**

4.4. Conclusion……..……..**100**

Chapter 5: General conclusions and recommendations……..……**101**

5.1. High abundance of cement proteins in the proteome of the *R. (B.) microplus* salivary gland proteome……..……..**101**

(i) Tick cement proteins……...…….**101-102**

(ii) Cement proteins studies on ticks……..….**102-103**

5.2. Transcriptome and proteome correlation……...…..**103-104**

5.3. Correlation between this study and other tick studies……..…..**104**

5.4. Methodology and data analysis shortcomings……...…..……**105**

5.4.1. Transcriptomics……..……..**105**

5.4.2. Proteomics……..…...**105-106**

References…….……**107-138**

Websites and databases……...……..**139**

List of tables

Table **Page**

1.1. Exposed antigens evaluated as recombinant anti-tick vaccines…..……**10**

1.2. Concealed antigens evaluated as recombinant anti-tick vaccines……..….**13**

1.3. List of anti-hemostatics and immunomodulators of *R. (B.) microplus*.……..…**18**

1.4. Summary of some tick transcriptomes and proteomes achieved by NGS technologies…….……..**23**

2.1. Comparison of features of NGS platforms/technologies used for transcriptome applications……**32**

2.2. Parameters for adapter trimming…….....……**53**

3.1. The results for nanodrop 260/280 and 260/230 ratios for RNA purity quantification…….……**63**

3.2. The results for cDNA libraries quantification with Qubit High Sensitive kit…....**63**

3.3. Statistics for quality trimming and creation of various data structures for *R. (B.) microplus*……...…**65**

3.4. Reduction of ORF numbers to obtain a representative set for analysis……..**66**

3.5. Protein families identified from the annotation of the *R. (B.) microplus* salivary gland transcriptome…….……**68**

3.6. Pathway classification of housekeeping genes in *R. (B.) microplus* salivary transcriptome……..…….……**68**

3.7. List of most abundant *R. (B.) microplus* salivary gland transcriptome secretory protein families.……**70**

4.1. *R. (B.) microplus* salivary gland proteome protein classification……..……..**87**

4.2. Some of the tick proteomics studies identified major protein families……..…**92**

List of figures

Figure **Page**

1.1. One-host life cycle hard tick, eg: *Rhipicephalus (Boophilus) microplus*…..……..**3**

1.2. Interactions between tick-borne pathogens, tick vectors and vertebrate hosts..**16**

2.1. The workflow of classical Sanger sequencing method workflow……..…....**28**

2.2. The next generation sequencing high-throughput sequencing workflow……....**30**

2.3. The time-line of classical DNA sequencing and next generation sequencing technologies ……..……..…….**31**

2.4. Improvements in high-throughput and read length of next generation sequencing technologies over the years……..……**33**

2.5. An overview of Roche GS FLX sequencing method……..…..**35**

2.6. An overview of Applied Biosystems SOLiD sequencing by ligation method…...**37**

2.7. An overview of Illumina Genome Analyzer sequencing method……..…**40**

2.8. A typical proteomics approach workflow……..……**41**

3.1. The NanoDrop results for RNA integrity results for BM1-5……..….**62**

3.2. 1% agarose gel electrophoresis depicting cDNA library integrity results……....**64**

3.3. Putative housekeeping proteins identified in *R. (B.) microplus* (ARC-OVR) transcriptome database……..…**69**

3.4. Putative secreted proteins identified in *R. (B.) microplus* (ARC-OVR) transcriptome database……..…**71**

3.5. Species count from our *R. (B.) microplus* (ARC-OVR) transcriptome study with the highest and best hits on the NCBI database……..……..**72**

4.1. SDS-PAGE separated pattern for *R. (B.) microplus* salivary gland membrane protein fraction for all feeding stages (BM1-5)……..…..**85**

4.2. SDS-PAGE separated pattern for *R. (B.) microplus* salivary gland membrane pellet fraction for all feeding stages (BM1-5)……..……....**86**

4.3. SDS-PAGE separated pattern for *R. (B.) microplus* salivary gland membrane soluble fraction for all feeding stages (BM1-5)……..……..……..**87**

4.4. The *R. (B.) microplus* salivary gland deep proteome coverage……..……..…**88**

4.5. Proteomics classification and abundance of the major putative housekeeping proteins from the *R. (B.) microplus* salivary gland proteins based on the LC-MS/MS analysis…….....……..…..……**90**

4.6. Proteomic identification of putative secreted proteins on *R. (B.) microplus* (ARC-OVR) transcriptome database……...……...…….**93**

List of abbreviations and acronyms

1DE	**One Dimensional Gel Electrophoresis**
ACN	**Acetonitrile**
APS	**Ammonium persulphate**
ATL	**A-Tailing Mix**
ATP	**Adenosine tri-phosphate**
BLAST	**Basic Local Alignment Search Tool**
BLAST P	**Accelerated protein-protein BLAST**
bp	**Base pair**
BSA	**Bovine serum albumin**
BUSCO	**Benchmarking Universal Single-Copy Orthologs**
CDNA	**Complementary Deoxyribonucleic Acid**
C-terminus	**Carboxyl terminus**
Da	**Dalton**
DNA	**Deoxyribonucleic acid**
DNA	**Deoxyribonucleic Acid**
dNTP	**2`-deoxynucleoside 5`-triphosate**
DTT	**1, 4-dithio-DL-threitol**
EPF	**Elute, Prime, Fragment Mix**
EST	**Expressed Sequence Tag**
EtOH	**Ethanol**
FSM-STR	**First Strand Master Mix-STR**
g	**gravity or gram**
Hrs	**hours**
iPr	**Isopropanol**
Kbp	**kilo base pairs**
KDa	**kilo Dalton**

KEGG	**Kyoto Encyclopedia of Genes and Genomes**
LC	**Liquid Chromatography**
LIG	**Ligation Mix**
MALDI	**Matrix Assisted Laser Desorption/Ionization**
Mb	**Mega bases**
Min	**minute**
mM	**milli molar**
MRNA	**Messenger Ribonucleic Acid**
MS/MS	**Tandem Mass Spectrometry**
NGS	**Next-Generation Sequencing**
Nt	**Nucleotides**
PAGE	**polyacrylamide gel electrophoresis**
PCR	**polymerase chain reaction**
PDB	**Protein Data Bank**
PMM	**PCR Master Mix**
PPC	**PCR Primer Cocktail**
RBP	**RNA Beads Purification**
RLT-DTT	**RLT-Dithiothreitol**
RNA	**Ribonucleic Acid**
SDS	**sodium dodecyl sulphate**
Sec	**seconds**
sp.	**species**
STL	**Stop Ligation Buffer**
TAE	**Tris acetate EDTA**
TOF	**Time of Flight; MS: Mass Spectrometry**
TPA	**Third Party Annotation**

SUMMARY

The cattle tick, *Rhipicephalus* (*Boophilus*) *microplus* is a tick of veterinary and health importance globally, transmitting *Babesia bovis* and *B. bigemina*. Tick control is important and needed to prevent livestock diseases caused by tick-transmitted pathogens. Traditionally, tick control methods have resulted in development of acaricide-resistant ticks, environmental pollution and meat and milk contamination. Therefore, there is a need for alternative method and vaccines directed against tick feeding. The aim of this study was to identify proteins involved in tick feeding, tick-host-pathogen interactions and tick reproduction. Consequently, these will help in identification of antigens with the ultimate goal of developing anti-tick vaccines. *R.* (*B.*) *microplus* female ticks were collected at five different feeding stages. RNA was isolated from the salivary gland extracts (SGEs). The cDNA libraries were synthesized and sequenced with the Illumina MiSeq technology. Transcriptome data was analyzed with CLC Genomics Workbench, Trinity and Minia. The SGEs were also used to isolate the fractions: membrane, soluble and pellet protein for proteomic analysis. The proteomics data was analysed with Mascot, X!Tandem and Scaffold. Both the transcriptome and proteome analysis revealed the presence of major secretory protein families such as Kunitz, lipocalins, serpins, cement proteins and metalloproteases, while the majority of transcripts coded for housekeeping genes

Chapter 1: Introduction and Literature Review

1.1. Ticks as a global problem

Ticks are obligate haematophagous ectoparasites (Singh and Girschick, 2003; de la Fuente and Kocan, 2006). Ticks are the vectors with a wide-range of pathogens such as bacteria, fungi, protozoa etc. They are also responsible for transmitting diseases of veterinary importance, among others Anaplasmosis and Babesiosis, which cause economic losses to livestock production globally (Jongejan and Uilenberg, 2004).

This study is about the southern cattle tick, *Rhipicephalus (Boophilus) microplus* (*R.(B.) microplus*), a vector of *Babesia bovis*, *Babesia bigemia* and *Anaplasma marginale,* the causative agents respectively, of Babesiosis (red water) and Anaplasmosis (gall sickness) (Peter et al. 2005). These diseases cause substantial stock losses in tropical and subtropical areas, globally (Parizi et al. 2009). *R. (B.) microplus* is the most highly adaptive and invasive species (Madder et al. 2011). As a result, several studies revealed that this tick species has recently been found in places where it has never existed (Tønnesen et al. 2004; Lynen et al. 2008; Madder et al. 2011).

1.2. Tick taxonomy

There are 896 tick species that parasitize vertebrates. These tick species consist of three families, namely, Ixodidae (702 species), Argasidae (193 species) and Nuttalliellidae (1 species) (Guglielmone et al. 2010). The livestock ticks of the following genera, *Amblyomma; Boophilus; Hyalomma* and *Rhipicephalus* are considered to be of the most economical importance in the Ixodidae family (Frans, 2000). Five *Boophilus* species have recently been put under the genus of *Rhipicephalus* (Horak et al. 2002). However, the name *Boophilus* is used as subgenus (Horak et al. 2002; Barker and Murrel, 2002). Ixodidae ticks possess a sclerotized scutum, the gnathostoma and numerous hypostome denticles (Sonenshine, 1991). Hard ticks of all life stages feed, on their hosts, for lengthy periods (several days to

weeks) (Sonenshine, 1991). Their feeding period also determined by the tick species, type of host and their life stage (Rajput et al. 2006). Within minutes to hours, ticks body size can increase by hundred fold after bloodmeal consumption (Sonenshine, 1991; Mans and Neitz, 2004). Following that, excess water (60-70%) is secreted in order for tick to concentrate the ingested blood. The salivary glands help in secreting the excess water back into the host (Tatchell, 1967, Kaufman, 1983; Sonenshine, 1991; Mans and Neitz, 2004).

Argasidae or soft ticks of the genera *Argas, Ornithodoros*, *Otobius* among others, are considered to have veterinary and medical importance. Soft ticks possess a leathery integument instead of scutum (Sonenshine, 1991). Soft ticks of all life stages are rapid feeders, and it can take minutes to hours for them to be engorged. The volume of ingested blood is determined by their leathery integuments expansion. In general, upon ingestion the soft tick body weight can increase by ten fold within minutes to hours (Balashov, 1972; Sonenshine, 1991; Mans and Neitz, 2004). The water that is derived from the blood-meal is secreted back into the host via their coxal organs (Sonenshine, 1991; Mans and Neitz, 2004).

The *Nuttalliellidae* possess only one species, *Nuttalliella namaqua* (*N. namaqua*). This species shares several characteristics with soft and hard ticks (Latif et al. 2012; Mans et al. 2012). *N. namaqua* adults and nymphs possess a leathery integument and are fast feeders like argasids (Mans et al. 2011). Like ixodids, *N. namaqua* larvae possess a partly sclerotized pseudo-scutum and their mouthparts are located in an apical position (Mans et al. 2011; Bedford, 1931; Latif et al. 2012), and are slow feeders (Latif et al. 2012; Mans et al. 2012). Given that it shares characteristics between hard and soft ticks, this species has been described as the “evolutionary missing link” between the hard and soft tick families (Bedford, 1931). *N. namaqua* has been described as a generalist in terms of its host preferences and feeding style (Mans et al. 2014). A generalist host strategy allows parasites to switch hosts and parasitize different hosts in different geographic areas (Krasnov et al. 2008).

1.3. *Rhipicephalus (Boophilus) microplus*

This study is about the hard tick, *Rhipicephalus (Boophilus) microplus*, an obligate haematophagous parasite of both domestic and wild animals, known as the southern cattle tick (Brake et al. 2010). This tick species is amongst the most significant tick species, and is originally from South-East Asia (Walker et al. 2003; Jongejan and Uilenberg, 2004). *R. (B.) microplus* can be found on and infest various vertebrate hosts that include buffaloes, cattle, dogs, goats, horses, sheep and some wild animals (CSFPH, 2007). This tick species has a one-host life cycle. It can take about two months to complete the cycle (**Figure 1.1**). *R. (B.) microplus* can produce several generations per year, mostly during prolonged rainy seasons. The large numbers of larvae are usually found on pastures in late spring and summer. They then go through into the cooler autumn and early winter months (Madder et al. 2012a).

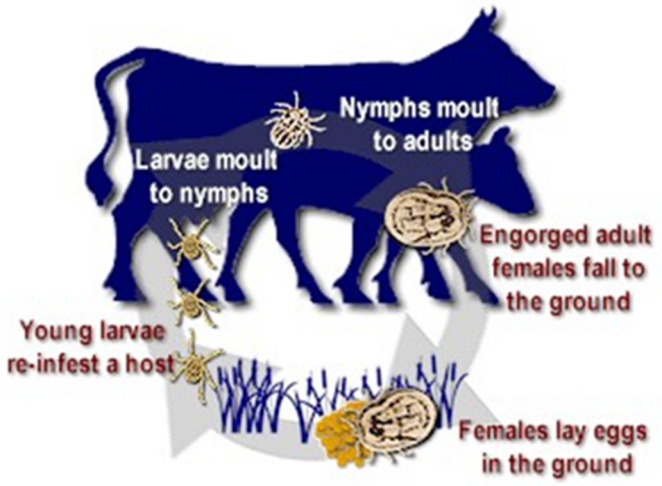

Figure 1.1: An example of one-host life cycle, like *Rhipicephalus (Boophilus) microplus*. All the life stages are on one vertebrate host. Adapted from: http://parasitipedia.net (accessed: 28 January 2015).

Globally, *Rhipicephalus (Boophilus) microplus* has over the years, spread throughout the subtropical and tropical regions (Estrada-Pena et al. 2006; Parizi et al. 2009). This tick species is currently found in continents such as Australia; Madagascar; South and Central America and East and Southern Africa (Jongejan and Uilenberg, 2004; Lynen et al. 2008). This species is considered the most economically important tick species from a global perspective (Guerrero et al. 2006). Babesiosis and Anaplasmosis could probably be the most important tick diseases affecting agriculture globally. The

Babesiosis-Anaplasmosis complex is responsible for severe economic losses in livestock and livestock products trading (de Castro, 1997).

1.4. Tick control and vaccines need

Control of ticks is important and needed to prevent livestock diseases caused by tick-transmitted pathogens (Latif and Walker, 2004). For a very long time the use of chemical acaricides has been considered as one of the best tick control methods (Martins et al. 1995). The application of such acaricides on vertebrate hosts aims at killing the ticks, but not harmful to livestock and/or applicators. In addition, the tissues of treated animals and the environment should not be contaminated (Drummond, 1983). However, this method of prevention has not yielded desired outcomes.

Over the years, the use of chemical acaricides was shown to have certain shortcomings. The control of tick diseases using chemical acaricides is normally accompanied by presence of chemical residues in milk and meat, chemical resistant tick strains, harmful effects on humans and environment (Willadsen and Kemp, 1988; Garcia-Garcia et al. 2000). Intensive use of acaricides eliminates livestock exposure to disease pathogens and results in failure to acquire natural immunity (Latif and Walker, 2004). Consequently, this will lead to livestock to have increased susceptibility to diseases and high disease endemic.

Furthermore, acaricides are expensive and pose a problem for end-users (Jongejan and Uilenberg, 2004; Rajput et al. 2006). Due to all these shortcomings with chemical acaricides, intensive and expensive dipping and/or spraying programmes remained the best alternative to eradicate and prevent tick infestation and tick-borne diseases. However, these programmes have not been meeting the desired results in terms of eradicating ticks and tick-borne diseases (Jongejan and Uilenberg, 2004). Consequently, all these lead to the need for the development of an alternate and complete control method (Graf et al. 2004), such as vaccination.

1.4.1. Importance of tick control

Tick control measures have relied on acaricides; however, this is accompanied by several disadvantages. These disadvantages include among others the acaricide-resistant ticks selection; environment, milk and meat contamination (Graf et al. 2004; de la Fuente and Kocan; 2006, Willadsen, 2006; Ghosh et al. 2007). All the problems and drawbacks associated with the use of acaricides have made acaricides to be less sufficient and not very effective. As a result, there is a need for alternative tick control methods. Among those control methods is the development of novel vaccines (de la Fuente and Kocan, 2006; Willadsen, 2006). Tick infestations and pathogen infections can be better combacted through vaccination. Vaccination offers the best alternative option due to its environmental-friendliness (Merino et al. 2013). Two tick vaccines for cattle tick infestation, TickGARD (Australia) and Gavac (Latin American countries) were commercially introduced in the early 1990's. They are both based on *R. (B.) microplus* midgut membrane-bound recombinant protein Bm86 (Willadsen et al. 1995; Canales et al. 1997; de la Fuente et al. 2007b). Several epidermal growth factor (EGF)-like domains are found in the Bm86 protein. Due to their nature, these domains may have a role in cell growth and blood coagulation. The Bm86 protein is expressed at all tick life stages as well as in the gut and ovaries of partially engorged females (Bastos et al. 2010). Andreotti et al. (2008) identified two distinct parts in the Bm86-CG sequence that might have a role in antibody recognition. In cattle, a positive correlation between antigen-specific antibodies and reduction of tick infestations and fertility is observed for Bm86-based vaccines (Rodriguez et al. 1995; de la Fuente et al. 1998; Merino et al. 2011a). However, the Bm86 mechanism of action remain unknown. When the Bm86 proteins are expressed, their action against tick infestation, lead to reduction of the number, weight and reproductivity of engorging female ticks (de la Fuente et al. 1998; de la Fuente et al.1999).

A number of studies have previously shown the effectiveness of the Bm86 vaccines when used for tick control. Canales et al. (2009) cloned the Ba86 and Bm86 genes from *R. (Boophilus) annulatus* and *R. (B.) microplus*, respectively. Vaccination of cattle with Bm86 caused reduction in *R. (B.) annulatus* and *R. (B.) microplus* numbers, and eggs weight and eggs fertility, tick weight and oviposition. On the other hand, Ba86 did not elicit any protective immune responses against both *R. (B.) annulatus* and *R.*

(B.) microplus. In another study, vaccines prepared from Bm86 resulted in reduction of engorged *R. decoloratus* adult female ticks, egg and tick weight (Odongo et al. 2007).

Vaccines based on Bm86 in most cases elicit protective immune responses against *R. (B.) microplus*. However, strain-to-strain sequence variations in the antigen locus between *R. (B.) microplus* isolates from different geographical regions and bovine populations have been shown to influence the vaccine efficacy (Penichet et al. 1994; Rodriguez et al. 1994, 1995a, b; Andreotti et al. 2008; Freeman et al. 2010; Ben Said et al. 2012). It was suggested that variations in the amino acid sequence exceeding 2.8% would be sufficient to cause vaccine inefficiencies (Lambertz et al. 2012). In South America strains divergences of up to 6.1 % and 4.6% in the sequence of the Bm86 and Bm95 protein, respectively, were found by Sossai et al. (Sossai et al. 2005). Therefore, the discovery of new tick antigens focusing on those displaying minimal genetic variability among *R. (B.) microplus* strains could improve vaccination efficacy and reduce variation in the protection level afforded by the Bm86-based vaccines (Parizi et al. 2012a, b).

Due to the reported variability in its efficacy and other industrial considerations, TickGARD is no longer commercially available. However, GAVAC is still in the market in North and South America despite reported shortcomings and a lack of widespread public acceptance of this vaccine (Guerrero et al. 2012; 2014). Given the shortcomings of the Bm86 vaccines, discovery of novel vaccines is needed to combat tick infestations, consequently, control ticks. However, production of vaccines is currently hindered by successful identification of effective antigens (Mulenga et al. 2000; Willadsen, 2001). The technological advancements in characterization of tick genomes, with the use of functional genomics, bioinformatics, transcriptomics, proteomics combined with reverse vaccinology have allowed for a rapid, systematic and comprehensive approach for numerous potential vaccine candidates identification (de la Fuente and Kocan, 2006; Maritz-Olivier et al. 2012).

African countries are dominated by tick species from the *Rhipicephalus (Boophilus)* genus. More than 50% of tick species in that genus are expected to expand their range. Of that range expansion, more than 70% is attributed to economically important tick species such as *R. (B.) microplus*, *R. (B.) decoloratus* or *R. appendiculatus*

(Estrada-Pena et al. 2006; Olwoch et al. 2007). It has been reported that *R. (B.) microplus* is rapidly expanding and is currently found in places in Africa where it never existed. Another phenomenon is that this tick has been reported to have displaced *R. (B.) decoloratus* on the African continent (Tønnesen et al. 2004). Given that there has never been an anti- *R. (B.) microplus* vaccine in Africa, it is therefore, important to sequence the genome of *R. (B.) microplus*. As part of that, our study aimed at analysing the *R. (B.) microplus* transcriptome in an attempt to identify potential candidate antigens that can be used in vaccine discovery.

A number of studies have attempted to characterize the *R. (B.) microplus* transcriptome on a large scale. That has only recently been made possible due to the availability of a partially assembled genome and large EST databases (Lee et al. 2005; Wang et al. 2007; Bellgard et al. 2012). The BmiGI website is one of such databases and was launched in 2005 (Guerrero et al. 2005). Studies that focus on the transcriptome of ticks are severely affected by the fact that most *R. (B.) microplus* nucleic acid sequences cannot be annotated (Wang et al. 2007; Bellgard et al. 2012). Up to 60% of the genome lacks similarity with sequences from other organisms (Wang et al. 2007), thereby negatively affecting the confidence of results for downstream analyses. Given that, the introduction of the new generation sequencing technologies has improved the throughput of transcriptome deep sequencing for a number of tick species (Karim et al. 2011; Schwarz et al. 2013, 2014). That has also been coupled with analysis of tick proteomes in an attempt to annotate their genomes, and subsequently identify anti-tick vaccines antigens. These efforts have afforded the tick vaccine research community with promising results towards achieving their goal of finding tick vaccines.

1.5. Vaccine types: Exposed vs hidden vaccines

Tick vaccines are based on either exposed or concealed antigens (Domingos et al. 2013). These antigenic targets have been explored as part of vaccine development. Exposed antigens are proteins or peptides produced in the salivary glands. During tick attachment and feeding on the host these antigens are secreted in saliva. These antigens are then taken up by host dendritic cells and processed before being

presented to T lymphocytes. As a result, that induce a cell or antibody-mediated immune response (Allen et al. 1979a; Nithiuthai and Allen, 1985; Larregina and Falo, 2005).

Concealed antigens are not exposed to the host immune system (Willadsen and Kemp, 1988), and are found in the tick gut wall. They normally interact with specific immunoglobulins taken up in the blood meal. However, any antigen not normally presented to the host is considered as a concealed antigen. It is also a potential vaccine candidate if such antigen is associated with some vital function for the tick. The antigen should encounter immunoglobulins entering the haemolymph to accomplish such function. Immunizations need to be repeated to maintain high antibody titers (Willadsen and Kemp, 1988). Concealed and exposed antigens possess different modes of action of vaccination; however, their results of vaccination are similar. They both reduce numbers of engorged ticks, increase mortality and reduce eggs laid by females.

1.5.1. Exposed vaccines

Numerous exposed antigens have been biologically tested as vaccines in anti-tick activities. However, in general, the results have not been a success (Nuttall et al. 2006). Consequently, none of these recombinant antigens is being developed and produced commercially. Some of exposed antigens evaluated as recombinant anti-tick vaccines are shown in **Table 1.1**. Mulenga et al. (1999) identified HL34 (saliva protein) and p29 (putative cement protein) from *H. longicornis* by immunoscreening and immunological approaches, respectively. Both recombinant proteins were tested on rabbits for their activities. The recombinant HL34 induced diseases and death in adult and nymphal *H. longicornis*. The immunization of rabbits with the recombinant p29 protein resulted to 40% and 56% mortality in *H. longicornis* larvae and nymphs, respectively. This protein also induced a reduction of 17% in adult ticks engorged weight.

A putative cement protein was identified from female and male *R. appendiculatus* salivary glands. This protein, 64P was found to cross-react with midgut epitopes, haemolymph and salivary glands of adult ticks, as well as *R. appendiculatus* nymphs and larvae whole-body extracts. The 64TRPs constructs act by disrupting tick

attachment and feeding, subsequently, causing tick mortality. These constructs are responsible for tick mortality of up to 80% (Trimnell et al. 2005). Moreover, an endoplasmic reticulum calcium-binding protein was found to be secreted in tick saliva (Jaworski et al. 1995; Ferreira et al. 2002). This protein has no known function and it was first identified in the salivary glands of female *A. americanum* ticks. Immunization of rabbits with recombinant *A. americanum* calreticulin led to the development of necrotic cutaneous lesions on tick challenge (Jaworski et al. 1995). Nevertheless, such action could not be seen in cattle immunized with recombinant calreticulin of *R. (B.) microplus*. Although the same recombinant protein was recognized by sera from dogs infested with *R. sanguineus* ticks (Ferreira et al. 2002).

Table 1.1: Exposed antigens evaluated as recombinant anti-tick vaccines.

Antigen	Tick species	References
Histamine-binding protein	*R. appendiculatus*	Paesen et al. 1999
P29	*H. longicornis*	Mulenga et al. 1999
HL34	*H. longicornis*	Tsuda et al. 2001
RIM36 cocktail	*R. appendiculatus*	Imamura et al. 2008
64TRPS	*R. appendiculatus*	Havlíková et al. 2009
Calceticulin	*A. americanum*	Jaworski et al. 1995, Ferreira et al. 2002
	D. variabilis, *R. (B.) microplus*	

1.5.2. Hidden (concealed) vaccines

There is no immune response induced by concealed antigens during tick attachment and feeding. However, these antigens are immunogenic as recombinant proteins. Some of concealed antigens evaluated as recombinant anti-tick vaccines are shown in **Table 1.2**. The well-defined example is the *R. (B.) microplus* Bm86 midgut membrane-bound recombinant protein (Willadsen et al., 1995). The blood from immunized host has specific immunoglobulins that interact with these gut wall antigens in a specific manner. In female ticks, this interaction cause gut wall lysis and interfere with blood digestion. Subsequently, this results in no egg production or kills the tick. Since Bm86 antigens are concealed antigens, natural infestations do not induce an anamnestic response. Hence, Bm86-based vaccinations are short-lived (Willadsen and Jongejan, 1999; Frisch, 1999), and require booster vaccinations in 6-month intervals (de la Fuente et al. 1998). A number of antigens have been isolated from various tissues and expressed as recombinant proteins. Such proteins have been tested as concealed vaccine candidates. The *R. (B.) microplus* Bm86 homologue, Bm95 glycoprotein was shown to protect against infestations by South African cattle tick strains not protected by Bm86 vaccination (Canales et al. 2009a). Other studies have previously shown Bm95 to protect against a wider range of tick strain infestations. This suggested that it could be a universal antigen against infestations by different geographical *R. (B.) microplus* strains (Garcia-Garcia et al. 2000; de la Fuente and Kocan, 2003).

Ba86 is a recombinant *R. (B.) annulatus* Bm86 orthologue protein that possesses more than 90% identity to Bm86 (Canales et al. 2008). Its efficacy together with that of recombinant Bm86 is higher against *R. (B.) annulatus* than for *R. (B.) microplus* infestations. These results suggested biological differences between *R. (B.) microplus* and *R. (B.) annulatus* (Canales et al. 2009a). Another Bm86 orthologue from *H. a. anatolicum*, was cloned and expressed by Azhahianambi et al. (2009) and was termed, HAA86. The following year it was reported that vaccination of cattle with the recombinant HHA86 antigen protected against homologous tick infestation. Another interesting phenomenon was discovered as this antigen also reduced transmission of *T. annulata* in ticks (Jeyabal et al. 2010).

A novel protein, ATAQ, from metastriate ticks was identified and shares structural similarities with Bm86. ATAQ is expressed in both midgut and Malpighian tubules, and Bm86 is only expressed in midgut. Because of that, ATAQ may offer increased cross-protection outcome against heterologous ticks compared to Bm86-based vaccines. Consequently, ATAQ may be a good vaccine candidate for tick vaccine (Nijhof et al. 2010). The other two candidates, serine proteinase inhibitors (HLS1, HLS2) are induced during blood feeding of *H. longicornis*. These serine proteinase inhibitors have potential in targeting tick homeostasis (Sugino et al. 2003, Imamura et al. 2005). In a rabbit model, mortality of both nymphal and adult *H. longicornis* was increased after vaccination with recombinant HLS1 and HLS2. A 21-kDa soluble cytoplasmic protein from *I. scapularis* protein, 4D8 resulted in infestation levels reduction, extended engorgement and smaller egg masses of adult *I. scapularis* upon vaccination (Almazán et al. 2005b). An *A. hebraeum* male-specific protein, voraxin that is induced during feeding, is another target antigen. There was a decrease in typical engorgement of mated female *A. hebraeum* in rabbits immunized with the voraxin recombinant protein (Weiss and Kaufman, 2004).

Table 1.2: Concealed antigens evaluated as recombinant anti-tick vaccines.

Antigen	Tick species	References
Bm86	*R. (B.) microplus*	Willadsen et al. 1995, Canales et al. 2009a
Bm91 (carboxy dipeptidase)	*R. (B.) microplus*	Willadsen et al. 1996
Bm95	*R. (B.) microplus*	Garcia-Garcia et al. 2000
Vitellin	*R. (B.) microplus*	Tellam et al. 2002
BmPRM (paramyosin)	*R. (B.) microplus*	Ferreira et al. 2002
HLS1	*H. longicornis*	Sugino et al. 2003
HLS2	*H. longicornis*	Imamura et al. 2005
Voraxin	*A. hebraeum*	Weiss and Kaufman, 2004
P27/30	*H. longicornis*	You, 2005
4D8	*I. scapularis*	Almazán et al. 2005

1.6. Tick salivary glands molecules: good targets for vaccine development

During their attachment and feeding, ticks face defense responses from the host. Hosts respond by various ways such as haemostatic plug formation, vasoconstriction and inflammatory responses and coagulant cascade activation. All these cause disruption in tick feeding and cause rejection of the ticks (Kazimírová and Štibrániová, 2013). Ticks have to circumvent these responses in order to successfully complete their blood meal. An enormous number of biologically active molecules found in ticks salivary glands aid to oppose the host's defense mechanism. These molecules possess anticoagulation, antiplatelet, anti-inflammatory and immunomodulatory activities and vasodilatory (Kazimírová and Štibrániová, 2013). These molecules are believed to have developed during host-parasite co-evolution and are essential for tick feeding and development (Nuttall and Labuda, 2008; Francischetti et al. 2009; Mans, 2011). The saliva of ticks possesses several bioactive molecules/compounds that modulate the immune-inflammatory and haemostatic system of their hosts (Hajnická et al. 2011).There are differences in these bioactive compounds (Kazimírová, 2008) and range from lipids to large proteins (Ribeiro, 1995; Valenzuela, 2002a). Due to the importance of tick saliva in host's immune and haemostatic response, it is important to isolate, identify and characterised molecules found in tick saliva (Valenzuela, 2004). As a result, tick saliva has been studied extensively over the years. The focus on tick saliva due to its ability to offer the opportunities to identify and mine biologically active molecules with anticoagulant activity, anti-inflammatory, immunosuppressive and vasodilating (Oliveira et al. 2011).

Transcriptomes have been generated from isolated and sequenced cDNAs from different tick species. Both classical and modern molecular biology techniques strategies have played a huge role to achieve this goal. The ability to isolate these molecules will contribute to the studying, understanding and evaluation of these molecules for their potential candidates as vaccines. These candidates can be selected based on their anti-haemostatic, anti-inflammatory or immunosuppressor functions. Alternatively, the candidates can be selected as protective molecules irrespective of their functions (Valenzuela, 2004). Subsequently, this will facilitate in identification of possible vaccine candidates for the control, of ticks and tick-borne

disease (Valenzuela, 2004). Some of the tick salivary biologically active molecules have been studied and their functions have been identified. However, the majority of them still have unknown functions (Ribeiro et al., 2006; Brossard and Wikel, 2008; Francischetti et al. 2009; Fontaine et al. 2011). Current studies are focusing on identifying and preparing tick salivary gland molecules in their recombinant form.

1.7. Tick feeding and host interactions

Ticks begin their feeding after finding a host and selecting a preferred feeding site. They use their chelicerae to saw and penetrate through the host's skin. Ixodid ticks secrete a polymerising milky cement to secure them on the skin (Nuttall et al. 2000). The cement is made up of collagen-related glycosylated proteins (Mulenga et al. 1999). The cement formation around the host wound help to prevent leakage of tick saliva and/or host fluids. After their blood meal, ticks detach their mouthparts and leave the cement core still embedded in the host's skin (Nuttall et al. 2000). During feeding, hard ticks accommodate large volumes of blood and expand their cuticle to accommodate those volumes. They take the majority of the blood meal on the last day of attachment (Kemp et al. 1982).

Tick feeding has both physical and chemical effects on the host. Consequently, these effects incite haemostatic, inflammatory and immune responses. In turn, ticks secrete their saliva molecules to fight against these responses (Valenzuela, 2004; Kazimírová, 2008). These saliva molecules have already been discussed above (**Tick salivary glands molecules: good targets for vaccine development**) section. These bioactive molecules are also implicated with pathogen transmission. The tick-host interaction is an infective interface, although the tick attachment to the host is the way of pathogen transmission. The transmission of pathogens via tick saliva constituents to the host, is known as saliva-activated transmission (SAT). Tick saliva molecules are SAT factors that promote pathogen transmission (**Figure 1.2**) (Nuttall and Labuda, 2004).

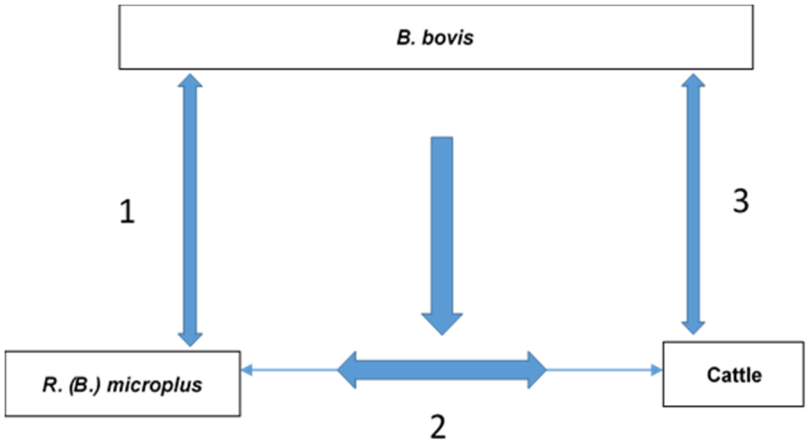

Figure 1.2: Tick-borne pathogens, tick vectors and vertebrate hosts' interactions, using *R. (B.) microplus* as example. (1) *B. bovis*-*R. (B.) microplus* interactions; (2) *R. (B.) microplus*-cattle interactions and (3) *B. bovis*-cattle interactions.

In many tick species, SAT factors are responsible for several different diseases, the majority being viral and bacterial diseases. The first report was on *Thogoto* virus transmitted by *R. appendiculatus* (Nuttall and Jones, 1991). Other studies reported transmission by flies and mosquitoes (Titus and Ribeiro, 1988; Belkaid et al. 1998; Osorio et al. 1996).

Bloodmeal is an essential requirement for hard ticks, such as *R. (B.) microplus* to develop in all their life stages (Liu and Bonnet, 2014). All *R. (B.) microplus* life stages occur on the host, which is normally the cattle. During the attachment of their hosts, ticks normally transmit pathogens such as *A. marginale* (Palmer et al. 1988; Aguirre et al. 1994). Once transmitted to the host, pathogens need adapt to the host' environmental conditions for effective survival and reproduction (Brossard and Wikel, 2004). During the feeding process, the exchange of fluids (saliva) occurs between the tick and its host. That saliva might contain some of the same pathogens that were transferred to the host. In turn, those pathogens can be re-transmitted to new vertebrate hosts during tick bloodmeal. In the case of *R. (B.) microplus*, certain pathogens are transferred via transovarial transmission to the next generation of ticks. This vertical transmission is necessary for those tick-borne pathogens infecting single-host tick species such as the *R. (B.) microplus*-transmitted *Babesia bovis* (Liu and Bonnet, 2014).

Larvae, nymphs and adult transmitted *A. marginale* to Holstein calves that were originally *Babesia* and *Anaplasma*-free (Aquirre et al. 1994). In another study, infection with extracts of both immature (nymphs) and mature (adults) stages of *R. (B.) microplus* resulted in successful transmission of *A. marginale* transmission into susceptible animals (Dalgliesh and Stewart, 1982). *Rhipicephalus (B.) microplus* is capable of transmitting pathogens other than *A. marginale, B. bigemina* or *B. bovis*. *R. (B.) microplus* can transmit other pathogens in other animals such horses and include *Babesia equi (B. equi)* during acute phase of infection (Knowles et al. 1992).

1.8. Tick feeding and survival: anti-hemostatics and immunomodulators

Both the transcriptome and proteome of many hard ticks have allowed for the identification of several anti-hemostatics and immunomodulators. These molecules have an important role in tick feeding and survival. Tick anti-hemostatics and immunomodulators are essential in hemostasis. Homeostasis helps in vascular constriction, platelet aggregation and blood coagulation. These anti-hemostatics and immunomodulators have been found in salivary glands, saliva, eggs and haemolymph. For the purpose of our study, we will only focus on anti-hemostatics and immunomodulators of *R. (B.) microplus* (**Table 1.3)**.

Table 1.3: List of anti-hemostatics and immunomodulators of *R. (B.) microplus*

Anti-hemostatic/immunomodulators	Source	Target	Reference
Boophilin	Engorged ticks	Thrombin	Macedo-Ribeiro et al. 2008
BmAP	Saliva	Thrombin, trypsin, plasmin	Horn et al. 2000
Microphilin	Saliva	Thrombin	Ciprandi et al. 2006
Gut membrane associated 5'-nucleotidase	Saliva	Platelet aggregation	Liyou et al. 1999
BmTI-A	Larvae	Kallikrein-kinin inhibitor	Sasaki et al. 2004
BmTI-D	Larvae	Kallikrein-kinin inhibitor	Sasaki et al. 2004
Serpin 15	Salivary glands, ovaries and guts	Thrombin	Rodriguez-Valle et al. 2015
Serpins	Salivary glands	Various	Tirloni et al. 2016
Kunitz	Salivary glands	Vessel	Soares et al. 2016
Kunitz	Larvae	Elastase/kallikrein	Tanaka et al. 1999
Bookase	Salivary glands	Bradkinin	Bistiani et al. 2002

1.9. Tick transcriptome studies

Over the years, sialotranscriptomes and/or transcriptomes of both soft and hard ticks have been studied intensively using expressed sequence tags (ESTs) based approaches. ESTs are normally produced from complementary deoxyribonucleic acid (cDNA) libraries and sequenced with traditional Sanger sequencing method. The number of these ESTs varies per study from 500 to 1000s (Valenzuela et al. 2002; Santos et al. 2004; Guerrero et al. 2005; Alarcon-Chaidez et al. 2007; Francischetti et al. 2008a, b; Chmelar et al. 2008; Anatriello et al. 2010; Ribeiro et al. 2011). ESTs constructed from cDNAs of *R. (B.) microplus* larvae, eggs, dissected organs, nymphs as well as both adult female and male ticks that were exposed to host odour, cold and heat shock, acaricides and infection with *Babesia bovis* have been analysed (Guerrero et al., 2005). Other studies have also used tick salivary glands and tick salivary glands in different feeding stages to study tick transcriptome (Santos et al. 2004; Chmelar et al. 2008). Ribeiro et al. (2011) revisited the *A. variegatum* sialotranscriptome using a combination of ESTs and proteomics approach. This approach could only reveal more secreted proteins, still leaving most proteins with unknown functions. That still posed a problem and proved the use of ESTs to have limitations.

Karim et al. (2011) provided the first transcriptome study employing the next generation sequencing approach. This study presented a revolution in tick sialotranscriptome analysis. Pyrosequencing was used to sequence the salivary gland cDNA library from adult female *A. maculatum* at different feeding times. This upcoming next generation sequencing approach allowed for a much deeper insight into rarer transcripts than previously done. This study produced 1.5 million reads instead of 1000s sequences enabling for deep coverage of rarer transcripts. The current catalogue of salivary transcripts of *A. maculatum* had a majority of genes encoding for housekeeping, secreted and unknown proteins. A single 454 pyrosequencing sequencing run can offer a fast and cost effective method for deep transcriptome sequencing. Subsequently, that has allowed for the considerable improvement on previously used sequencing methods (Meyer et al. 2009).

Karim's pioneering work laid a good foundation and it was followed by other studies seeking to characterize several tick transcriptomes. These studies sought to get deep coverage by using one or more of the next generation sequencing technologies simultaneously. The success of using two NGS technologies on a study provided highthroughput in terms of high transcriptome coverage. One good example of such was the study by Schwarz et al (2013) where four *I. ricinus* salivary gland cDNA samples were prepared from early and late-feeding nymphs or adults. These samples were sequenced to produce 441,381 and 67,703,183 reads with 454 pyrosequencing and Illumina sequencing, respectively. Assembly of these reads yielded 272,220 contigs. As a result, annotation of 34,560 contigs allowed for 8,686 coding sequences to be deposited to GenBank. Of those contigs 13% was associated with secreted proteins family. This family was differentially represented among the four experimental samples. Different transcripts showed high expression and were more represented in certain samples. This data improved and expanded the available genomic information for *I. ricinus* for future transcriptome assemblies and functional analysis of identified proteins.

The following year Schwarz et al (2014) expanded their study by combining the transcriptome and proteome of both the salivary glands and midgut of *I. ricinus*. This study combined the next generation sequencing and proteomics to deliver an unprecedented and quantitative description of gene expression in the midgut and salivary glands of *I. ricinus* upon attachment on the host. The transcriptome was improved by the generation of an additional 315 million Illumina sequence reads from four cDNA libraries of *I. ricinus* nymphs, and both adult females and males. This was the first in-depth combined analysis of the midgut and salivary gland proteome and transcriptome of *I. ricinus* upon attachment on the host. Over 1,500 genes were profiled and they showed tissue-specific and developmental stage-specific expression at both transcriptomic and proteomic level. About 25% of identified proteins had higher expression in the salivary glands, but only 7.3% had corresponding high transcript levels in the same tissue. These results were promising, especially for an organism with an unsequenced genome. This study proved that proteomics is possible in non-model organisms with limited transcriptome information, provided that sufficient transcriptome coverage of the studied sample is available. This also applicable in our study of *R. (B.) microplus* where transcriptomics was coupled with proteomics to

characterize salivary gland transcriptome at different feeding stages of female *R. (B.) microplus* ticks on the vertebrate host. It can be argued that the success of this study might hold a promise to the success of our own study; however, their experimental approaches differ.

Schwarz et al (2013) *de novo* assembled 441,381 454 pyrosequencing reads and 67,703,183 Illumina reads from nymphal and adult *I. ricinus* salivary glands. Those reads were combined with 268,914,130 reads from the nymphal and adult salivary glands and midgut of *I. ricinus* (Kotsyfakis et al. 2015a). The total reads from both studies allowed for extraction of 25,808 coding sequences (CDS) larger than 150 nucleotides. As a result, a total of 148,813,058 reads were mapped on the 25,808 CDS. The CDS were classified into five categories: secreted, housekeeping, transposable elements, viral and unknown. Together, secreted and housekeeping accounted for 85% of the CDS and over 96% of the mapped reads. This indicated the success of the massive *de novo* sequencing and improved coverage achieved by combining two NGS technologies and three sets of reads from the current study and the previous study (Schwarz et al. 2013). The secreted class contained many families including, lipocalins, Kunitz, and metalloproteases. Members of these families showed stage and time-specific expression and diversity. Their diversity and selective expression may suggest antigenic variation (Kotsyfakis et al. 2015b). In this study, *de novo* assembly of 926,596 pyrosequencing reads and 49,328,982 Illumina reads were produced from sequencing the *I. ricinus* haemocytome.

The previous study was followed by characterization of *I. ricinus* haemocytome (Kotsyfakis et al. 2015b). The total reads from this study were combined with over 189 million reads from Schwarz et al (2014) to produce over 319 million Illumina reads. From those reads, a total of 15,716 coding sequences were obtained and 2,860 reads could be mapped into the total reads. Those 2,860 reads were classified into seven major classes, which include secreted and housekeeping proteins. The putative secreted proteins accounted for 33% and 44% of coding sequences and 44% of the mapped reads, respectively. The secreted proteins included serine proteases, cathepsins, metalloproteases and protease inhibitors. The housekeeping proteins were also predominant and involved those that play roles in metabolic pathways or cytoskeleton which accounted for 50% of the coding sequences and 46% of mapped

reads. Protein with no known function represented 8,5% of the total coding sequences and 2.6% of the mapped reads. The additional 2,860 coding sequences from this study with the previous deposition of 24,687 coding sequences, increased public available *I. ricinus* genomic sequences. These sequences will provide mining platform for biochemical and physiological studies aiming at interrogating ticks genomes.

Transcriptomes have been used to analyse tick tissues other than salivary glands or midgut. Egekwu et al (2014) used Illumina sequencing and 454 pyrosequencing for characterization of female *I. scapularis* synganglion genes. In this study, it was shown that Illumina GAIIx sequencing was more efficient than 454 pyrosequencing in terms of total read coverage, GO term searching and annotation of important gene transcripts and gene categories. Total read coverage by Illumina sequencing ranged from 34 to 117 million reads for the samples, whereas coverage by 454 pyrosequencing was only 4, 6 million reads. More than 60% of transcripts were predicted by Illumina sequencing. Transcriptomes have proven useful for global analysis and categorizing of genes from selected tissues or even whole animals. Next generation sequencing has made it possible to examine the complex range of genes associated with selected biological functions. The application of NGS technologies in transcriptome and proteome of several tick species has shown success in terms of highthroughput. All the above-discussed studies have generated millions of sequence reads. Those reads allowed for de novo assembly of millions and thousands of contigs. As a result, thousands of coding sequences and proteins were produced/identified. The pattern in terms of number of reads, contigs and proteins produced/identified is similar to that of the findings in our study. The results from some these studies, including our study, are summarized in **Table 1.4**.

Table 1.4: Some of the transcriptomes and proteomes studies achieved by NGS technologies

Tick species	Organ	Number of contigs/transcripts	Number of mapped reads	Number of CDS/proteins	NGS technology	Reference
I. ricinus	Salivary glands and midguts	N/A	148,813,058	25,808	Illumina sequencing and 454 Pyrosequencing	Kotsyfakis et al. 2015a
I. ricinus	Salivary glands and midguts	N/A	50,255,578	15,716	Illumina sequencing and 454 Pyrosequencing	Kotsyfakis et al. 2015a
A. maculatum	Salivary glands	N/A	1,626,969	15,914	454 Pyrosequencing	Karim et al. 2011
I. ricinus	Salivary glands	272,220	68,144,564	8686	Illumina sequencing and 454 Pyrosequencing	Schwarz et al. 2013
I. ricinus	Salivary glands	N/A	315,000,000	25,808	Illumina sequencing	Schwarz et al. 2014
I. scapularis	Synganglion	92,717	152,876,879	N/A	Illumina sequencing and 454 Pyrosequencing	Egekwu et al. 2014
R. (B.) microplus	Salivary glands	N/A	29,129,146	30,537	Illumina sequencing	Genu, 2017 (MSc dissertation)
R. appendiculatus	Salivary glands	21,410	Approximately 430,000,000	12,761	Illumina sequencing	De Castro et al. 2016

1.10. Proteomics studies on ticks

Proteomics remains as one of the significant area of studies in the post-genomics era. Moreover, proteomics can be explained as the systematic identification of protein sequence, abundance, activity, interactions, post-translational modifications, etc. All these can be characterized in a given sample, for example, salivary glands from hematophagous arthropods (Pandey and Mann, 2000). Proteomics has appeared to be a powerful new instrument that includes strategies for the characterization of dynamic interactions that cannot be analysed by genomics/transcriptomics approaches alone. The information provided by proteomics on the protein content of cells and/or tissues may be different at the transcriptomic level. Such information may assist in identifying antigens that might have potential to be used in tick vaccine development (Villar et al. 2010; de la Fuente et al. 2010, 2012; Marcelino et al. 2012).

The first paper published on tick proteome was by Madden et al. (2002). In this study, 2-DE gels and MALDI-MS were used for characterization of protein profiles of *A. americanum* and *A. maculatum*. That pioneer study was followed by several publications of papers in the tick field using the same proteomics method (Untalan et al. 2005; Oleaga et al. 2007; Rachinsky et al. 2007; Vennestrøm and Pensen, 2007, Kongsuwan et al. 2010; Villar et al. 2010, 2012; Marcelino et al. 2012. The sequence databases with genomic and proteomic information have been assembled for various arthropod species (including ticks) of agricultural and medical importance. Such databases are constantly increasing and that has enabled the expansion of research into the field of proteomics (Rachinsky et al. 2007).

Both hard and soft tick salivary glands have been studied extensively with the application of the proteomic approach (Madden et al. 2002; Valenzuela et al. 2002a; Oleaga et al. 2007; Chmelar et al. 2008). Untalan et al. (2005) reported the early only large-scale *R. (B.) microplus* proteome study, which resulted in sequence information for 20 abundantly expressed larval, representing multiple cuticular proteins, cytoskeletal protein and salivary gland-associated protein, several housekeeping proteins and tropomyosin. On the other hand, considerable efforts have also been dedicated to studying salivary gland transcriptomes and proteomes.

One of the successful proteomic analysis was done on *R. (B.) microplus* saliva. Such study resulted in identification of 187 tick proteins and 68 bovine proteins (Tirloni et al. 2014). *R. (B.) microplus* saliva is rich in mainly, among other proteins; secreted conserved proteins, antimicrobial peptides, glycine-rich proteins, housekeeping proteins, lipocalins, peptidase inhibitors and host proteins. The results from this study improved the knowledge about tick salivary modulators that play a role(s) in host defense during tick feeding. Moreover, those proteins offer unique understanding about tick-host relationship (Tirloni et al. 2014). The difference in protein composition in tick saliva suggest that tick salivary protein profiles undergo changes during tick feeding (Leboulle et al. 2002, McSwain et al. 1982, Binnington, 1978).

Recently, Kim et al (2016) identified both tick and host proteins from saliva of *I. scapularis* every 24 hours through the first five days of feeding and up to the completion of feeding. These saliva proteins were characterized by LC-MS/MS. This allowed for identification of 582 tick and 83 from the tick and rabbit proteins in saliva of *I. scapularis*. This study provided identification of *I. scapularis* tick saliva proteins that are involved in tick attachment onto host skin, feeding and tick-borne diseases transmission.

The studies discussed above are some of several hard tick proteome studies. Their main aim was to identify protein mostly, but not limited to, tick saliva/salivary glands, midgut and ovaries. They all employed different proteomics approaches; some studies used more than one method to improve the outcome and to meet their expectations.

Proteomic knowledge from these studies could also be useful in various studies of identified proteins, for example phylogenetic analysis (Louw et al. 2013). Comparison of the sequence divergence of different or same protein families could benefit from proteomic approach. In soft ticks, transcriptomic and proteomic studies have been carried out to investigate the phylogenetic linkage between soft and hard ticks and understand the mechanism of hematophagy (Oleaga et al. 2007, Francischetti et al. 2008a, b; Mans et al. 2008b).

1.11. Overall aim and objectives

1.11.1. Overall aim

The aim of this study was to characterize the salivary gland transcriptome and proteome of female *Rhipicephalus (Boophilus) microplus* at different stages of feeding.

1.11.2. Objectives

• Construction of cDNA library from different feeding stages of *R. (B.) microplus* female ticks.

• Next generation sequencing with the Illumina TruSeq RNA Sample Preparation method and MiSeq sequencing of different female feeding stages.

• Assembly and annotation of transcriptome.

• Proteomic analysis of the salivary gland extracts from different female feeding stages.

Chapter 2: Experimental approach and methodology

2.1. Introduction

2.1.1. Transcriptomics and next-generation sequencing (NGS)

The transcriptome can be described as a complete set of RNA molecules and their quantity in a given organism/organ/tissue/cell for a specific developmental stage or physiological condition. It is composed of both coding and non-coding RNAs and their degradation products. The transcriptome undergoes constant qualitative and quantitative changes that are triggered by external stimuli. The analysis of the genome functional elements requires the understanding of the transcriptome. Transcriptomics serve to characterize all types of transcripts in an organism. That, in turn, helps to measure how much is the production of each transcript during different biological and physiological conditions (Wang et al. 2009; Żmieńko et al. 2011, Martin and Wang, 2011).

The sequencing of different organism's transcriptomes started over two decades ago. All sought to characterize transcriptome components by applying different techniques. These techniques include expressed sequence tags (ESTs) (Gerhard et al. 2004), serial analysis of gene expression (SAGE) (Velculescu et al. 1995), massively parallel signature sequencing (MPSS) (Brenner et al. 2000). These methods possess certain drawbacks, as they are laborious and relatively expensive. Next-generation sequencing (NGS) technologies have recently been applied to transcriptome analysis research and/or studies. This approach is called RNA-sequencing (RNA-seq) and unlike the classical methods, it does not require the cloning step. The RNA-seq-based techniques produce exceptional quantities of data. These techniques provide the possibility to explore many transcriptomes in short periods of time simultaneously (Wang et al. 2009; Żmieńko et al. 2011). In transcriptome sequencing there is no need for prior knowledge about the genome sequence. Therefore, it can be applicable to any organism and allows not just analysis but also transcript discovery (Żmieńko et al. 2011). The most famous traditional DNA sequencing method, Sanger, which produce ESTs is discussed below.

(a) The traditional sequencing method/Sanger sequencing method

Large-scale sequencing projects have commonly required the cloning, amplification and purification steps prior to Sanger sequencing (Sanger et al. 1977). The Sanger sequencing method is based on the complementary DNA (cDNA) strand synthesis. For successful synthesis, natural 2'-deoxynucleotides (dNTPs) and 2', 3'-dideoxynucleotides (ddNTPs) are needed to serve as nonreversible synthesis terminators (Sanger et al. 1977). The DNA synthesis reaction is randomly terminated by the addition of a ddNTP to the growing oligonucleotide chain. The synthesis produce truncated products with varying different length sizes. The products are then separated (resolved) by size polyacrylamide gel or capillary electrophoresis. Finally, the DNA sequence of the template strand is revealed by terminal ddNTPs (**Figure 2.1**).

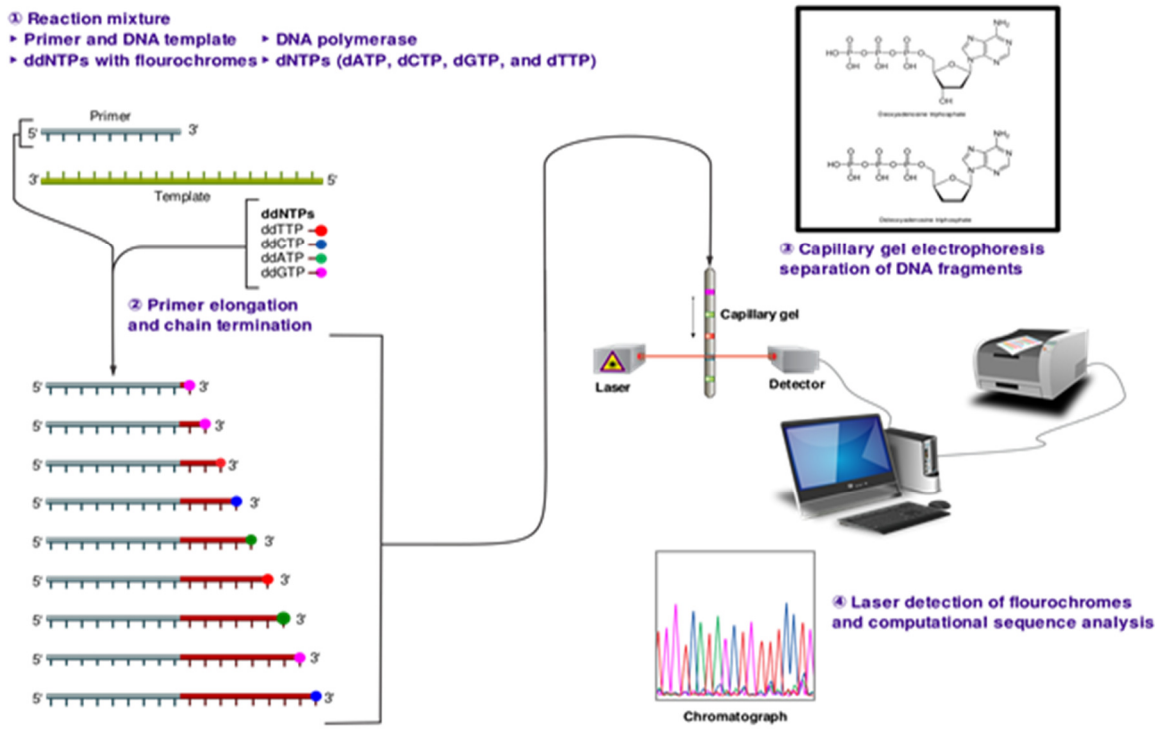

Figure 2.1: The workflow of classical Sanger sequencing method. Adapted from: http://en.wikipedia.org/wiki.Sanger_sequencing (03 May 2015).

The basic sequencing method requires four different reactions per template. For each reaction a different ddNTP terminator, ddATP, ddCTP, ddTTP or ddGTP is needed. Advancements in fluorescence detection enabled the combination of the four terminators in one single reaction. This was achieved by labeling them with fluorescent dyes of different colours (Smith et al. 1986; Prober et al. 1987). In later years the

capillary gel electrophoresis replaced the original slab gel electrophoresis (Margulies et al. 2005). That enhanced the separation of fragments (Cohen et al. 1988) (**Figure. 2.1**).

(b) Next Generation Sequencing (NGS) technologies

A number of NGS companies and technologies have been created, and the corresponding field of bioinformatics has developed significantly. The rapid advancement of instruments, chemistries and techniques led to NGS instruments to changing within months and, chemistries and analysis algorithms changing within weeks (Levy and Myers, 2016). Several reviews over the last several years have described the technological landscape of sequencing (Metzker, 2010; Morey et al. 2013; Reuter et al. 2015).

In principle, massive parallel sequencing systems are powerful technology rigs that integrate basic molecular biology, automated fluidics devices, high-throughput microscopic imaging and information technologies. In order for one to be able to use these systems, a broad understanding of the complexities underlying molecular biology and biochemistry is required. The high-throughput instruments are essentially high-tech machines and understanding the engineering principles gives the user the ability to command and troubleshoot the massive parallel sequencing systems. With the introduction of these technologies, it is required that users acquire computational skills and develop systematics data analysis pipelines (Myllykangas et al. 2012).

Commercial high-throughput sequencing platforms share three critical steps: DNA sample preparation, immobilization and sequencing (**Figure 2.2**). These high-throughput sequencing systems require amplification of the sequencing library DNA to form spatially distinct and detectable sequencing features. Amplification can be performed *in situ*, emulsion or in solution to generate clusters of clonal DNA copies.

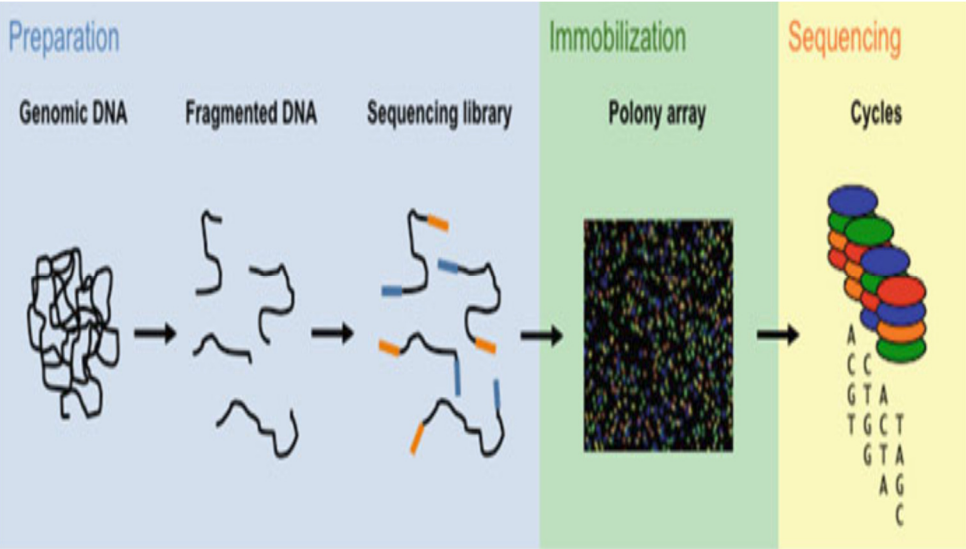

Figure 2.2: The NGS high-throughput sequencing workflow. The high-throughput sequencing involves three main steps in high-throughput sequencing: preparation, immobilization, and sequencing. (Adapted from Myllykangas et al. 2012).

Next generation sequencing (NGS) has brought about a most effective method known as RNA-sequencing (RNA-seq). The RNA-seq is a powerful method for mapping and measuring transcripts in biological samples. In 2005, two new next-generation or deep-sequencing sequencing (NGS) technologies were introduced both based on sequencing-by-synthesis. They both held a promise to take over from and/or improve old sequencing methods. The 454 system (Roche Applied Science) uses pyrosequencing technology (Margulies et al. 2005) and the Solexa system (Illumina, Inc.) that detect fluorescence signal (Porreca et al. 2007). In 2008, Applied Biosystems introduced their SOLiD sequencer, another short-read platform. However, these new methods provided shorter read lengths than capillary sequencing, with average size of 35-500 base pairs (bp) (Metzker, 2010). With the help of new complementary computational tools, NGS enables high-throughput DNA sequencing and assembly, thus yield deep coverage of transcriptomes (Pareek et al. 2011, Strickler et al. 2012). This offers dramatic increases in cost-effective sequence throughput and a variation of methodologies for characterization of transcriptomes, including single-end and pair-end cDNA sequencing. The timeline of advances in DNA sequencing over the years is shown in **Figure 2.3** and compared in **Table 2.1**.

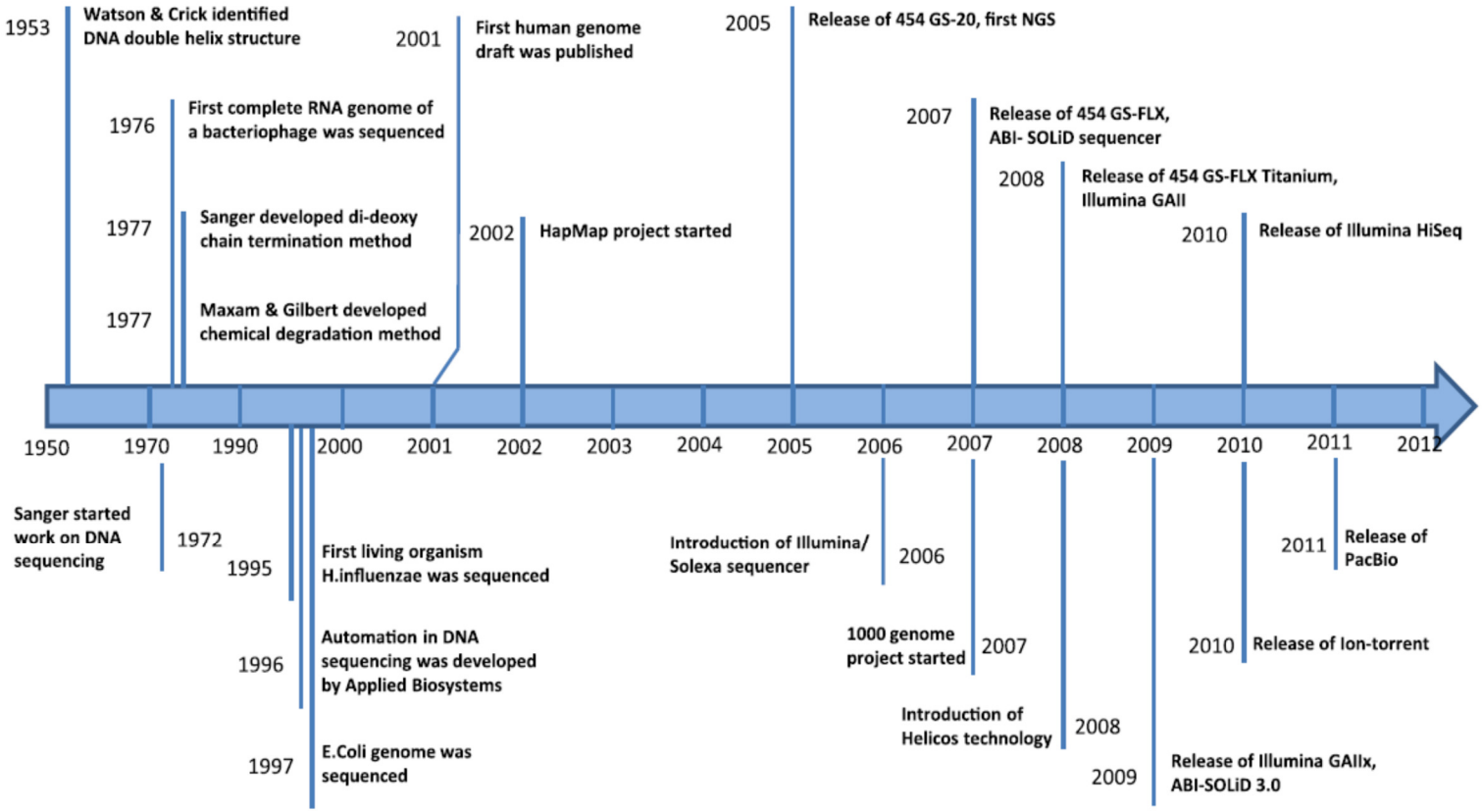

Figure 2.3: The time line of next-generation sequencing (NGS) technologies. A number of projects had used DNA sequencing to study various organisms. (Adapted from: Gupta and Gupta, 2014).

Table 2.1: Comparison of features of NGS platforms/technologies used for transcriptome applications (Metzker, 2010)

Platform	Library/ template preparation	NGS chemistry	Read length (bases)	Run time (days)	Gb per run	Machine cost (US$)	Pros	Cons	Biological applications	Refs
Roche/454's GS FLX Titanium	Frag, MP/ emPCR	PS	330*	0.35	0.45	500,000	Longer reads improve mapping in repetitive regions; fast run times	High reagent cost; high error rates in homo-polymer repeats	Bacterial and insect genome *de novo* assemblies; medium scale (<3 Mb) exome capture; 16S in metagenomics	D. Muzny, pers. comm.
Illumina/ Solexa's GA_{II}	Frag, MP/ solid-phase	RTs	75 or 100	4†, 9§	18†, 35§	540,000	Currently the most widely used platform in the field	Low multiplexing capability of samples	Variant discovery by whole-genome resequencing or whole-exome capture; gene discovery in metagenomics	D. Muzny, pers. comm.
Life/APG's SOLiD 3	Frag, MP/ emPCR	Cleavable probe SBL	50	7†, 14§	30†, 50§	595,000	Two-base encoding provides inherent error correction	Long run times	Variant discovery by whole-genome resequencing or whole-exome capture; gene discovery in metagenomics	D. Muzny, pers. comm.

NGS technology is rapidly developing and has become one of the most popular and crucial techniques used today in several fields of biological sciences (Mitra et al. 2015). The currently dominant platform is Illumina, used for approximately 85% of samples deposited in NCBI's Sequence Read Archive (SRA) in (Leinonen et al. 2011). This technology has been shown to be a very promising approach for deep sequencing of the transcriptome. Read length and throughput have, over the years, increased and improved, respectively (**Figure 2.4**).

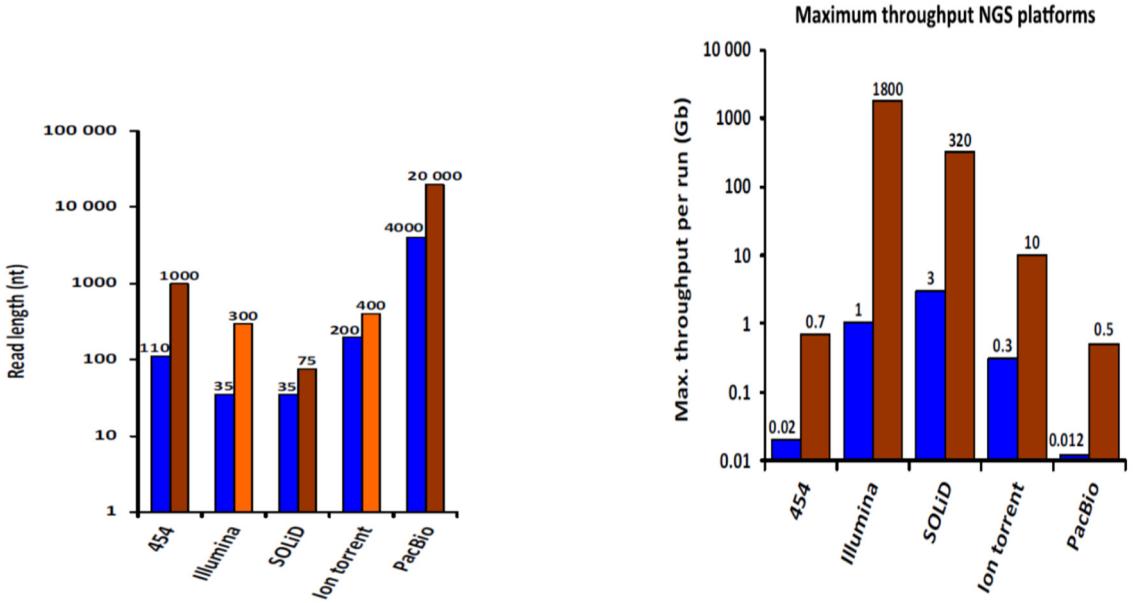

Figure 2.4: Over the years read length has improved with new developments to increase the throughput of NGS technologies. **Left -** blue bars indicate maximum read length of the first NGS instruments. Orange/dark orange indicate current large and bench-top NGS instruments. **Right –** blue bars indicate maximum throughput of the first commercially available NGS instruments, while dark orange bars indicate current maximum throughput. Adapted from van Dijk et al., 2014.

However, there are several challenges facing these NGS methods, including the expensive NGS experiments, data storage challenges and analysis required for NGS datasets, as well as the rapidly evolving technologies (Wall et al. 2009). Each NGS method is discussed in detail below.

(i) 454 sequencing

The GS FLX sequencing system (Margulies et al. 2005), originally developed by 454 Life Sciences and later acquired by Roche (Basel, Switzerland), was the first commercially available high-throughput sequencing platform (Margulies et al. 2005). In October 2013, Roche announced that it would shut down 454, and stop supporting the platform by mid-2016 (https://www.genomeweb.com).This technology avoids the cloning step and replace it with emulsion PCR for a highly efficient *in vitro* DNA amplification (Tawfik and Griffiths, 1998). Emulsion PCR makes use of DNA fragments carrying streptavidin beads in emulsion droplets. These droplets are used as for amplification reaction that produce thousands of a unique DNA template. Each bead containing DNA template and its clonally related templates are analyzed by pyrosequencing reaction. The sequencing throughput of this technology is enhanced by use of picotiter plate that allows for thousands of pyrosequencing reactions to be carried out in parallel (Margulies et al. 2005).

The pyrosequencing approach (Ronaghi et al. 1996) is a sequencing-by-synthesis technique. This technique makes use of chemiluminescence to measure the release of inorganic pyrophosphate (PPi). In pyrosequencing reaction, the template DNA is immobilized and dNTPs solutions are added one at a time. That releases PPi upon incorporation of complementary nucleotide. That is then detected by light production as a result of a chemiluminescence enzyme in the reaction mix. The DNA template sequence is determined by the order of correct nucleotides that had been incorporated (**Figure 2.5**). The current 454 platform is can generate up to 80-120 Mb of sequence in 200 to 300 bp reads in a 4-hour run (Roche Applied Sciences).

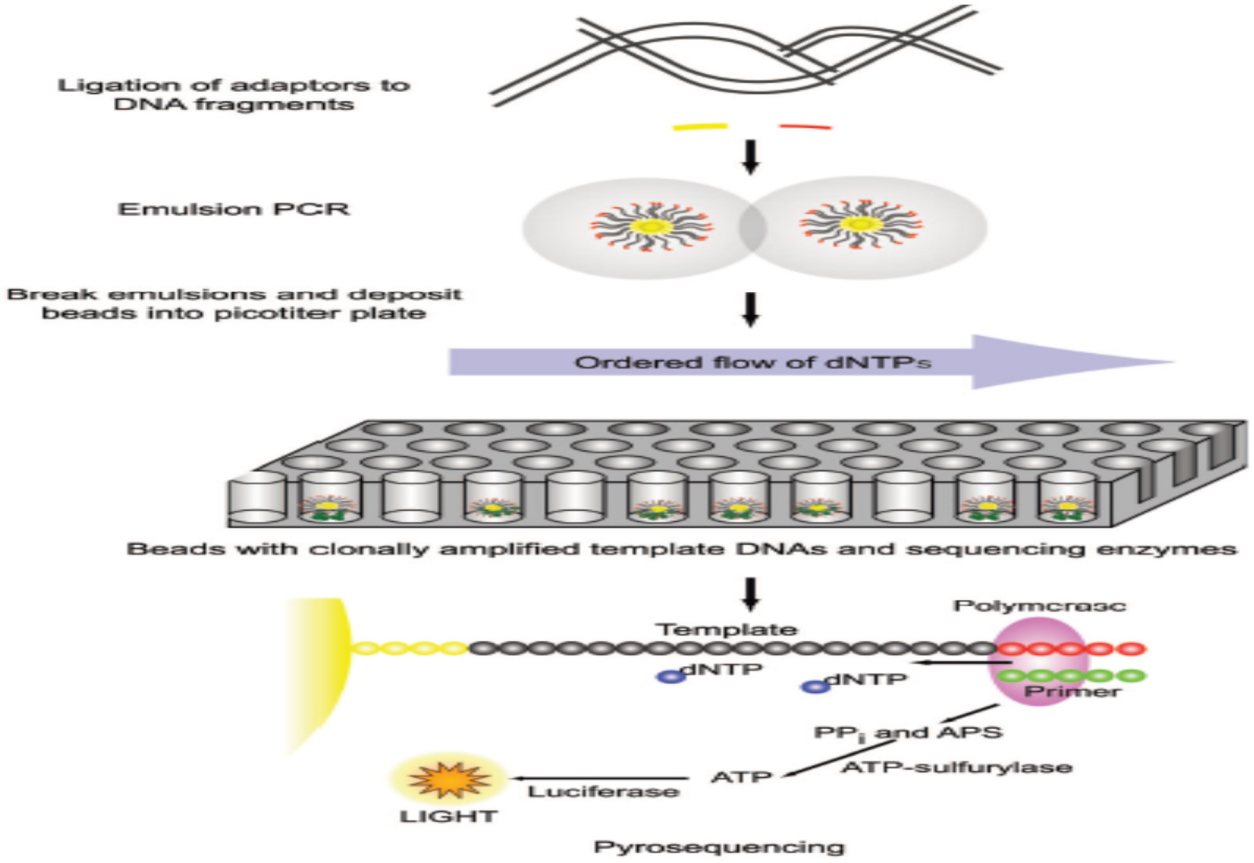

Figure 2.5: Roche 454 GS FLX sequencing. Template DNA is fragmented, end-paired, ligated to adapters and clonally amplified by emulsion PCR. (Adapted from Voelkerding et al., 2009).

(ii) ABI/SOLiD sequencing

The SOLiD sequencing system by Applied Biosystems (Foster City, CA) applies fluorescent labelled oligonucleotide panel and ligation chemistry for sequencing (Smith et al. 2010; Valouev et al. 2008). SOLiD, Applied Biosytems – This sequencer is featured in Valouev et al. (2008) and Smith et al. (2010), based on the Polonator technology (Shendure et al. 2005), an open source sequencer that utilizes emulsion PCR to immobilize the DNA library onto a solid support and cyclic sequencing by ligation chemistry (Myllykangas et al. 2012). This technology employs a massively parallel sequencing by hybridization-ligation approach. The ligation chemistry of this method is based on the polony sequencing technique (Shendure et al. 2005). Like the 454 technique, SOLiD employs the emulsion PCR single-molecule amplification. The molecules that are produced from the amplification reaction are sequenced on a glass surface by 16 dye-labeled dinucleotide combinations. The dinucleotides are labeled with four different fluorescent dyes (each dye is used to label four dinucleotides). Each position on the oligonucleotide chain is effectively probed twice. The nucleotide identity is determined by analyzing the colour that results from two successive ligation reactions (**Figure 2.6**). The SOLiD instrument can produce 1-3 Gb of sequence data in 35-bps reads per run in an 8-day run (Morozova and Marra, 2008).

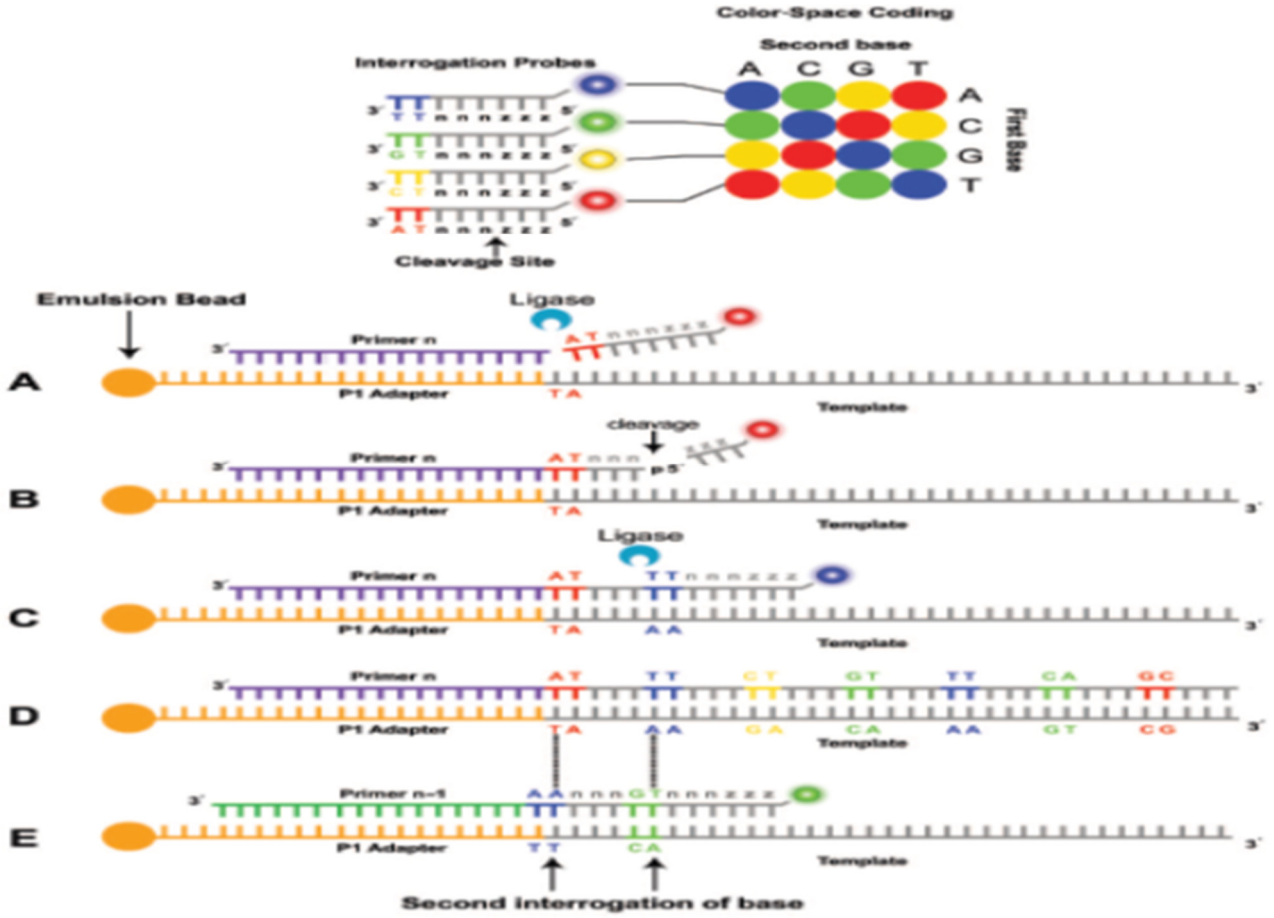

Figure 2.6: Applied Biosystems SOLiD sequencing by ligation. (Adapted from Voelkerding et al., 2009).

The two NGS technologies discussed above have their own unique advantages as well as disadvantages/limitations. The latter can pose challenges for the *de novo* sequencing of the transcriptome, which is the ultimate goal of this study. Among other things, both the 454 pyrosequencing and SOLiD sequencing have high sequencing cost and low raw sequence output. In 2014, Brewer et al (2014) reported that 454 GS FLX produced 0.7 Gb raw sequence output and cost $ 10/Mb, and SOLiD v4 produced 120 Gb raw sequence output and cost $ 0,13/Mb. On the other hand, Illumina HiSeq 2000 produced 600 Gb raw sequence output and cost only between $ 0,05 – 0,15/Mb. Given all these shortcomings, Illumina sequencing was the technology of choice due to its well-known advantages it provides. Its cheapest sequencing and high-throughput made it the option. Moreover, ARC-OVR has an in-house sequencing facility located at the Biotechnology Platform. That would mean even less sequencing expenses, as there would be no transportation required. Considering all those factors, Illumina was the best sequencing choice for our study.

(iii) Illumina sequencing

The first short read sequencing technology, Genome Analyzer, was developed by Solexa, which was later acquired by Illumina Inc. (San Diego, CA) (Bentley et al. 2006; 2008). The Illumina Genome Analyzer is the system that is based on immobilizing linear sequencing library fragments using solid support amplification. DNA sequencing is enabled using fluorescent reversible terminator nucleotides (Myllykangas et al. 2012).

The Illumina/Solexa approach uses solid surface such as single-molecule array or flow cell and adapters to attach one end of single-stranded DNA molecules. Subsequently, the molecules form a bridge by bending over and hybridizing to complementary adapters. Consequently, that forms the template for the synthesis of their complementary strands. About 1000s clonal copies of each single template molecule are produced after amplification. The templates are sequenced by DNA sequencing-by-synthesis method. The colour-labeled terminators are incorporated into growing oligonucleotide chains. The terminators are labeled with four different colours to distinguish among the different bases. Then, the template sequence of each cluster is deduced by reading off the colour of each successive nucleotide addition step. While the Illumina approach is more effective than pyrosequencing, it produces shorter reads

(Bentley, 2006) and substitution errors have also been noted (Hutchison, 2007). A typical Illumina's HiSeq 2500 machine is capable of producing 300 million 250 base pairs paired reads and 1 terabyte per run using Rapid run mode v2 500 cycle reagents at approximately one-third the cost per base compared to MiSeq (de Muinck et al. 2017) (**Figure 2.7**).

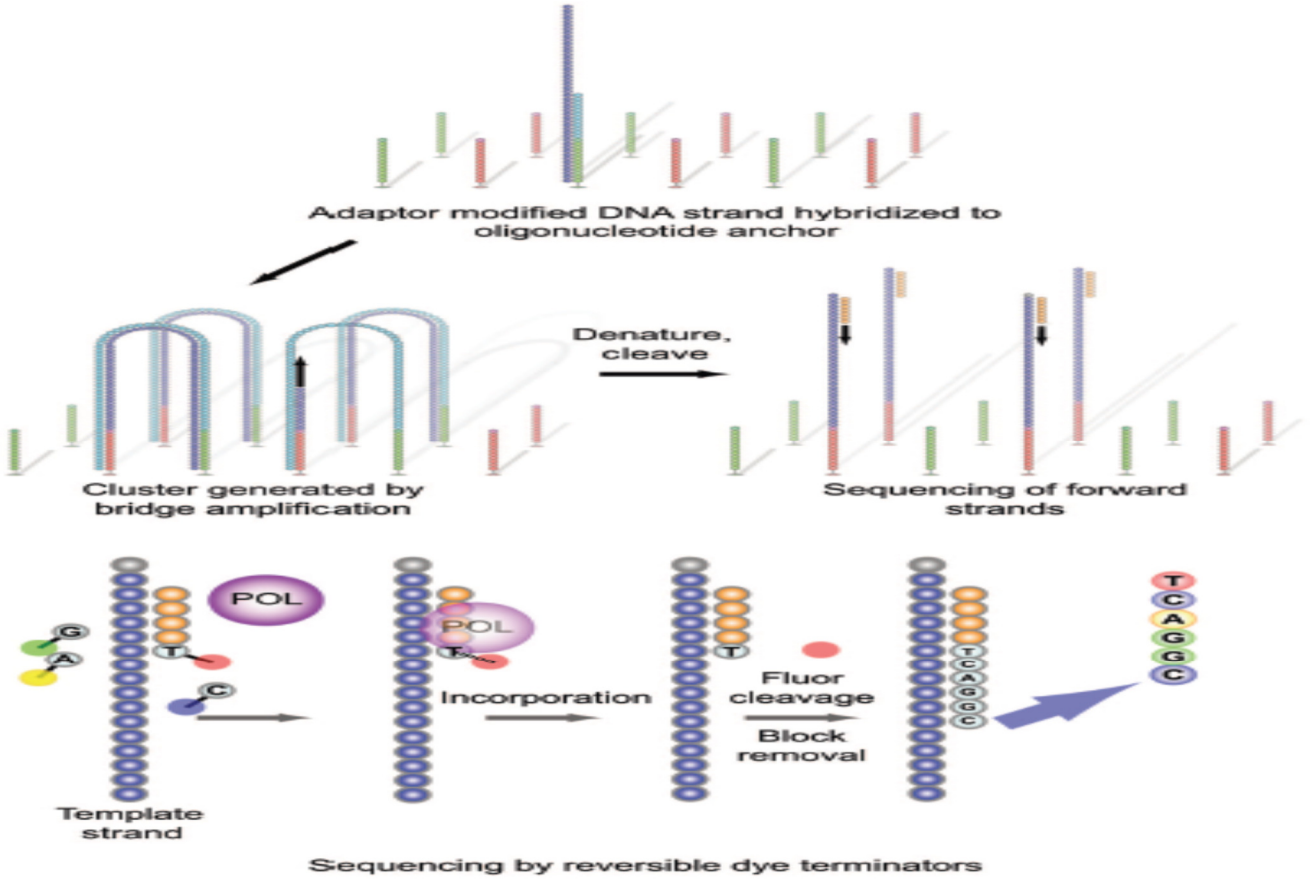

Figure 2.7: Illumina Genome Analyzer sequencing. (Adapted from Voelkerding et al., 2009).

2.1.2. Proteomics approach on tick antigens identification

The word ***proteome***, was pioneered by Wilkins in 1994 (Wilkins et al. 1996). It refers to the complete set of proteins (complementary to the genome) produced by a cell, a tissue or an organism, varying with time and under given physiological conditions (Kellner, 2000; Deracinois et al. 2013). The proteome also contain information about modifications made on proteins (Kellner, 2000). The analysis of the proteome is called proteomics, which is a large-scale study of proteins expressed in an organism, tissue or cells (Sun et al. 2014). Proteomics attempts to describe the molecular basis of physiological processes that are due to protein expression changes over time in response to various stimuli or to change of conditions (Sun et al. 2014). It supplements gene sequence data with protein information about where and in which ratio and under what conditions proteins are expressed (Kellner, 2000).

Generally, a proteomic approach comprises of four experimental steps: The first step is dedicated to the cell or sample conditioning (cell growth conditions, cell collection, cell storage, cell disruption). The second step corresponds to the sample preparation (extraction, concentration, purification to remove contaminants such as lipids or nucleic acids, storage of proteins), while the third step is related to methods of separation and the fourth is related to the quantification and identification of proteins (Bodzon-Kulakowska et al. 2007). The fifth step would be protein classification as per their biological functions and families they belong to (**Figure 2.8**).

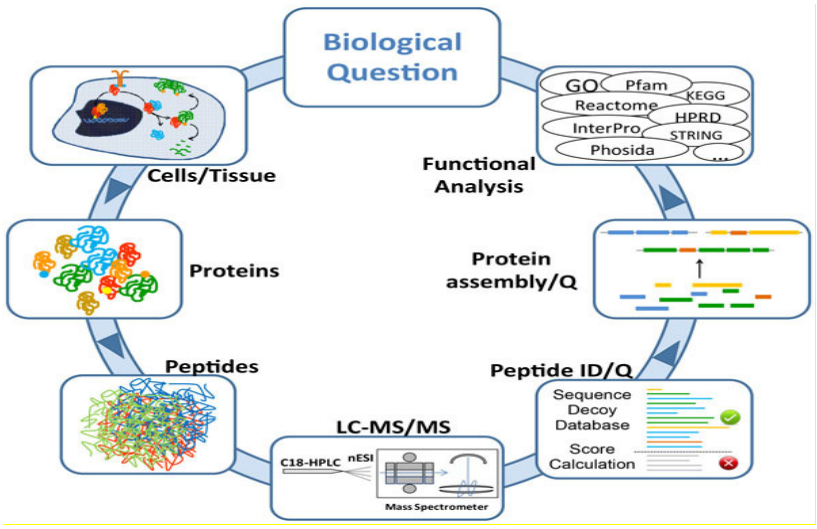

Figure 2.8: A typical proteomic approach workflow**. This approach involved the sampling step,** protein extraction and digestion, peptides separation, peptides detection, protein assembly and functional analysis. (Adapted from: Schmidt et al. 2014).

Proteomics has emerged as a high-throughput method for identification and analysis of proteins (Sun et al. 2014). Proteomics is a technique that can be applied broadly to various studies that seek to provide answers to different scientific questions or hypothesis. It has been used to profile proteomes of various organisms such as bacteria, animals, insects, plants and cell lines to mention a few (Molloy et al. 2000; Park, 2004; Shao et al. 2010; de Graaf et al. 2010; Lundberg et al. 2010). Proteomics has, over the years, undergone great advances in the research field (Domon and Aebersold, 2006). On the other hand, proteomics has also become a field of research facilitated by numerous advancements of protein separation, mass spectrometry, genome sequencing/annotation and protein search algorithms (Thelen, 2007).

Moreover, proteomics is regarded as one of the rapid growing areas in research, mainly because of the demand for global-scale analysis of proteins. Such analysis is expected to surpass analysis of genes as it yields more direct understanding of function and regulation (Bertone and Snyder, 2005). Proteomics coupled with transcriptomics as well as next generation sequencing (NGS) holds a promise to provide a high-throughput strategy to facilitate the analysis of proteins produced by a number of ongoing genome sequencing projects seeking to fully sequence various organisms' genomes.

2.2. Transcriptomics

2.2.1. Tick feeding, collection and ethics clearance

The ethics clearance certificate (**clearance number: 2013/CAES/000**) for this study was applied for and obtained from Unisa after approval by Unisa ethics committee. *Rhipicephalus (B.) microplus* female ticks used for the purpose of this study were fed on cattle at the Agricultural Research Council-Onderstepoort Veterinary Research animal facilities. They were collected at five different feeding stages, that is small (BM1); medium (BM2) (Day 2); medium (BM3) (Day 3); large (BM4) (Day 4) and when they dropped off (BM5) (Day 5).

2.2.2. Tick salivary glands isolation/preparation of salivary gland extract (SGE)

Female ticks were embedded in a wax-filled petri dish with their dorsal parts visible. The ticks' dorsal surface were separated and removed by viewing under a dissecting microscope and cutting the lateral edge of the cuticle with sterile scalpel blade. Intact salivary glands were carefully removed with fine-tipped forceps and were placed in 500 µl of ice-cold phosphate-buffered solution (PBS) (20 mM Tris-HCl, 0.15M NaCl, pH 7.6 at 25 °C). Salivary glands were subsequently immersed in 500 µl RNAlater® (Ambion Inc., Austin, USA) and incubated at 10°C for 24 hours. After 24 hours, excess RNAlater was removed and glands frozen at -70°C until use.

2.2.3. RNA isolation from tick salivary glands

Total RNA was isolated using Qiagen RNeasy® Mini Kit (Qiagen, CA, USA). There were five 2 ml Eppendorf tubes each with about 200 mg of the glands for each feeding stage. A volume of 600 µl of RLT-*Dithiothreitol* (RLT-DTT) buffer was added to each tube on ice. Each sample was homogenized with a 25 gauge syringe and needle by pushing it up and down through the syringe ten times (10X). The previous step was repeated using a 25-gauge syringe and needle, and/or until the lysate was clear. Then, the lysate was centrifuged at 14 000 x *g* for 3 minutes at room temperature (25 °C). Supernatant was transferred into a new 1.5 ml Eppendorf tube and 1 ml of 70% ethanol was added, and mixed by gently pipetting five times. Seven hundred microliters of the suspension was transferred to an RNeasy spin column/2ml collection tube and centrifuged at 14000 x *g* for 15 seconds. The previous step was repeated with the leftover suspension (~700 µl). This was followed by centrifuging the empty spin

column/2ml collection tube at 14000 x *g* for 15 seconds. Seven hundred microliters of RW1 wash buffer was added to the RNeasy spin column and centrifuged at 14000 x *g* for 15 seconds, before discarding the flow-through. A volume of 500 µl of RPE wash buffer was added straight onto the column and centrifuged at 14000 x *g* for 15 seconds, before discarding the flow-through. The empty tube was centrifuged for 2 minutes at 14000 x g, followed by another dry spin at 14000 x *g* for 15 seconds in a new collection tube. The spin column tube (pink) was transferred to a clean 1.5 ml Eppendorf tube and 40 µl of RNAse-free water was added to the spin column. It was then incubated for 1 minute and centrifuged at 14 000 x *g* for 1 minute to elute the total RNA. All procedures were performed at 25 °C. The RNA was then stored at - 70°C until further use in downstream processes.

2.2.4. Determination of RNA quantification by NanoDrop

The NanoDrop ND2000 (Thermo Scientific, Wilmington, USA) was used for quantification of total RNA at 260 nm. The purity of the total RNA was determined by the ratio of 260/280 nm. The upper optical pedestal and lower pedestal were cleaned by placing 4 µl of RNAse-free water. Then, they were wiped off thoroughly with a clean paper towel. Undiluted RNA (4 µl) was placed onto the lower measurement pedestal and the upper pedestal was lowered, before the measurement was taken. Upper and lower pedestals was cleaned after each reading, by wiping with a clean paper towel. Each RNA sample was quantified in this manner.

2.2.5. Next-generation sequencing: salivary gland cDNAs for transcriptome analysis

(a) Messenger RNA purification and fragmentation

This procedure was performed using Illumina® TruSeq® RNA Sample Preparation kit v2 Guide (Low Sample) Protocol (Illumina Inc., CA, USA). One and half microgram (1.5 µg) of total RNA was used for each sample. A volume that is equivalent to 1.5 µg was calculated for each sample, and below are the volumes. Prior to perform the reactions, each sample was diluted out with nuclease-free water to give a final working volume of 50 µl.

Sample	Mass (1.5 µg) equivalence in volume (µl)	Nuclease-free water vol. (µl)
BM 1	5 µl,	45
BM 2	3 µl,	47
BM 3	3 µl,	47
BM 4	5 µl	45
BM 5	3 µl	47

A volume of 50 µl of RNA Beads Purification (RBP) was added to each tubes (labeled "RBP") containing total RNA, and mixed gently by pipetting up and down six times. Following the mixing, tubes were placed and heated on a heating block at 65°C for 5 minutes. Samples were then incubated on ice for a minute and at room temperature for 5 minutes. Tubes were placed on a magnetic stand and incubated at room temperature for 5 minutes. After incubation, all of the supernatant was removed carefully from all the tubes and discarded without disturbing the pellets. Tubes were then removed from the magnetic stand and pellets were washed with 200 µl of Bead Washing Buffer (BWB) and pipetting up and down six times. Samples were returned to the magnetic stand and incubated at room temperature for 5 minutes, after which all of the supernatant was removed carefully without disturbing the pellet, and discarded. With tubes off the magnetic stand, 50 µl was added to each tube and mixed by gently pipetting up and down six times. Samples were heated at 80°C on the heating block for 2 minutes. The heating step was followed by incubation on ice for a minute and at room temperature for 30 seconds. Following incubation, 50 µl of Bead Binding Buffer (BBB) was added to each sample and mixed by gently pipetting up and down six times. Following that was incubation at room temperature for 5 minutes and on the magnetic stand (room temperature) for another 5 minutes. Thereafter, all of the supernatant was removed carefully from all the tubes without disturbing the pellets.

Samples were removed from the magnetic stand. Pellets were washed by adding 195 µl of BWB each sample and mixed by gently pipetting up and down six times. Thereafter, samples were incubated on the magnetic stand (at room temperature) for 5 minutes. All of the supernatant was removed from each sample without disturbing the pellet before removing them from the magnetic stand. A volume of 19.5 µl of Elute,

Prime, Fragment Mix (EPF) was then added to each sample and mixed by gently pipetting up and down six times. After mixing, samples were transferred into new 0.5 ml PCR tubes (labeled “RBP”). PCR tubes were then put in the PCR machine and fragmentation was performed by 1 cycle at 94°C for 3 minutes, and hold at 4°C. Fragmentation step was immediately followed by first strand cDNA synthesis.

(b) First strand cDNA synthesis

The polymerase chain reaction (PCR) was used for the synthesis of first strand of cDNA. PCR tubes from the previous step were placed on a magnetic bar (fitted with 1.5 ml tube-adapters) at room temperature for 5 minutes. Without disturbing the pellets 17 µl of the supernatant was removed from each tube, and transferred into new 0.5 ml PCR tubes. To each sample 8 µl of First Strand Master Mix-STR (FSM-STR) was added and mixed by pipetting up and down six times. After mixing tubes were put in PCR machine and the synthesis was performed according to the following conditions:

- **25°C for 10 minutes**
- **42°C for 50 minutes**
- **70°C for 15 minutes**
- **4°C on hold**

When the reaction reached 4°C, tubes were taken out and proceeded to second strand synthesis.

(c) Second strand cDNA synthesis

To each PCR tubes (labeled “CDP”) 25 µl of Second Strand Master Mix (SSM) was added and mixed by pipetting up and down six times. After mixing, tubes were put in PCR machine and the synthesis was performed according to the following conditions: 16°C for 1hour and 25°C on hold. After the PCR step sample reactions were transferred into new 1.5 ml Eppendorf tubes (labeled “CDP”), to which 90 µl of AMPure XP Beads was added and mixed by pipetting up and down ten times. Following that was incubation at room temperature for 15 minutes and on the magnetic bar (at room temperature) for 5 minutes. With tubes still on the magnetic bar, 135 µl of supernatant was removed from each tube without disturbing the pellets. Pellets were then washed with 200 µl of freshly prepared 80% ethanol and incubation at room temperature for 30 seconds. After the washing step all of the supernatant was removed without disturbing the pellets.

The second wash was performed with 190 µl of freshly prepared 80% ethanol and incubation at room temperature for 30 seconds. After the second washing step, all of the supernatant was removed without disturbing the pellets. Tubes were left on the magnetic bar for 15 minutes or until pellets were completely dry (no ethanol residues on the inside of the tubes). Pellets were resuspended in 52.5 µl of RSB by gently pipetting up and down ten times. Sample suspensions were then incubated at room temperature for 2 minutes and on the magnetic bar (at room temperature) for 5 minutes. After incubation 50 µl of the supernatant from each sample was transferred into new 1.5 ml Eppendorf tubes (labeled “IMP”). Samples were stored at - 20°C until used within seven days.

(d) Perform End Repair

Frozen “IMP” tubes were thawed on ice and briefly vortexed prior to end repair reaction. Ten microliters of RSB and 40 µl were into each sample and mixed thoroughly by pipetting up and down ten times. Samples were then transferred into new PCR tubes and PCR procedure was performed according to the following conditions: 30°C for 30 minutes, and hold at 25°C. PCR reaction mixes were then transferred into new 1.5 ml Eppendorf tubes. Samples were cleaned up by adding 160 µl of AMPure XP beads and gently pipetting up and down ten times. Mixing was followed with incubation at room temperature for 15 minutes and on the magnetic bar (at room temperature) for 5 minutes. The volume of 127.5 µl of supernatant was removed from each sample and discarded without disturbing the pellet.

Pellets were then washed with 200 µl of freshly prepared 80% ethanol and incubation at room temperature for 30 seconds. Thereafter, all of the supernatant was removed and discarded from each sample without disturbing the pellets. The ethanol washing step was repeated as above and the pellets were dried at room temperature (with tube caps opened), or until pellets were completely dry. Sample tubes were removed from the magnetic bar and pellets were resuspended in 17.5 µl of RSB and gently pipetting up and down ten times. Sample suspensions were then incubated at room temperature for 2 minutes and on the magnetic bar (at room temperature) for 5 minutes. After incubation 15 µl of supernatant was removed from each sample and transferred into new PCR tubes (labeled “ALP”). Samples were stored at - 20°C until used within seven days.

(e) 3' Ends Adenylation

A volume of 2.5 µl of RSB and 12.5 µl of A-Tailing Mix (ATL) were added into each PCR tubes (labeled "ALP") and mixed by gently pipetting up and down ten times. All PCR tubes (labeled "ALP") were put in PCR machine and the reaction was run with following conditions: 37°C for 30 minutes and 25°C on hold**.** This step was immediately followed by adapters ligation step.

(f) Adapters Ligation

RNA Adapter Indices was assigned to each sample. Indices assignment were done as follows: **BM1 – 6, BM2 – 13, BM3 – 14, BM4 – 5 and BM5 – 19.** Volumes of 2.5 µl of RSB and 2.5 µl of Ligation Mix (LIG) were added to each PCR tube. 2.5 µl of each RNA Adapter Indices was added to its tube and mixed by gently pipetting up and down ten times. Samples were then put in the PCR machine and the reaction was run with following conditions: 30°C for 10 minutes and 25°C on hold. The ligation reaction was stopped by adding 5 µl of Stop Ligation Buffer (STL) into each sample and mixed by gently pipetting up and down ten times. Reaction mixes were then transferred into new 1.5 ml Eppendorf tubes (labelled "ALP"). The ligation step was followed by the clean-up step. About 42 µl of AMPure XP Beads was added into each tube and mixed by gently pipetting up and down ten times. Mixing was followed with incubation at room temperature for 15 minutes and on the magnetic bar (at room temperature) for 5 minutes and 79.5 µl of supernatant was removed from each sample and discarded without disturbing the pellet. Tubes were left on the magnetic bar for washing step.

Pellets were then washed by adding 200 µl of freshly prepared 80% ethanol to the pellets and incubation at room temperature for 30 seconds. All of the supernatant was removed and discarded. The ethanol washing step was repeated with 190 µl and pellets were dried at room temperature (with tube caps opened), or until pellets were completely dry. Pellets were resuspended in 52.5 µl of RSB and mixed by gently pipetting up and down ten times. That was followed by incubation at 25°C for 2 minutes and on the magnetic bar (at room temperature) for 5 minutes. After incubation 50 µl of the suspension for all samples was transferred into new 1.5 ml Eppendorf tubes (labelled "CAP").

A volume of 50 µl of AMPure XP Beads was added into each tube (labelled “CAP”) and mixed by gently pipetting up and down ten times. Mixing was followed with incubation at room temperature for 15 minutes and on the magnetic bar (at room temperature) for 5 minutes and 95 µl of supernatant was removed from each sample and discarded without disturbing the pellet. Tubes were left on the magnetic bar for washing step. Pellets were washed by adding 200 µl of freshly prepared 80% ethanol to the pellets and incubation at room temperature for 30 seconds. All of the supernatant was removed and discarded.

The ethanol washing step was repeated with 190 µl and pellets were dried at room temperature (with tube caps opened) for 15 minutes, or until pellets were completely dry. With tubes off the magnetic bar, pellets were resuspended in 22.5 µl of RSB and mixed by gently pipetting up and down ten times. That was followed by incubation at 25°C for 2 minutes and on the magnetic bar (at room temperature) for 5 minutes. After incubation 20 µl of the suspension for all samples was transferred into new 0.2 ml PCR tubes (labelled “PCR”).

(g) DNA Fragments Enrichment

PCR mixes were prepared by adding 5 µl of PCR Primer Cocktail (PPC) and 25 µl of PCR Master Mix (PMM) into each PCR tube and mixed by gently pipetting up and down ten times. After mixing tubes were put in PCR machine and the synthesis was performed according to the following conditions:

- **98°C for 30 seconds**
- **98°C for 10 seconds**
- **60°C for 30 seconds**
- **72°C for 30 seconds**

15 cycles

- **72°C for 5 minutes**
- **4°C on hold**

PCR products were transferred into new 1.5 ml Eppendorf tubes (labelled “PCR”) for clean-up step. A volume of 50 µl of AMPure XP Beads was added into each tube and mixed by gently pipetting up and down ten times. Mixing was followed with incubation at room temperature for 15 minutes and on the magnetic bar (at room temperature) for 5 minutes and 95 µl of supernatant was removed from each sample and discarded without disturbing the pellet. Tubes were left on the magnetic bar for washing step.

Pellets were washed by adding 200 µl of freshly prepared 80% ethanol to the pellets and incubation at room temperature for 30 seconds. All of the supernatant was removed and discarded. The ethanol washing step was repeated with 190 µl and pellets were dried at room temperature (with tube caps opened) for 15 minutes, or until pellets were completely dry. With tubes off the magnetic bar, pellets were resuspended in 32.5 µl of RSB and mixed by gently pipetting up and down ten times. That was followed by incubation at 25°C for 2 minutes and on the magnetic bar (at room temperature) for 5 minutes. After incubation 20 µl of the suspension for all samples was transferred into new 1.5 ml tubes (labelled “TSP1”). Those were the finally prepared cDNA libraries.

2.2.6. Complementary DNA (cDNA) Library Validation

The cDNA libraries were validated by Qubit® HS quantification and agarose gel electrophoresis analysis. Qubit® dsDNA HS Assay kit (Life Technologies™, CA, USA) was used to prepare the assay and its standards. The concentration of each cDNA sample was measure with Qubit® 2.0 Fluorometer (Life Technologies™, CA, USA). Samples were then resolved in 1% agarose gel electrophoresis to check their integrity.

(a) Qubit® HS (High Sensitivity) Quantification Assay

All reagents were be kept at room temperature prior to use. Qubit® HS (High Sensitivity) working solution was prepared by mixing 1 µl of Qubit® HS reagent with 199 µl of Qubit® HS buffer and mixed by vortexing for 2–3 seconds. Standards were prepared by adding 10 µl of each standard into 190 µl Qubit® HS working solution and mixed by vortexing for 2–3 seconds. Each standard had a total volume of 200 µl. Samples were set up by mixing 2 µl of each sample with 198 µl of Qubit® HS working solution and mixed by vortexing for 2–3 seconds. All tubes were incubated at room temperature for 2 minutes. The tubes were inserted in the Qubit® 2.0 Fluorometer and readings were taken. The stock concentrations of original samples were determined by using the dilution calculator feature of the Qubit® 2.0 Fluorometer.

(b) Agarose Gel Electrophoresis Analysis

Assessment of the purity of cDNA libraries was carried out with agarose gel electrophoresis. A 1% agarose gel was prepared by mixing 100 g of agarose powder with 100 ml of 1X TEA buffer. The mixture was boiled in the microwave and cooled down under running tap water with gently swirling. Five µl of ethidium bromide was added into the mixture and mixed with gently swirling. The mixture was then poured in the casting tray and welled-combs were placed on the top of the casting tray. The gel was left to solidify before it was placed in a tank filled with 1X TAE buffer. Twenty microlitre of cDNA library for each sample was mixed with 4 µl of 6X DNA loading dye. The mixtures were then loaded into 1% agarose gel wells along with the DNA ladder. Electric voltage of 100 volts was applied for 20 minutes. The gel bands were visualised using UV transilluminator (Biorad Laboratories Inc., CA, USA) and gel photos were taken. Gel bands of each sample were cut out of the gel into two slices with sharp scalpel. The slices were of high molecular weight (~600bp and above) and low molecular weight (~300bp and below). The gel slices were put into two separately 2 ml Eppendorf tubes, labelled as HMW (High Molecular Weight) and LMW (Low Molecular Weight). Tubes were kept at -20°C for gel clean up procedure.

2.2.7. Gel Clean Up

The gel clean-up was carried out with Qiagen MinElute Gel Extraction Kit (Qiagen, CA, USA). Only gel slices of high molecular weight (~600 bp and above) were used for clean-up. Gel slices were weighed out to be 450 mg or less. If needed gel slices were cut and divided to have a mass of 450 mg or less.

Three volumes of Buffer QG was added to 1 volume of gel and incubated at room temperature. During incubation, samples were vortexed for 2 – 3 minutes and gently inverted several times until gels were completely dissolved. One volume of isopropanol was added to each gel sample and mixed by inverting the several times. MinElute columns with 2 ml collection tubes were placed in a rack and 700 µl of each gel sample was transferred into MinElute columns - collection tubes. Following that was centrifugation at 11 000 x *g* at room temperature for a minute. The flow-though was discarded and MinElute columns were placed back to the same collection tubes **(this step was repeated for samples with a total volume more than 700 µl)**. A volume of 500 µl of Buffer QG was added into each gel sample and centrifuged at 11 000 x *g* at room temperature for a minute. Bounded cDNA was washed with 750 µl of Buffer

PE and centrifugation at 11 000 x *g* at room temperature for a minute. The flow-though was discarded and MinElute columns were centrifuged for an additional minute. MinElute columns were placed into clean 1.5 ml tubes and cDNA was eluted by adding 16 µl of Buffer EB and incubation at room temperature for 5 minutes. Incubation was followed by centrifugation at 11 000 x *g* at room temperature for a minute. The eluted volume of cDNA was ±30 µl.

2.2.8. Validation of purified cDNA library

Purified double stranded cDNA libraries cDNA libraries were validated by Qubit® HS quantification and agarose gel electrophoresis analysis. Qubit® dsDNA HS Assay kit (Invitrogen™, Life Technologies™) was used to prepare the assay and its standards. The concentration of each cDNA sample was measured with Qubit® 2.0 Fluorometer (Life Technologies™, CA, USA). Samples were then resolved in 1% agarose gel electrophoresis to check their integrity.

2.2.9. Qubit® HS Quantification Assay

Qubit® HS working solution was prepared by mixing 1 µl of Qubit® HS reagent with 199 µl of Qubit® HS buffer and mixed by vortexing for 2–3 seconds. Standards were prepared by adding 10 µl of each standard into 190 µl Qubit® HS working solution and mixed by vortexing for 2–3 seconds. Each standard had a total volume of 200 µl. Samples were set up by mixing 2 µl of each sample with 198 µl of Qubit® HS working solution and mixed by vortexing for 2–3 seconds. All tubes were incubated at room temperature for 2 minutes. The tubes were inserted in the Qubit® 2.0 Fluorometer and readings were taken. The stock concentrations of original samples were determined by using the dilution calculator feature of the Qubit® 2.0 Fluorometer.

2.2.10. Agarose Gel Electrophoresis analysis

Assessment of the purity of cDNA libraries was carried out with agarose gel electrophoresis. The 1% agarose gel was prepared by mixing 100 g of agarose powder with 100 ml of 1X TEA buffer. The mixture was boiled in the microwave and cooled down under running tap water with gently swirling. Five µl of ethidium bromide was added into the mixture and mixed with gently swirling. The mixture was then poured in the casting tray and welled-combs were placed on the top of the casting tray. The gel was left to solidify before it was placed in a tank filled with 1X TAE buffer. Twenty µl of cDNA library for each sample was mixed with 4 µl of 6X DNA loading dye. The mixtures

were then loaded into 1% agarose gel wells along with the DNA ladder. Electric voltage of 100 V was applied for 20 minutes or bands had reached the bottom of the gel. The gel bands were visualised using UV transilluminator (Biorad Laboratories Inc., Hercules, CA, USA). The remaining sample (±10 µl) of each purified cDNA library was used for the transcriptome sequencing procedure.

2.2.11. Transcriptome sequencing

(a) Sequencing with Illumina MiSeq sequencing

A volume of (±10 µl) of each purified cDNA library was used for the sequencing procedure. The sequencing was done by the Agricultural Research Council-Biotechnology Platform facility, using the Illumina® Miseq® sequencing machine. Samples were multiplexed together using different Illumina® specific indexes and run as 250 bp by 250 bp runs.

2.2.12. *De novo* assembly of transcriptome data

(a) Data filtering

A bioinformatics pipeline was used to quality trim and assemble the data (Professor Ben Mans). Reads were imported as single unpaired reads into CLC Genomics Workbench (6.0.5) (Qiagen), quality trimmed with a score of 0.001 (Phred score = 35) and no ambiguous nucleotides. TruSeq adapters were removed by default parameters **Table 2.2**).

Table 2.2: Parameters for adapter trimming

Name	Sequence	Strand	Alignment score	Action
TruSeq Universal Adapter	AATGATACGGCGACCACCGAGAT CTACACTCTTTCCCTACACGACGC TCTTCCGATCT	Plus	Mismatch: 2, Gapcost: 3, Cutoff: 15, Cutoff at end: 2	Remove adapter
TruSeq Universal Adapter reverse	AGATCGGAAGAGCGTCGTGTAG GGAAAGAGTGTAGATCTCGGTG GTCGCCGTATCATT	Minus	Mismatch: 2, Gapcost: 3, Cutoff: 15, Cutoff at end: 2	Remove adapter
TruSeq Adapter end	GATCGGAAGAGCACACGTCTGA ACTCCAGTCAC	Plus	Mismatch: 2, Gapcost: 3, Cutoff: 15, Cutoff at end: 2	Remove adapter

TruSeq adapter reverse	GTGACTGGAGTTCAGACGTGT GCTCTTCCGATC	Minus	Mismatch: 2, Gapcost: 3, Cutoff: 15, Cutoff at end: 2	Remove adapter
TruSeq Adapter end short	GATCGGAAGAGC	Plus	Mismatch: 2, Gapcost: 3, Cutoff: 10, Cutoff at end: 4	Remove adapter
TruSeq adapter reverse short	GCTCTTCCGATC	Minus	Mismatch: 2, Gapcost: 3, Cutoff: 10, Cutoff at end: 4	Remove adapter

This yielded a single read dataset (Single). Removal of duplicates yielded a single duplicate removed dataset (Sddup). Paired reads were imported and quality trimmed and adapters removed as described above. Trimmed paired reads were merged using default overlap detection parameters *(*Mismatch cost =2, Gap cost = 3, Maximum unaligned end mismatches = 0, Minimum score = 8*)* to give a merged dataset (Merged). Removal of duplicate reads from this dataset yielded a merged duplicate removed dataset (Mddup). Merged and Single datasets were combined to give a single-merged dataset (SM) and mddup/sddup datasets were combined to give a SMddup dataset. The following datasets were generated: All, BM0, BM1, BM2, BM3, BM4 and BM5, each of which had the Merged, Mddup, Single, Sddup, SM and SMddup subsets. This gave 42 datasets that was used to assemble the transcriptomes.

(b) Transcriptome *de novo* assembly and annotation

Transcriptome assembly was performed using CLC Genomics workbench (6.0.5) (Qiagen) or Trinity v2.4.0 (Grabherr et al. 2011). In CLC Genomics workbench, the following kmer sizes were used: 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 and 64. Assembly parameters were: mismatch cost-2, insertion cost-3, deletion cost-3, length fraction-0.9, similarity-0.9, minimum contig length-240, kmer size-variable, bubble size-automatic. Trinity was used with default parameter settings. The number of assemblies produced in this manner totalled 420 for CLC Genomics Workbench and 42 for Trinity.

(c) Creating a non-redundant open-reading frame (ORF) set and selection based on RPKM cut off

All contigs were exported in FASTA format and open reading frames (ORFs) greater than 240 bp were extracted from FASTA files using the Perl script “Orffinder” from Douglas Senalik (https://github.com/vikas0633/perl/blob/master/orffinder.pl). A non-redundant ORF dataset was created by combining ORFs from all assemblies and clustering at 95% identity level using CD-Hit-v4.6.4 (Li and Godzik, 2006). Chimeric

sequences were also removed using the same program, creating a non-redundant dataset which retained the longest ORFs. The combined datasets (All_Single) were mapped against the non-redundant ORF dataset using CLC Genomics Workbench. All ORFs with a RPKM ≥0.5 were retained for further analysis.

(d) Identification of correct open reading frames

ORFs were searched against an ACARI database using BLASTX (Altschul et al. 1990), to determine the correct reading frames. ORFs were extracted from these based on the best hit (E value ≤1E-4) produced for reading frames 1, 2, 3, 4, 5 and 6. ORFs that gave E values above the cut-off were classified as unknowns and removed from further analysis. ORFs from reading frames 2-6 were re-extracted using Orffinder. Final ORFs were translated to protein sequences using CLC Genomics Workbench for further analysis.

(e) Identification of secretory and house-keeping proteins

Translated protein sequences were analyzed against an in-house annotated ACARI protein database using BLASTP (Altschul et al.1997), with a cut-off = E≤1E-4. The ACARI protein database is an in-house curated database of protein sequences derived from GenBank or Vector-base (Mans et al. 2016). Where no protein sequences are available, nucleotide and EST data were retrieved from GenBank or Vectorbase and ORFs were extracted as described above and annotated as third party annotation (TPA). The ACARI database is curated based on KEGG (Kanehisa et al. 2016) annotation for housekeeping proteins, annotations from Acari genome projects and PSI-BLAST analysis (Altschul et al. 1997) using tick secretory proteins from literature to identify secretory protein families. This first pass analysis assigns proteins to housekeeping or secretory classes based on sequence homology. To confirm this, ORFs were submitted to the SignalP 4.1 Server (Petersen et al. 2011), TMHMM server v.2.0 (Krogh et al. 2001) and Phobius (Käll et al. 2004).

(f) Quality assessment of the transcriptome using BUSCO

Open reading frames were analysed using BUSCO (Simão et al. 2015), as implemented for transcriptome analysis using 1066 conserved arthropod genes.

2.3 Proteomics

2.3.1. Salivary gland extracts (SGE) preparation

Salivary glands were dissected in saline and stored frozen at -70°C. Salivary glands were suspended in 500 µl extraction buffer (20 mM Tris-HCl, pH 8) and sonicated using a Branson sonifier cell disruptor B-30 (Branson Sonic Power Co.) for 3 X 6 pulses at 30% duty cycles and an output control of 3 to prepare salivary gland extract. This was followed by centrifugation at 14000x*g* to pellet the membrane fraction (M fraction). The supernatant was removed and heated at 80°C for 5 minutes and precipitated proteins (P fraction) were pelleted by centrifugation at 14000x*g* (5 min). Supernant (S fraction) was removed. The M, P and S fractions were dissolved in 10 µl SDS-PAGE sample buffer.

2.3.2. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) analysis

NuPAGE® Novex 4-12% Bis-Tris precast polyacrylamide gels (Life Technologies, Carlsbad, CA, USA) were used to run the protein gel electrophoresis. All three samples i.e: M, S and P were fractionated by gel electrophoresis. Ten µl of each sample was mixed with 1 µl of 10X reducing agent (Dithiothreitol), and gently mixed by pipetting up and down. Samples were then heated at 70°C for 10 minutes. Samples were cooled to room temperature before loading. Gels were unpacked and assembled in gel tank with 1X SDS MES running buffer. Samples were ran along with pre-stained protein molecular weight marker (Life Technologies, Carlsbad, CA, USA) at 200 volts (V) and 125 amperes (A) for 10 minutes (±3 cm into the gel). Gels were shaken in fixing solution for 10 minutes and another 10 minutes in staining solution Stainer A (Coomassie Colloidal Blue staining kit). That was followed by addition of 5 ml of Stainer B (Coomassie Colloidal Blue staining kit) and shaken for 3 hours. After staining, gels were rinsed with 100 ml deionized water for a minute twice.

2.3.3. Gel processing

Using a glass plate and sharp scalpel, each sample lane was cut into 10 slices (±1X1 mm) based on the weight of the resolved bands. Each slice was cut into 10 pieces and transferred into a clean Eppendorf tube. To each tube 200 µl of 100% acetonitrile was added and followed by incubation at room temperature for 2 minutes twice. Then acetonitrile was removed and tubes were left open on the bench for 20 minutes to dry

the gel pieces. Gel pieces were kept in a cool dry place until further use. There were four feeding stages (BM1-BM4) and each feeding stage had three fractions. That resulted into a total amount of 120 samples.

2.3.4. Proteomics studies

A total of 160 samples (40 samples/protein fraction) were submitted to the *Centre for Proteomic and Genomic Research (CPGR)* Proteomics platform, Cape Town for analysis. Procedures were performed as per the CPGR protocols.

2.3.5. Proteome data analysis

(a) Database searching

A mass and charge state deconvolution was used to extract the tandem mass spectra. There was no deisotoping performed. Mascot (Matrix Science, London, UK; version 2.5.1) and X! Tandem (The GPM, thegpm.org; version CYCLONE (2010.12.01.1) were used to analyzed all MS/MS samples. Mascot was set up to search the Ixodidae_Bmic_Bdec20150613_20150720 database (134116 entries) assuming digestion by trypsin. X!Tandem was set up to search a subset of the Ixodidae_Bmic_Bdec20150613_20150720 database also assuming trypsin digestion. Mascot and X! Tandem were searched with a fragment ion mass tolerance of 0.020 Da and a parent ion tolerance of 20 PPM. Cysteine carbamidomethyation (CAM) was specified in Mascot and X! Tandem as a fixed modification. Deamidated of asparagine and glutamine and oxidation of methionine were specified in Mascot as variable modifications. Glu->pyro-Glu of the N-terminus, ammonia-loss of the N-terminus, gln->pyro-Glu of the N-terminus, deamidated of asparagine and glutamine and oxidation of methionine were specified in X! Tandem as variable modifications.

(b) Criteria for protein identification

The validation of MS/MS based peptide and protein identifications was performed by Scaffold (version Scaffold_4.5.0, Proteome Software Inc., Portland, OR). Peptide with greater than 98.0% probability and an FDR less than 0.1% using the Scaffold Local FDR algorithm were considered true positives. Proteins with greater than 99.0% probability and an FDR less than 1.0%, with at least 2 identified peptides were considered true identifications. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003). Proteins with similar peptides that could

not be differentiated using MS/MS analysis, were grouped as isoforms to satisfy the principles of parsimony.

Chapter 3: Transcriptome sequencing and analysis - data analysis and discussion

3.1. Introduction

The characterization of the transcriptome is a highthroughput way to discover tick salivary gland components. Previously the characterization of salivary gland transcriptomes/sialotranscriptomes of several tick species have been described by expressed sequence tags (ESTs) analysis obtained by classical Sanger sequencing of cDNA libraries (Ribeiro et al. 2006, 2011, 2012; Alarcon-Chaidez et al. 2007;Chmelar et al. 2008; Anatriello et al. 2010; Francischetti et al. 2011), which generally yielded a few thousands ESTs. The number of ESTs varies per study from five hundred to several thousands (Nene et al. 2002; Valenzuela et al. 2002; Santos et al. 2004; Ribeiro et al. 2006, 2011; Alarcon-Chaidez et al. 2007; Mans et al. 2008b; Francischetti et al. 2008a,b; Chmelar et al. 2008; Aljamali et al. 2009; Anatriello et al. 2010).

Early tick transcriptome studies that were performed on *Ixodes* species (Valenzuela et al. 2002; Francischetti et al. 2005; Ribeiro et al. 2006; Chmelar et al. 2008) provided insights into the qualitative aspects of tick salivary gene expression. This provided a foundation for understanding the molecular interface between ticks, hosts and transmitted pathogens. The first tick sialotranscriptome revealed the existence of many protein families in secreted saliva of *I. scapularis* (Venenzuela et al. 2002). Subsequently, these protein families were identified and confirmed in other tick species (Francischetti et al. 2005, 2008a; Alarcon-Chaidez et al. 2007; Wang et al. 2007; Chmelar et al. 2008; Aljamali et al. 2009). A study by Ribeiro et al (2011) revisited the *A. variegatum* sialotranscriptome to improve the throughput by combining the ESTs and proteomics (Ribeiro et al. 2011). This approach could only reveal more secreted proteins, still leaving most proteins with unknown functions. That still posed a problem and proved the use of ESTs to have limitations. The next generation sequencing NGS) would help to identify more proteins, and by doing so increasing the high-throughput. However, the NGS just leads to even more proteins with unknown function. Consequently, this led to the need to explore newer and better ways to achieve highthroughput in an effort to identify tick sialotranscriptome proteins. Such proteins

could play significant roles in tick-host interactions; tick feeding; tick survival. More importantly, such proteins can be good candidates for anti-tick vaccines. Years later, the next generation sequencing (NGS) was introduced in tick sialotranscriptome studies (Karim et al. 2011).

The newest insights into improved tick sialotranscriptomes are progressively being driven by NGS technologies. Consequently, comprehensive transcriptional data can be reliably acquired at the level of individual tissues and cell types. Subsequently, that allows for de novo assemblies of transcriptome data. In turn, such transcriptome can provide genetic and phylogenetic information even in the absence of an assembled genome sequence. In the past years, there has been a gradually rapid adoption of NGS technologies to undercover underlying aspects relevant to the biology, feeding, tick-host interactions and reproduction of different tick species. For more unsequenced tick species, *de novo* transcriptome assembly from RNA-seq data can provide sufficient information to reveal their genome coding sequences, SNPs and differential expression. With such information being continuously generated, that should be publicly available to other researchers via public databases such as NCBI's Sequence Read Archive (SRA) (Rinker et al. 2016).

Compared to previous sequencing methods, NGS produce larger datasets of giga base rather than kilo base scale produced by traditional methods (Metzker, 2010). Tick sialotranscriptomes have been described in two studies by NGS technologies. Those studies yielded a highthroughput data on salivary gland transcripts from *A. maculatum* (Karim et al. 2011) and *I. ricinus* (Schwarz et al. 2013). NGS-based transcriptomics projects can produce hundreds of thousands of assembled contigs (Schwarz et al. 2013) and over 50 000 unique transcripts (Tan et al. 2015; Xu et al. 2015). The unprecedented transcriptome coverage by NGS enables statistically reliable analysis of gene expression dynamics of secreted salivary proteins throughout the course of tick feeding, the comparison of tissue-and developmental stage-specific transcripts accumulation and metabolic pathways analysis (Kotsyfakis et al. 2015a).

The current study aimed at characterizing the *R. (B.) microplus* salivary gland transcriptome. Due to lack of a fully sequenced genome for *R. (B.) microplus*, a *de novo* approach was used to achieve this aim. This was made possible by Illumina next generation sequencing technology. In the past it has been shown that tick transcriptomes contain a higher abundance of housekeeping proteins than secreted proteins (Anatriello et al. 2010; Francischetti et al. 2011; Karim et al. 2011; Garcia et

al. 2014; Kotsyfakis et al. 2015b). Tick transcriptome studies have sought to undercover secreted proteins as they play vital roles in tick physiology; feeding; host interaction and reproduction. The most represented secreted protein families are Kunitz-domain proteins; lipocalins; metalloproteases and basic tail proteins (Karim et al. 2011; Schwarz et al. 2013; Garcia et al. 2014; Kotsyfakis et al. 2015b).The pioneering tick transcriptome work by Karim et al. (2011) provided the first transcriptome study employing the next generation sequencing approach.

3.2. Results

3.2.1. RNA isolation and purity

RNA was isolated from salivary gland extracts prepared from five different feeding stages of *R. (B.) microplus* ticks. Nanodrop was used to determine the purity and yield of the total RNA extracted and the results are illustrated in (**Figure 3.1**). The RNA concentration values were within the recommended range for Illumina sequencing protocols, which is 0.1 – 4 µg. This provided sufficient quality and quantity to produce a good library for sequencing.

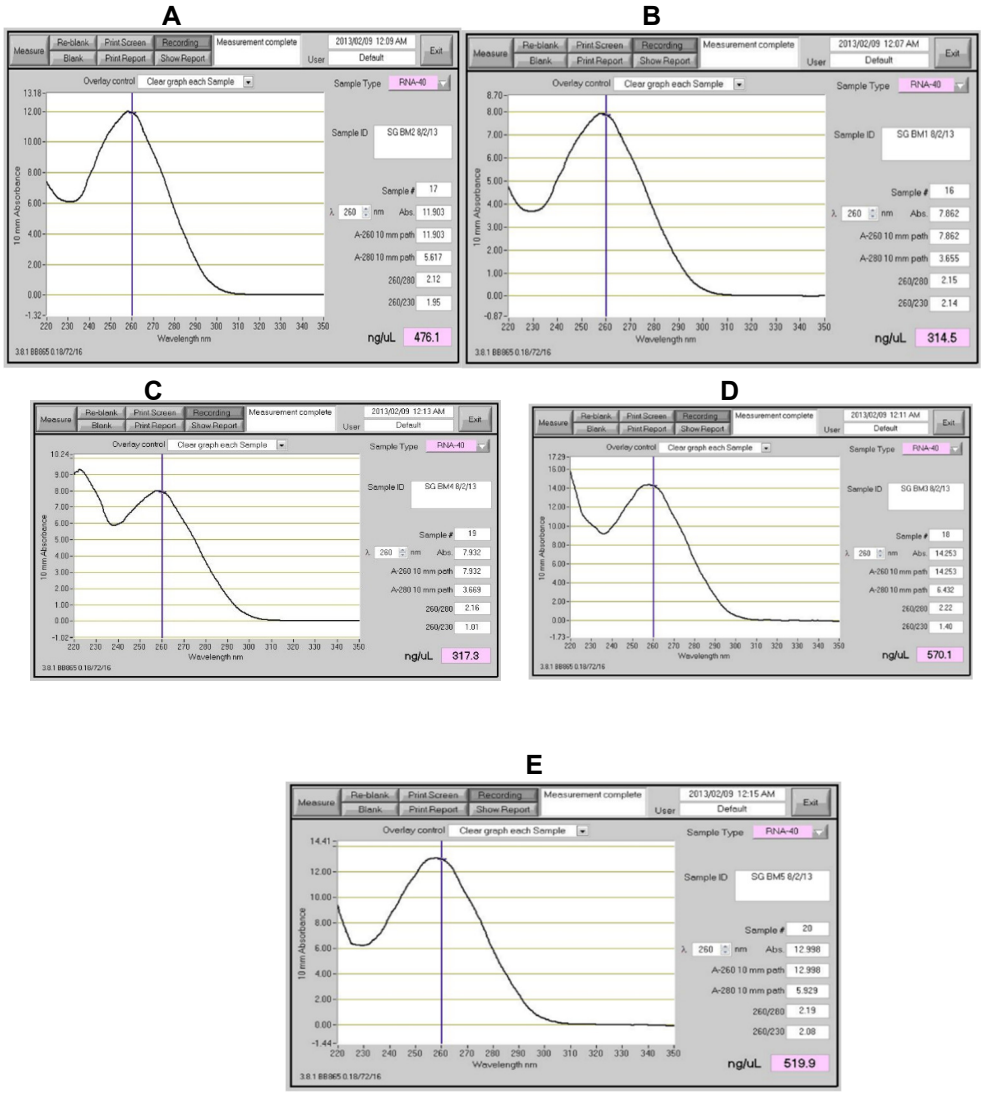

Figure 3.1: RNA integrity results as measured with NanoDrop. A to E is BM1 to BM5, respectively.

Table 3.1. The results for NanoDrop 260/280 and 260/230 ratios for RNA purity quantification.

Sample	260/280 ratio	260/230 ratio
BM 1	2.15	2.14
BM 2	2.12	1.95
BM 3	2.22	1.40
BM 4	2.16	1.01
BM 5	2.19	2.08

3.2.2 cDNA library validation

The cDNA library validation showed that the concentrations of cDNA library synthesis was adequate for Illumina MiSeq sequencing. The synthesized libraries indicated smears ranging from the sizes of ~500 bp – 2500 bp. These sizes suggested the integrity of high molecular size transcripts. The bands for sequencing were cut at 600 bp (and above) as it shown in **Figure 3.2**. The below part of the band was kept as a backup. Subsequently, this suggested that the mRNA integrity was good (not degraded) (**Table 3.2**).

Table 3.2. The results for cDNA libraries quantification with Qubit High Sensitive (HS) kit.

cDNA library	Concentration (ng/µl)
BM 1	3.2
BM 2	39.3
BM 3	38.4
BM 4	35.7
BM 5	43.7

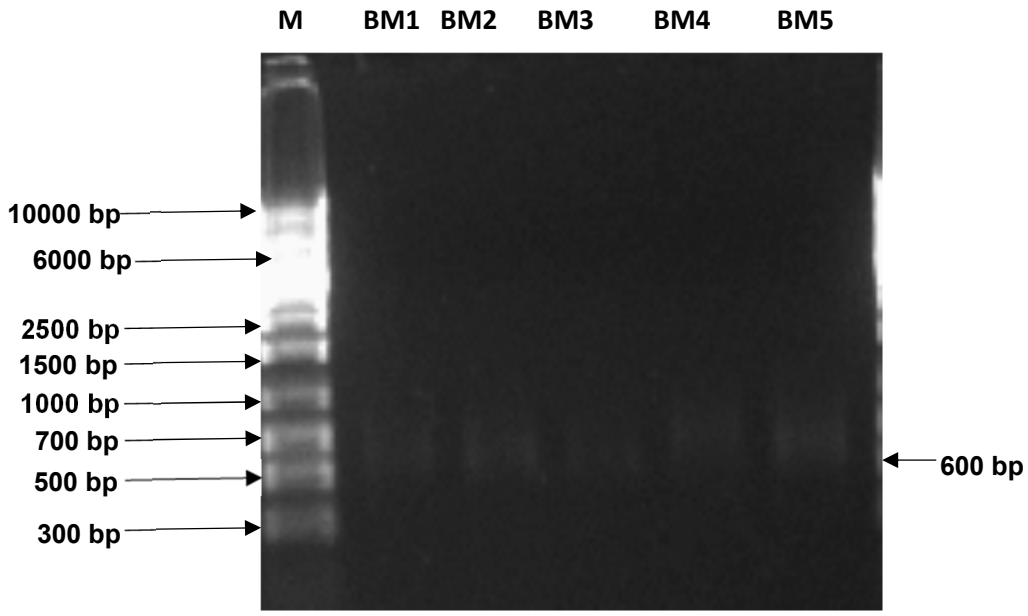

Figure 3.2: A 1% agarose gel electrophoresis depicting cDNA library integrity results. M (1 Kb Ladder DNA marker), BM1 – 5 are the different tick feeding stages. The bands for sequencing were cut at 600 bp (and above). The part below the 600 bp band was kept as a backup.

3.2.3. Illumina sequencing and quality trimming

Literature has indicated differential expression of salivary gland transcripts during tick feeding. In order to account for this possibility and test the hypothesis that differential expression occurs, ticks were fed and sampled at different feeding points. From those, five separate salivary gland transcriptomes cDNA libraries of *R. (B.) microplus*, were generated. The cDNA libraries prepared from the RNA of these feeding stages were subjected to RNASeq using Illumina MiSeq technology. A total number of ~113 million untrimmed reads were generated and, after quality trimming and generation of various data structures represented 42 data structures that all possessed different numbers of reads (**Table 3.3**).

Table 3.3: Statistics for quality trimming and creation of various data structures for *R. (B.) microplus*. Samples include BM1, BM2, BM3, BM4, BM5 and all samples combined (All). Data structures include raw reads before quality trimming (Untrimmed), paired reads that was quality trimmed and merged (Merged), Merged reads from which duplicates were removed (Mddup), single reads that was quality trimmed (Single), Single reads from which duplicates were removed (Sddup), combined Merged and Single reads (SM), combined Sddup and Mddup reads (SMddup).

Samples	Untrimmed	Merged	Mddup	Single	Sddup	SM	SMddup
BM1	22,056,581	5,609,572	4,060,830	16,499,335	11,503,179	22,108,907	15,564,009
BM2	25,844,254	7,022,173	5,078,768	19,694,787	14,131,827	26,716,960	19,210,595
BM3	21,558,672	6,584,118	4,415,584	6,609,949	5,097,206	13,194,067	9,512,790
BM4	22,163,383	5,115,562	3,678,815	16,753,227	12,489,630	21,868,789	16,168,445
BM5	21,786,218	4,797,721	3,604,754	16,166,905	12,011,727	20,964,626	15,616,481
All	113,409,108	29,129,146	20,838,751	75,724,203	55,233,569	104,853,349	76,072,320

3.2.4. *De novo* assembly

De novo assembly resulted in 462 assemblies from which open reading frames (ORF's) were extracted resulting in 15,310,502 total ORF's (**Table 3.4**). The ORF's were clustered using CD-HIT to produce a non-redundant dataset resulting in 549,475 ORF's. Chimeric sequences were removed using CD-HIT indicating that 14% of ORF's were chimeric assemblies. Only ORF's with an average coverage > 4 were retained resulting in 83,020 ORF's. Even though it is assumed that the largest ORF represents the correct reading frame (in this case reading frame 1), BLASTX analysis was used to confirm this. In a number of cases, the BLASTX hit indicated that the correct reading frame was not the largest. After extraction of the correct reading frames the ORF number were reduced to 30,537. This indicates a significant loss in transcripts, most probably due to truncated ORFs in the incorrect reading frame.

Table 3.4: Reduction of ORF numbers to obtain a representative set for analysis.

ORF reduction step	**Number of ORFs**
ORF's before clusterization	15,310,502
ORF's after clusterization on CD-HIT	549,475
ORF's after removal of chimeric sequences,	471,158
ORF's with coverage > 4	83,020
ORF's after BLASTX analysis and reading frame correction	30,537

3.2.5. BUSCO analysis to determine the quality of the transcriptome

Benchmarking Universal Single–Copy Orthologs (BUSCO) assessments are used to robustly estimate completeness of the datasets, however, some technical limitations may increase proportions of "fragment" and "missing" BUSCOs, especially for large genomes. BUSCO quality assessments provide high-resolution quantification citeable in the simple C [D], F, M, n notation for genomes, gene sets and transcriptomes. This

facilitates informative comparisons for instance, of newly sequenced draft genome assemblies to those of important organism models (for example *R. (B.) microplus*), or to quantify iterative improvements to assemblies or annotations (Simâo et al. 2015).

Analysis of the final set of nucleotide sequences using BUSCO (Bmic: 85.1% [S: 42.4%; D: 42.7%], F: 1.6%, M: 13.3%, n=1066) indicated that the transcriptome possess 85.1% of the 1066 conserved arthropod genes in the BUSCO database. Based on the BUSCO parameters 42.4% of these are single, 42.7% duplicated, 1.6% fragmented and 13.3% missing. It was important to compare the quality of the transcriptome from this study with those of other related tick species. The sequences from this study were analysed, and compared to the available data in databases. Sequencing technologies, assemblers and bioinformatics tools can play a significant role in the final transcriptome completeness validation. The *D. variabilis* whole-body transcriptome yielded over 271 million clean reads. The BUSCO analysis of this transcriptome confirmed over 90% completeness (Rosendale et al. 2016). Cramaro et al (2017) used PacBio sequencing system to produce just above 5 million ultra-long reads from *I. ricinus* adult females and males. The estimation of the genome assembly completeness by BUSCO retrieved 55.5% completeness for unique orthologs. The study of *I. scapularis* allowed for the BUSCO transcriptome validation of 68.9 [2.4], 21.0, 10.1, which indicated a 68.9% completeness with 2.4% duplicates; 21% partially recovered genes and 10.1% unrecovered genes (Hoy et al. 2016). In comparison with these studies, our study had better transcriptome completeness. However, Rosendale et al. (2016) showed over 90% completeness in *D. variabilis* whole-body transcriptome. Overall, this suggests that our study proved to have deeper coverage due to deep sequencing by the Illumina sequencing. However, Rosendale study (Rosendale et al. 2016) may be more complete because it sequenced a whole-body transcriptome compared to a salivary gland one.

3.2.6. Functional annotation

The reads were classified into four main classes: housekeeping proteins (HKP), secreted proteins (SP), proteins annotated as hypothetical or unknown function (UNK) and reads with no BLAST hits (NA). In summary, 19 979 ORFs were associated with housekeeping proteins (65.4%), 2 649 were secreted proteins (8.7%), 5 706 were unknown (18.7) and 2 203 were NA (7.2%) (**Table 3.5**).

Table 3.5: Protein families identified from the annotation of the *R.* (*B.*) *microplus* salivary transcriptome.

Protein class	**Number of identified proteins**
Secreted proteins	2 649
Housekeeping proteins	19 979
Proteins with unknown functions	5 706
Proteins with no hits	2 203

3.2.6.1. Housekeeping proteins

In this study the majority of main putative housekeeping proteins identified were associated with protein translation (558), serine proteases (289), transport (261), transposable elements (188), amino acid metabolism (177) and energy metabolism (177) (**Table 3.6**) and (**Figure 3.3**).

Table 3.6: Pathway classification of main housekeeping genes in *Rhipicephalus* (*Boophilus*) *microplus* salivary transcriptome.

Category	**Count**
Translation	558
Serine proteases	289
Transport	261
Transposable elements	188
Amino acid metabolism	177
Energy metabolism	177

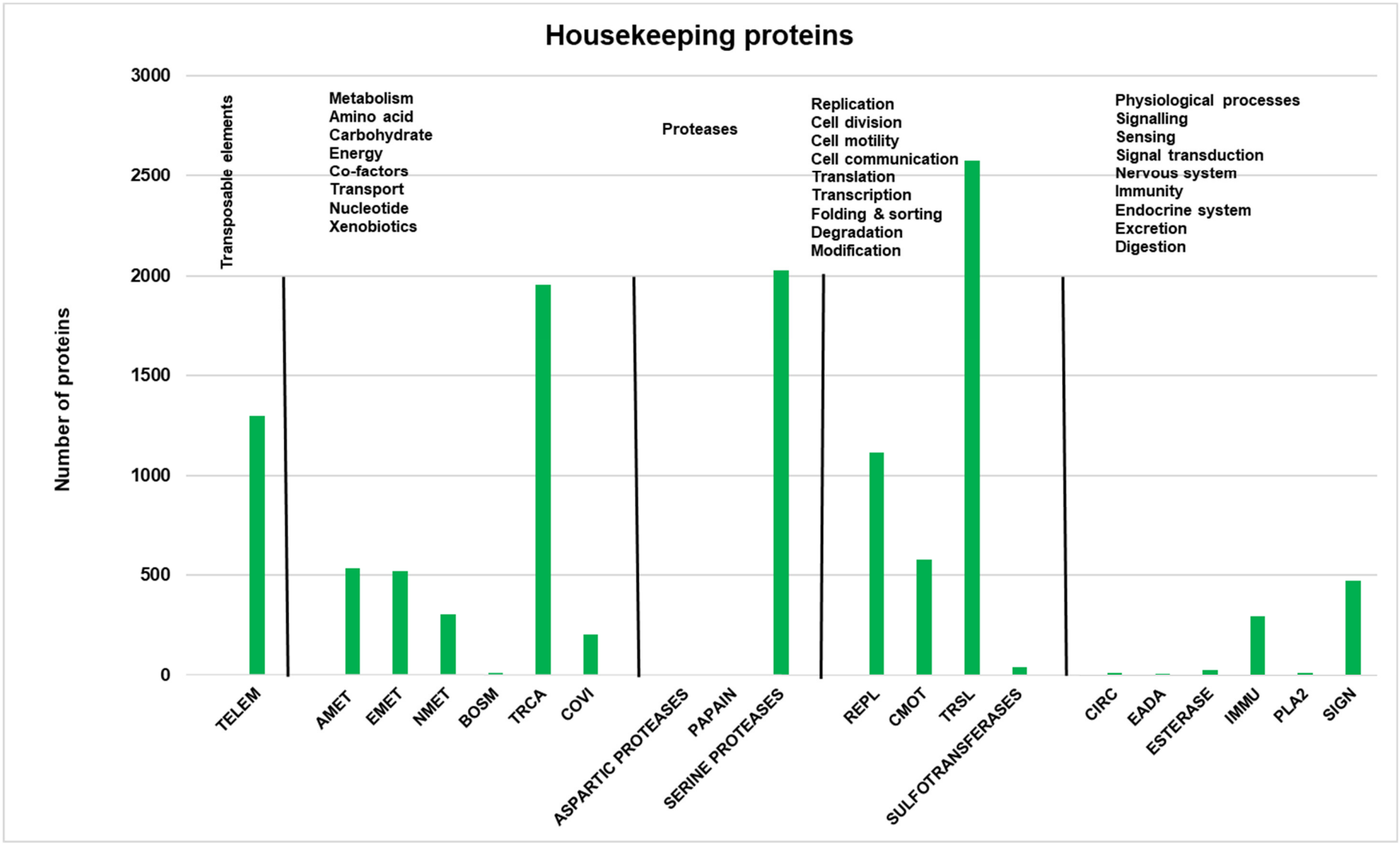

Figure 3.3: Putative housekeeping proteins identified in *R. (B.) microplus* (ARC-OVR) transcriptome database. Proteins associated with FSDE (folding, sorting and degradation) and metabolism were expressed at higher levels relative to other protein families. **Abbreviations:** TELEM (transposable elements), AMET (amino acid metabolism), EMET (energy metabolism), BOSM (co-factors), TRCA (transport and catabolism), CDVI (cell growth and death), REPL (replication), CMOT (cell motility), TRSL (translation), CIRC (circulation), EADA (environmental adaptation), IMMU (immunity), PLA2 (phospholipase A2), SIGN (signalling).

3.2.6.2. Putative secreted proteins

The major secreted protein families that were identified in this study were lipocalins, BTSP, BPTI/Kunitz, cement proteins and metalloproteases. In total, lipocalins were the largest group of secreted proteins in the total *R. (B.) microplus* transcriptome database (ARC-OVR) with 504 (**Figure 3.4**) and (**Table 3.2**). This pattern corresponds with what has been shown by other tick transcriptome studies. The BTSPs were the least abundant protein family shown, with a total of 122 proteins. Overall, The *R. (B.) microplus* salivary gland transcriptome BLAST search results yielded only 2,649 out of total 30,537 (8.67%) ORFs that were annotated as secreted proteins (**Figure 3.5**), the majority of which retrieved *R. pulchellus* as best BLASTP for total protein hits. This may be due to the good transcriptome database deposited for *R. pulchellus*, or that *R. pulchellus* is closer related to *R. (B.) microplus* to the exclusion of other *Rhipicephalus* species. Moreover, this could suggest unique *R. (B.) microplus* ORFs were found in this current study since they did not retrieve much *R. (B.) microplus* sequences. Proteins from the tick *R. (B.) microplus* were the second largest group in species count as shown in (**Figure 3.5**). The current *R. (B.) microplus* transcriptome database contributes more proteins than is currently available for *R. (B.) microplus*. This could be due to the advantage of using both molecular biology techniques and high-throughput next-generation sequencing techniques, like Illumina MiSeq sequencing.

Table 3.7: List of most abundant *R. (B.) microplus* salivary gland transcriptome secretory protein families.

Protein family	**Secreted proteins from our *R (B) microplus* database**
Lipocalins	504
Metalloproteases	131
Cement	148
BPTI	184
BTSP	122

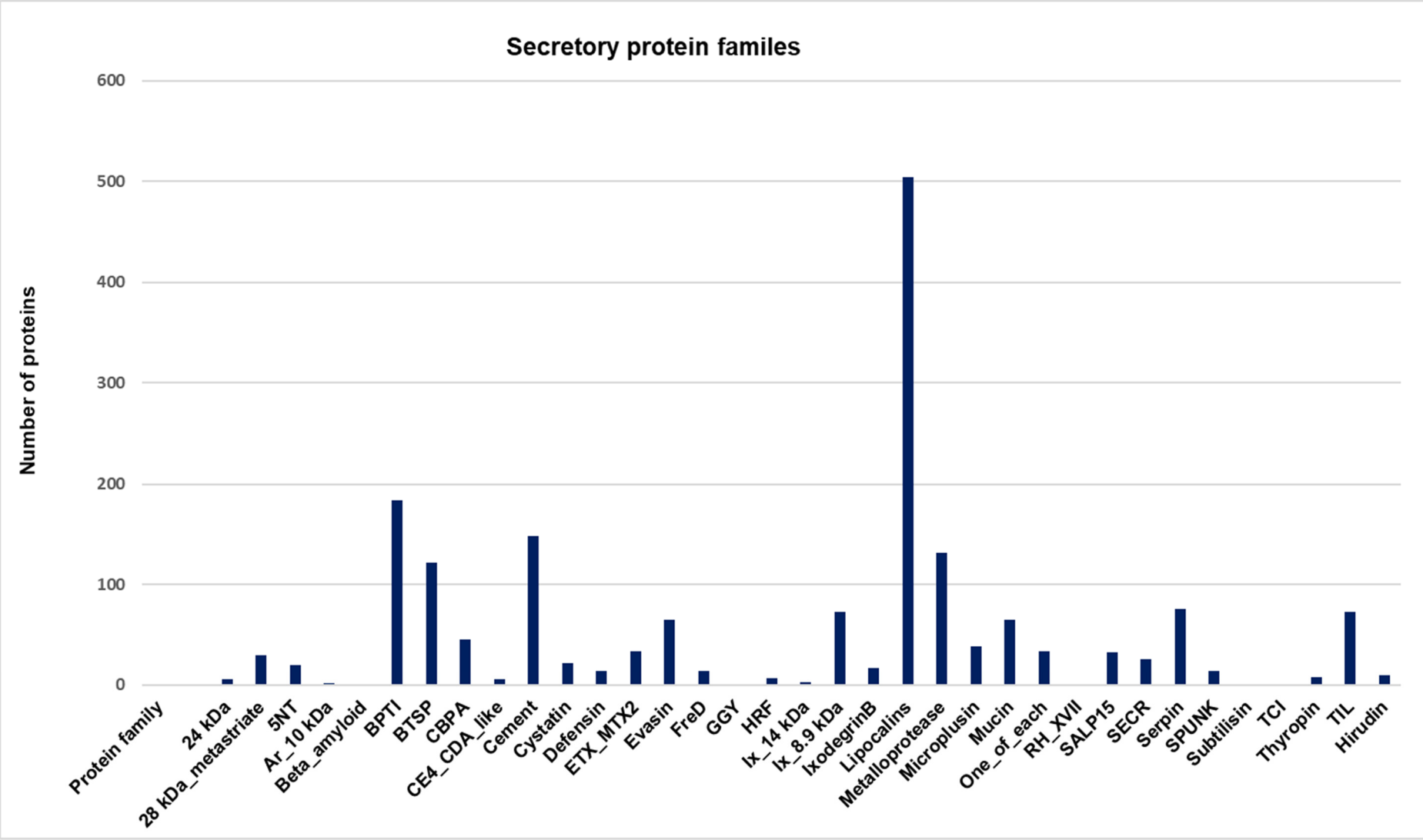

Figure 3.4: Putative secreted proteins identified in *R. (B.) microplus* (ARC-OVR) transcriptome database. That could be attributed the deeper sequencing and coverage obtained by the study design. Lipocalins cement proteins and metalloproteases protein families were the highest identified. **Abbreviations:** 5NT (5'nucleotidase), AG5 (antigen 5), BPTI (bovine pancreatic trypsin inhibitor), BTSP (basic tail secreted protein), CBPA (choline binding protein A), FreD (fibrinogen-related domain), HRF (histamine release factor), Ix (Ixodegrin), SALP15 (tick salivary protein 15), SECR (tick secreted protein), serpin (serine protease inhibitor), TCI (tick carboxypeptidase inhibitor), TIL (trypsin inhibitor-like cysteine rich domain).

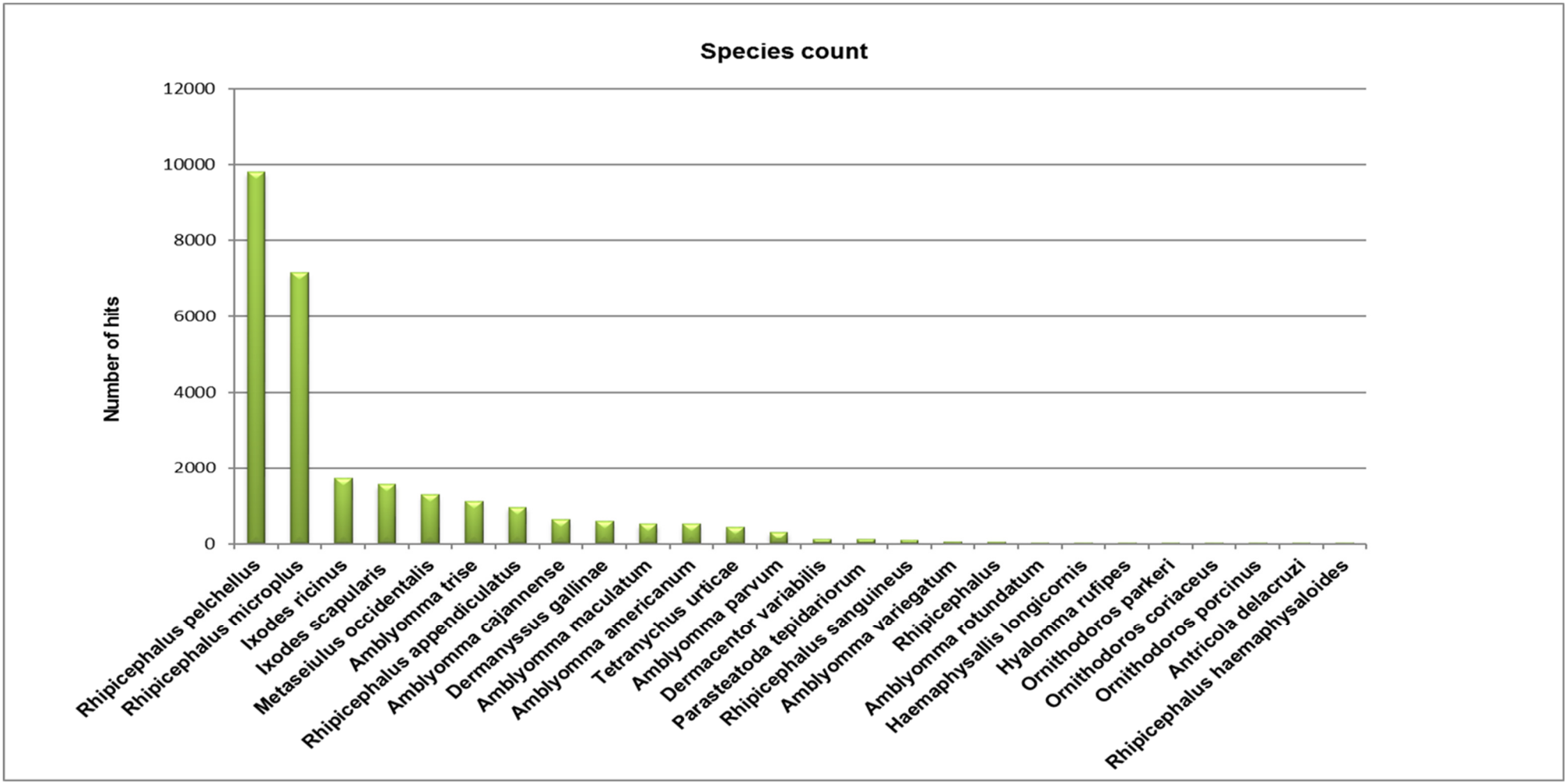

Figure 3.5: Species count from our *R. (B.) microplus* (ARC-OVR) transcriptome study with the highest and best hits on the NCBI database.

3.3. Discussion

In total 30,537 of open reading frames (ORFs) were assigned to KEGG housekeeping protein classes, secretory protein families, proteins with unknown functions and those that belonged to the NA class. The majority of housekeeping proteins were associated with protein translation and serine proteases. Lipocalins, metalloproteases and proteins of cement function were most abundant of the secreted proteins (**Table 3.7**). The high percentage of proteins with unknown functions (18,7%) might contain secreted proteins that are not yet characterized. Unexpectedly, the greatest number of BLAST hits were not obtained for *R. (B.) microplus*, but *for R. pulchellus* from an annotated Acari protein database (**Figure 3.5**). This suggests that unique *R. (B.) microplus* ORFs were found in this study. This was reflected in both higher numbers of housekeeping and secretory protein family members found compared to the *R. (B.) microplus* sequences available. The majority of housekeeping proteins were those involved in biological processes such as secretion; transcription and translation, which fits the profile expected for salivary glands. It is known that ticks produce (express) proteins as they progress with their blood feeding and some of those proteins play vital roles in processing the blood meal and helping the tick to remain attached to the host. The dominant secretory protein families found were those known to be involved in regulation of host haemostatic and immune defences as well as tick attachment. These included BPTI; BTSP; lipocalin; metalloproteases and cement proteins. Typically, this phenomenon has been found in other tick sialotranscriptomes (Ribeiro et al. 2006, 2011; Mans et al. 2008b; Francischetti et al. 2009). The secretory proteins were the lowest in number, suggesting that they may be expressed at higher levels compared to housekeeping proteins.

3.3.1. Housekeeping proteins

Housekeeping proteins are essential to maintain the cellular processes of organisms. These proteins are regularly expressed throughout the life of all organisms. Such proteins include, among others, proteins responsible for metabolism, transcription, signal transduction, protein synthesis, modification and export (Kim et al. 2016). Most studies have previously shown that the majority of housekeeping proteins expressed in all living organisms are responsible for metabolism, carbohydrates, amino acids and cytoskeleton (Francischetti et al. 2009; Schwarz et al. 2014; Kim et al. 2016). This is the same when it comes to ticks and tick feeding. Ticks express housekeeping proteins during their feeding.

Detoxification (enzymes) is one of the vital process during tick feeding. In this study, many enzymes responsible for detoxification were also abundant such as sulfotransferases. These might play a role in detoxification of dopamine (Pichu et al. 2011; Yalcin et al. 2010), a secretagogue found in salivary glands (Pichu et al. 2011). NADH-ubiquinone oxidoreductases, ATP synthases and cytochromes had higher expression frequency. The latter possess some defense mechanism. Graham et al. (2016) recently isolated a cytochrome P450 gene, CYP3006G8 from *R. (B.) microplus*. This was the first CYP to be shown to putatively metabolize pesticide. Pesticides are obviously detrimental to ticks, and ticks need to overcome that to have successful blood meal. Probably ticks evolved the CYP to deal with human derived pesticides. Therefore, the expression of these CYPs could be expected, perhaps at all feeding/life stages and under different stimuli.

Heat shock proteins (such as HSP70), proteases and glycosyltransferases were also abundant in this study. This group of proteins is involved in post-translational modification and protein degradation. The tick salivary glands start to degenerate towards the end of the tick feeding and are almost completed within four days of the tick detaching (Sauer et al. 1995; Freitas et al. 2007). It is possible that these proteins may play a role in salivary gland degeneration.

3.3.2. Putative secreted proteins

Our data revealed tick protein families that have been previously found in abundance by other studies. A core set of protein families are conserved between ixodids and argasids (Mans et al. 2008a). These families are often found in abundance and they include amongst others, lipocalins; BTSP; BPTI/Kunitz and metalloproteases (Mans et al. 2008a). Lipocalins were the largest group of secreted proteins (1.65%), followed by the metalloprotease group (1.61%) (**Figure 3.4**). For the purposes of our study, only a few classes of proteins from the secretory group will be further discussed. These are lipocalins, metalloproteases, cement proteins and Kunitz/BPTI, and are discussed below.

(a) Lipocalins

Mammals secrete serotonin also known as 5-Hydroxytryptamine (5-HT)[2] and histamine as a defense mechanism in general. Serotonin helps to induce platelet aggregation and vasoconstriction (Van Nueten et al. 1985), while histamine allows monocytes and neutrophils permeability and filtration in the tick feeding site (Jutel et al. 2001b). To circumvent these mammalian hosts' responses, ticks scavenge both the serotonin and histamine. Both serotonin and histamine were shown to be scavenged by lipocalins in some ticks (Paesen et al. 1999; Sangamnatdej et al. 2002).

Lipocalins possess a highly conserved barrel structure with which they interact with their hydrophobic ligands. These proteins are expressed in abundance in tick salivary glands (Francischetti et al. 2005; Ribeiro et al. 2006; Alarcon-Chaidez et al. 2007). Tick lipocalins have been described to be involved in histamine binding (Paesen et al. 1999), anti-haemostasis (Rodriguez-Valle et al. 2013), anti-complement activity (Nunn et al. 2005) and toxicity (Mans et al. 2003). Lipocalins also known to play a role in TXA2 scavenging and inhibition of platelet aggregation (Mans et al. 2008c, LTB4 scavenging (Beaufays et al. 2008; Mans et al. 2008c; LTC4 scavenging (Mans et al. 2008b).

Lipocalins are known to bind both the serotonin and histamine. Mans et al (2008c) isolated two abundant lipocalins (monomine and monotonin) from *A. monolakensis* salivary glands. Monomine (AM-10) and monotonin (AM-38) were shown to bind

histamine and serotonin (5-HT), respectively. They both showed high affinity binding and that would suggest that they can be effective lipocalins scavengers at the feeding site thereby inhibiting host inflammatory responses. Furthermore, mast cells and neutrophils release to help in fighting inflammatory responses (White and Kaliner, 1987). The cutaneous inflammation mediated by histamine, is one of the host's defence reactions to tick feeding. Given the importance of lipocalins in ticks feeding survival, it can be speculated that *R. (B.) microplus* tick saliva glands can possess proteins like histamine binding lipocalins to sequester histamine in order to circumvent the inflammation response. However, further work still needs to be done to test such proteins' biological functions in ticks.

(b) Metalloproteases

Metalloproteases are multifunctional proteins with diverse biological functions in several organisms, including ticks. These proteins have been identified in various tick species. They are associated with blood meal related functions (Francischetti et al. 2003; Harnnoi et al. 2007; Decrem et al. 2008a; Barnard et al. 2012). Saliva metalloproteases play a vital role in fibrin and fibrinogen breakdown (Francischetti et al. 2003). To prove that, RNAi was used to inhibit *I. ricinus* metalloproteases expression. That prevented blood meal completion and inhibited saliva fibrinolytic activity (Decrem et al. 2008a, b). Francischetti et al (2003) confirmed that *I. scapularis* saliva contains metalloproteases with fibrinolytic, antifibrinogen and antifebrin specific activities. Metalloproteases transcripts were also found in *I. ricinus* salivary glands. These were in accordance with the idea that metalloproteases participate in hemostatic modulation and in facilitating blood meal completion (Francischetti et al. 2003; Decrem et al. 2008a).

A few studies used vaccination to determine the efficiency of tick metalloproteases in modulating host immune responses. Rabbit vaccination with *I. ricinus* recombinant metis metalloproteases was shown to prevent blood meal completion, without altering tick survival rate. Recently, a recombinant reprolysin metalloprotease (BrRm-MP4) from *R. (B.) microplus* was tested against *R. (B.) microplus* infestation in cattle. Vaccination with rBrRm-MP4 provided immune response, and reduced tick feeding success, engorged female numbers and their reproduction. The rBrRm-MP4 provided

an overall of 60% protection, and that made it a good potential anti-tick candidate (Ali et al. 2015a).

The reprolysin family of metalloproteases was shown to be dominant in the salivary glands of *R. (B.) microplus*, while the astacin family was prominent in the midgut and ovary (Barnard et al. 2012). BLASTP analysis of the transcriptome from the current study indicated that the majority of metalloproteases belong to the reprolysin family in support of the previous observations. A metalloprotease enzyme, BooKase (*Boophilus Kininase*) was identified in *R. (B.) microplus* salivary glands. This enzyme inactivates bradykinin and lowers its levels at host attachment site, relieving pain (Bastiani et al. 2002). The metalloprotease transcription upregulation in salivary glands has been shown to correlate with blood feeding (Francischetti et al. 2003; Harnnoi et al. 2007; Decrem et al. 2008a; Barnard et al. 2012). Tirloni et al. (2014) found that *R. (B.) microplus* partially engorged female ticks have significant amounts of metalloproteases compared to the fully engorged ones. That could support the fact that host defence modulation is crucial during the early feeding stages.

(c) Basic pancreatic trypsin inhibitor/Kunitz

Both hard and soft tick sialotranscriptomes have been studied extensively over the years. These studies have revealed several protein families that are believed to play significant roles in tick feeding; tick reproduction and tick-host- pathogen interactions. Some of these protein families are serine protease inhibitors that play roles, mainly, in tick feeding. The majority of such serine proteinase inhibitors belong to the BPTI/Kunitz family. In the past, several molecules have been studied extensively. These include anticoagulant peptide (TAP) (Waxman et al. 1990); a blood coagulation factor Xa inhibitor, purified from *O. moubata*; thrombin inhibitors, ornithodorin (Van de Locht et al. 1996) and savignin (Mans et al. 2002b) from *O. moubata* and *O. savignyi*, respectively. Factor Xa and FVIIa/tissue factor inhibitors, Ixolaris (Francischetti et al. 2002) and Penthalaris (Francischetti et al. 2004) *from I. scapularis*. Several *R. (B.) microplus* trypsin inhibitors (BmTIs) with inhibitory activity toward human plasma kallikrein (HuPK), human neutrophil elastase (hNE) and plasmin present in larvae and eggs have been described (Tanaka et al. 1999, Andreotti et al. 2001, Sasaki et al. 2004). These Kunitz/BPTI-domain containing proteins have been placed in different

categories, based on their biological activities. These activities include, antimicrobial activity; ion channel blocking; blood uptake and digestion; tick egg production development; FVIIa/TF complex activity blockage; plasmin, elastase and plasma kallikrein blockage (Lima et al. 2010; Tanaka et al. 1999; Sasaki et al. 2004; Francischetti et al. 2002, 2004).

R. (B.) microplus chymotrypsin inhibitor (BmCI) was identified from tick body fat (Lima et al. 2010). This inhibitor inhibits both chymotrypsin and human neutrophil elastase (hNE) that play a role in immune response and blood coagulation. Increased levels of BmCI were observed in hemocytes following *M. anisopliae* infection, suggesting its role in tick defense system (Lima et al. 2010). BmCI was been shown to have other biological activities other than its role in the immune response and blood coagulation. A study by Lima et al (2010) demonstrated that BmCI shares similarities with dendrotoxins. Dendrotoxins have the ability to block ion channels such as Na^+, K^+, Ca^+ that play a role in cell proliferation (Lang et al. 2005) and apoptosis (Nutt et al. 2002). Lima et al (2010) demonstrated that rBmCI cytotoxic effects in mammalian cells caused by an apoptosis-inducing activity.

Tick inhibitors, BmTA and B were identified from *R. (B.) microplus* larvae, and they have shown to inhibit activities of trypsin, elastase and plasma kallikrein. Cattle immunization with BmTI showed gave an overall protection of 72.8%.

A number of saliva anticoagulants were discovered in ticks. The first one, Ixolaris containing two Kunitz/BPTI domains was identified in salivary glands of *I. scapularis* (Francischetti et al. 2002). This inhibitor has two Kunitz domains and similar to human tissue factor pathway inhibitor (hTFPI). It has the same activity as hTFPI to block FVIIa/TF complex and inhibit extrinsic pathway (Francischetti et al. 2002). Later another inhibitor with five tandem Kunitz domains called Penthalaris was identified. Penthalaris binds to the FX and FXa, inhibiting the FVIIa/TF-induced FX activation at high concentrations. As a result, it was implicated in inhibition of coagulation (Francischetti et al. 2004). Given all these significant roles of the members of the BPTI/Kunitz family, it could be suggested that ticks really need these proteins for the completion of their feeding process and survival on the host. The identification of these proteins in this study is in accordance with findings from previous related studies.

(d) Cement proteins

This protein family was found to be among the highly abundant protein families in *R. (B.) microplus* salivary glands. This was in accordance with other tick sialotranscriptome studies (Tirloni et al. 2014; Kim et al. 2016). This supports the hypothesis that cement proteins play a crucial role during the early tick feeding phases. Ticks make use of their specialized mouthparts to attach to their hosts. They remain attached on the host with the help of cement proteins (Kemp et al. 1982). These cement proteins are rich in glycine and have been identified in tick sialome studies. A number of those proteins were anticipated as salivary cement proteins with roles in tick-host attachment during feeding (Francischetti et al. 2009).

Cement proteins form a cone that ensures that ticks remain attached to the host during feeding (Valenzuela, 2004). Ideally, these proteins are expected to be expressed and up regulated during the early phases of tick feeding. Tirloni et al (2014) identified cement proteins in partially engorged female *R. (B.) microplus* ticks saliva. The same phenomena was observed in *A. americanum* (Mulenga et al. 2007). These cement proteins were also identified in other studies, *O. moubata* (Diaz-Martin et al. 2013b) and *R. sanguineus* (Oliveira et al. 2013). The up regulation of these proteins during the early feeding phases is in accordance with the fact that they facilitate tick attachment to the host.

Like the overall aim of this study, which is part of tick vaccine development, some of these proteins have been evaluated as anti-tick vaccines (Bishop et al. 2002; Trimnell et al. 2005). Some tick cement proteins were identified and one of those is TRP64. TRP64 was identified in *R. appendiculatus*. More than being an important protein with a role in tick attachment, this protein has shown to have other important functions in ticks. A number of studies have evaluated TRP64 as an anti-tick antigen. The guinea pigs that were vaccinated with TRP64 showed increased mortality, decreased weight and egg mass in *R. appendiculatus* ticks (Trimnell et al. 2002). Vaccination of mice with TRP64 showed both tick feeding inhibition and reduction in encephalitis virus transmission by *I. ricinus* (Labuda et al. 2006). Given that, it can be anticipated that cement proteins from this study might possess anti-tick and antimicrobial properties. Further work is needed to characterize such proteins and possibly determine their

biological roles in anti-tick activities. Therefore, it cannot be concluded, yet, that this study had revealed novel tick cement proteins.

3.3.3. Improved *R. (B.) microplus* salivary gland transcriptome coverage

The *R. (B.) microplus* salivary gland transcriptome yielded higher protein coverage for housekeeping proteins. There was also a high number of unknown proteins, which could be also housekeeping proteins that were discovered in the *R. (B.) microplus* salivary gland transcriptome (**Table 3.5**). These proteins could not be assigned to any protein family, majority of which were classified as tick salivary gland hypothetical proteins. These proteins still have to be characterized and be assigned to their respective protein families. Some, if not, majority of such proteins could be secretory proteins that were never identified by other tick saliva/salivary gland transcriptome studies. Thus, the higher number of unknown proteins could be attributed to the high coverage obtained by the more advanced experimental design employed in this study. The next-generation sequencing technologies such as Illumina MiSeq have, over the years, proved to be efficient in providing high-throughput transcriptome data. A study on *I. ricinus* midgut and salivary glands transcriptome and proteome, yielded improved protein coverage when the Illumina HiSeq was used for sequencing (Schwarz et al. 2014). Ten libraries produced 35.5 Gb of raw sequence reads, which was greater than the transcriptomic data produced from their previous work (Schwarz et al. 2014).

3.4. Conclusion

In conclusion, the current study yielded more housekeeping proteins than secretory protein as expected based on the published literature. The putative secreted proteins included, lipocalins; metalloproteases; Kunitz and cement proteins. This study improved and increased transcriptome coverage in terms of identified proteins contributing a higher number of *R. (B.) microplus* proteins than currently available in GenBank. This will allow for deposition of more *R. (B.) microplus* proteins into the public protein databases such NCBI. Moreover, the identification of proteins from the transcriptome was further validated by determining the proteins at the proteome level. That would then determine the correlation between the transcriptome and proteome of the *R. (B.) microplus* salivary gland proteins.

Chapter 4: Proteome analysis - data analysis and discussion

4.1. Introduction

Transcriptomic studies have provided useful information with regard to the general characterization and analysis of tick salivary glands. However, proteomic analysis of saliva/salivary glands provides unique information regarding proteins that are actually secreted (Tirloni et al. 2014). Proteomics aims at the large-scale analysis of proteins in various organisms, including ticks (Marcelino et al. 2012). In studies seeking to characterize the protein profiles, proteomic tools are useful. These tools allow large scale, high-throughput, identification and functional analysis of the proteome molecules (Kim et al. 2015). Some of these molecules play specific roles in tick physiological processes such blood feeding and/or digestion processes in ticks (Imamura et al. 2009; You and Fujisaki, 2009; You et al. 2001).

The first tick proteomic study used two-dimensional gel electrophoresis (2DE) followed by MALDI-TOF mass spectrometry analysis (Madden et al. 2002. These studies did not result in the identification of proteins, probably due to the lack of appropriate tick protein databases (Untalan et al. 2005; Vennestrøm and Jensen, 2007). The limitations that hampered the application of proteomics technology in ticks included the scarce amount of material available for proteomics methodologies (Villar et al. 2010) and the little information available in sequence databases for non-model organisms such as ticks. That has created the need to provide suitable databases for the correct assignment of the identified proteins. The first tick genome sequencing assembly of the blacklegged tick, *I. scapularis* (Gulia-Nuss et al. 2016) together with the release of transcriptomics datasets for other species (Chmelař et al. 2016) has increased the number of tick proteins present in the UniProt database (http://www.uniprot.org) from 26,066 proteins (taxonomy Ixodida) to 123,969 proteins, therefore enabling the identification of previously unknown proteins.

The majority of tick proteomics studies have focus on the analysis of tick saliva or salivary gland protein extracts from soft and hard tick species (Valenzuela et al. 2002;

Madden et al. 2004; Oleaga et al. 2007; Francischetti et al. 2008b; Mans et al. 2008a; Francischetti et al. 2011; Diaz-Martin et al. 2013; Mudenda et al 2014; Schwarz et al. 2014; Tirloni et al. 2015). Some of them also combined the proteomics approach with transcriptomics analysis (Francischetti et al. 2008b; Francischetti et al. 2011; Mudenda et al. 2014; Schwarz et al. 2014). Other tick studies sought to profile other tick organs using the combination of proteomics and transcriptomics (Sonenshine et al. 2011; Cramaro et al. 2015). The recent application of (protein identification technology) PIT has allowed for the identification of novel secreted proteins in *D. andersoni* saliva that are potentially relevant for immunomodulation and blood feeding in the first stages of tick attachment (Mudenda et al. 2014). Proteomics has also been applied in analysis of other tick tissues such as midgut, synganglion, spermatophore and ovaries in order to characterize tick biology and find potential vaccine antigens (Schwarz et al. 2014; Di Venere et al. 2015; Tan et al. 2015; Tirloni et al. 2015; Sonenshine et al. 2011). A recent publication reviewed the current knowledge on the transcriptomics and proteomics of tick tissues from a systems biology perspective (Chmelař et al. 2016).

The application of PIT for the characterization of stress response in *D. reticulatus* unfed tick larvae revealed more than one hundred potential antigens. These candidates could reduce tick survival and reproduction (Villar et al. 2014a). The analysis of nymphs and adults from *A. americanum* ticks by quantitative proteomics revealed differences between developmental stages in relevant biological processes such as blood digestion, heme detoxification and innate immunity (Villar et al. 2014b). The findings from this study could help to design antigens for vaccines targeting ticks at their early development stages.

In this this study, *R. (B.) microplus* salivary gland proteome was analysed for comparison purposes with the complementing transcriptome. Proteome analysis was carried out by salivary gland extracts preparation and that was subjected to one-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (1D SDS-PAGE). A specific protocol was optimised and adopted to yield proteomic data of high quality for the purposes of this study.

4.2. Results

R. (B.) microplus salivary gland soluble, membrane and pellet protein fractions were tryptic digested and subjected to mass spectrometry to validate the analysis of the transcripts possibly being secreted proteins (**Figure 4.1-4.3**). The smears observed in the 1DE gels indicate the complexity of the salivary gland extract, which does not resolve into distinct bands when electrophoresed for 3 cm. However, this separation linked with the pre-fractionation steps that collected membrane (M), heat-denatured pellet (P) and soluble (S) fractions was considered sufficient steps to reduce the overall complexity of each fraction to an extent that would allow diagnostic depth in the MS/MS analysis, while reducing the cost of sample analysis.

The Mascot data was viewed by the Scaffold 4 software program, and then analysed by X!Tandem. The four data sets were analysed independently/separately, and ultimately combined to analyse the total number of proteins obtained. As it was observed with the transcriptome findings, the same phenomena was also observed here. The majority of the proteins identified were from our own *R. (B.) microplus* database. They have BLAST hits in the database of which some is TPA (third party analysis of *R. (B.) microplus*). The overall analysis revealed that cement proteins accounted for the majority of secretory proteins.

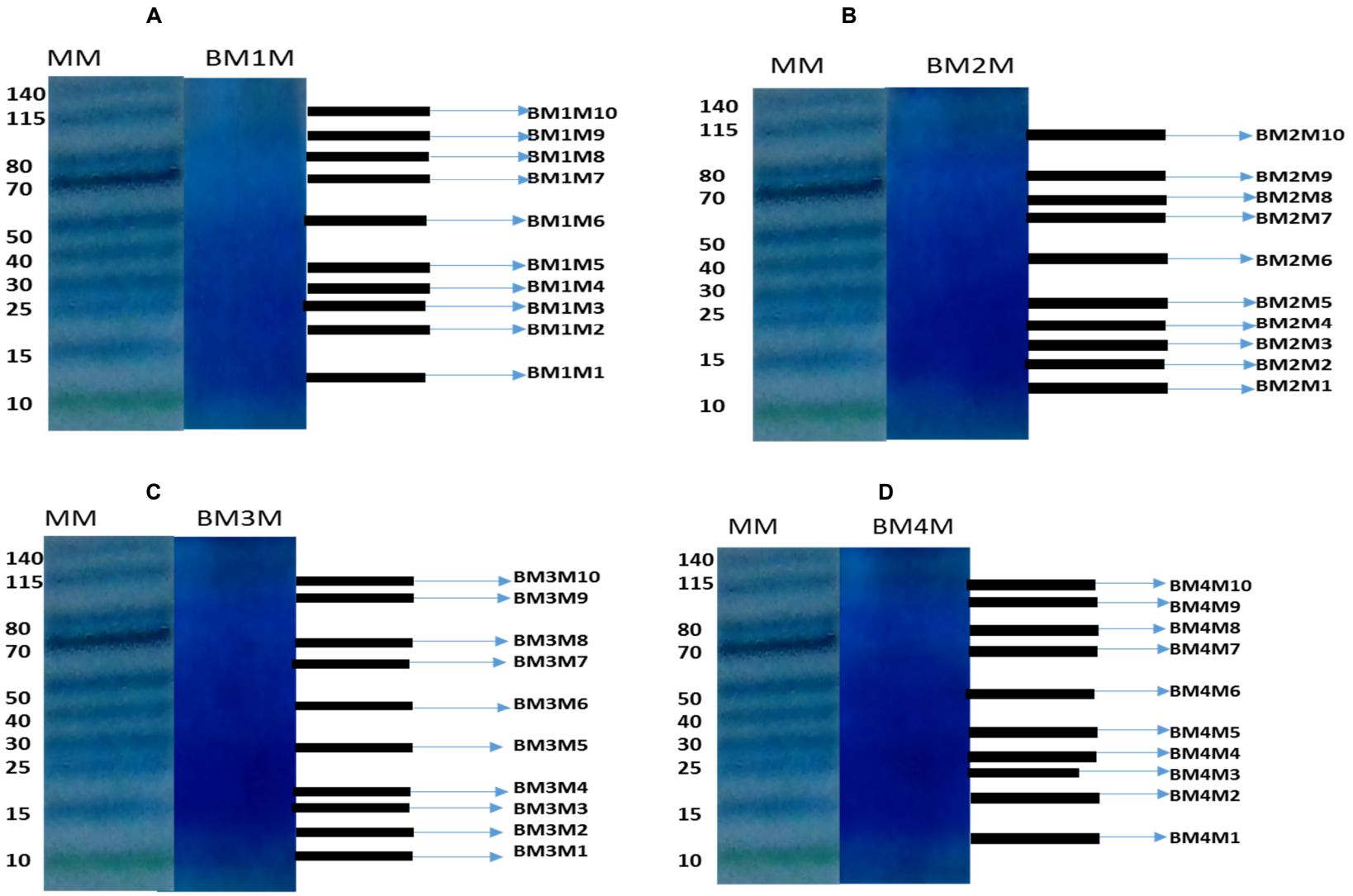

Figure 4.1: An SDS-PAGE separated pattern of *R. (B.) microplus* membrane protein fraction for all feeding stages, BM1 – BM4 stained with Coomassie Brilliant Blue R-250 dye. MM is the protein molecular weight standard. A to D represent BM1 to BM4, respectively, from which 10 gel pieces were cut out for each feeding stage for proteomics studies.

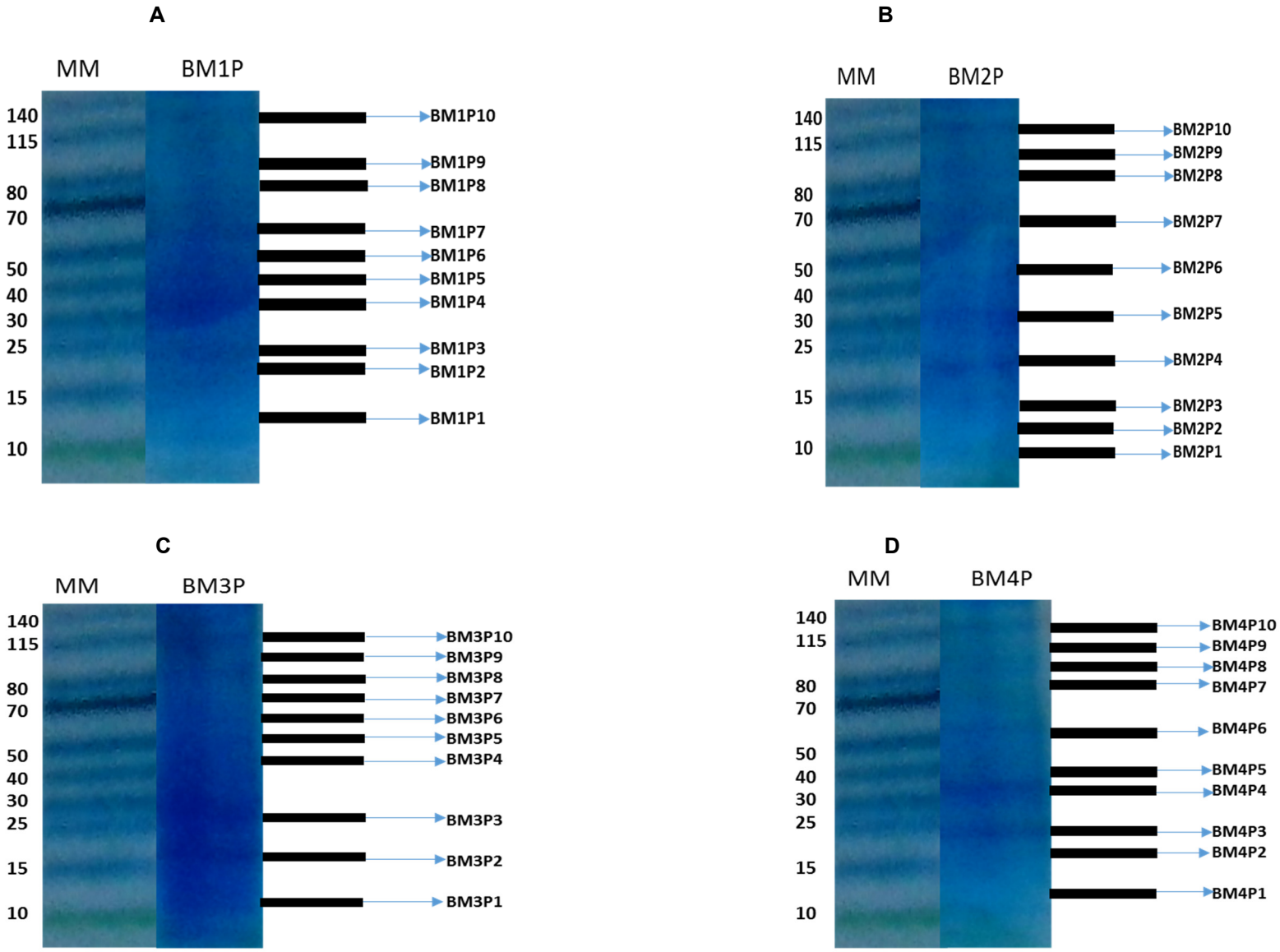

Figure 4.2: An SDS-PAGE separated pattern of *R. (B.) microplus* pellet protein fraction for all feeding stages, BM1 – BM4 stained with Coomassie Brilliant Blue R-250 dye. MM is the protein molecular weight standard. A to D represent BM1 to BM4, respectively, from which 10 gel pieces were cut out for each feeding stage for proteomics studies.

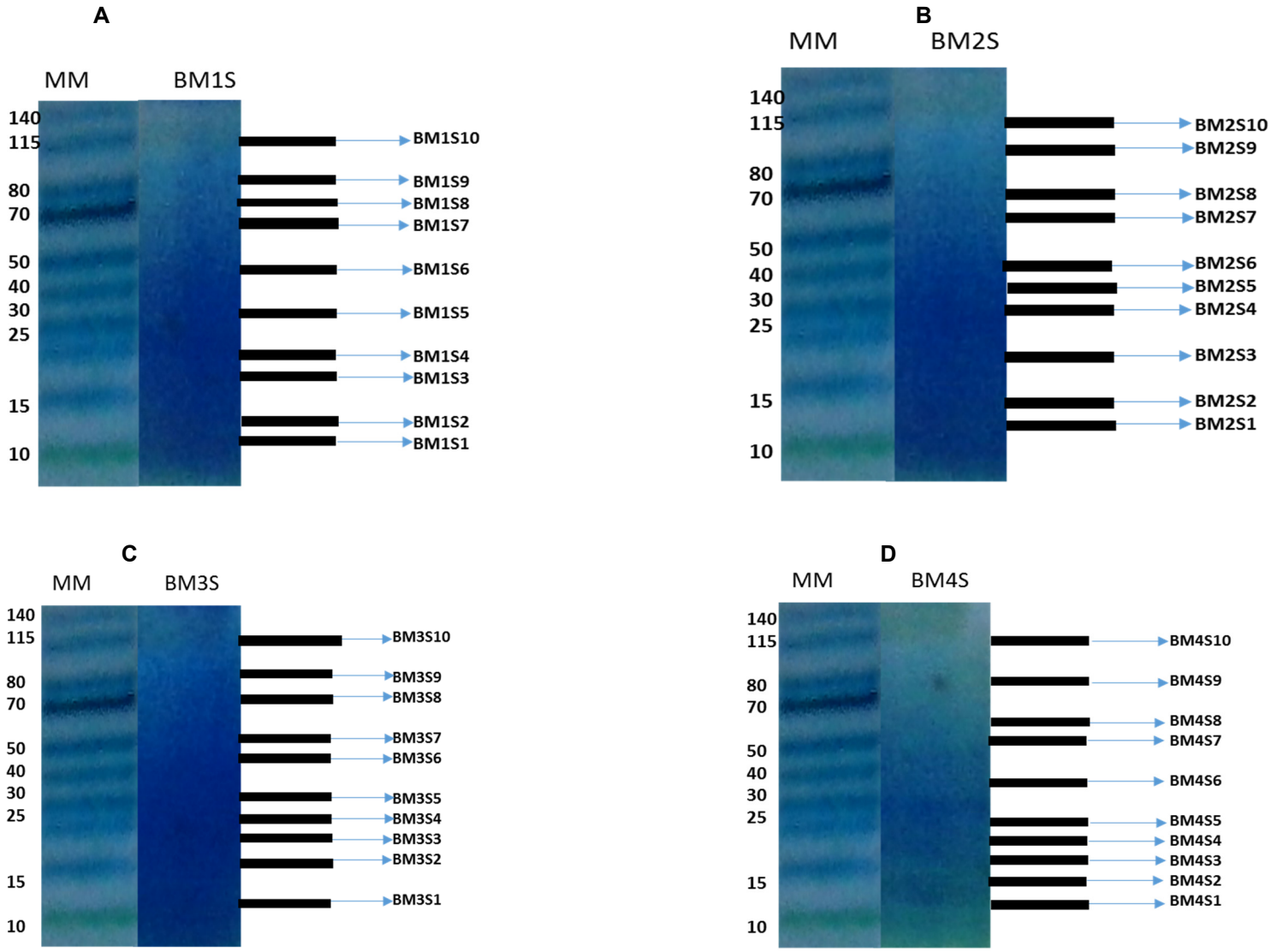

Figure 4.3: An SDS-PAGE separated pattern of *R. (B.) microplus* soluble protein fraction for all feeding stages, BM1 – BM4 stained with Coomassie Brilliant Blue R-250 dye. MM is the protein molecular weight standard. A to D represent BM1 to BM4, respectively, from which 10 gel pieces were cut out for each feeding stage for proteomics studies.

4.2.1. Deep *R. (B.) microplus* salivary glands proteome coverage

The proteomic approach used in this study resulted in deep proteome discovery of proteins from the *R. (B.) microplus* salivary glands. The clusterization of the proteomics data yielded a total number of 5 030 proteins (**Figure 4.4**). Of those 3 661 were identified in the transcriptome, 527 in *Boophilus decolaratus* (*B. decolaratus*) (database obtained from Mr. Philasande Gaven, a fellow MSc colleague working on the same larger project), which a very close related tick to *R. (B.) microplus*. Eight hundred and forty three proteins accounted for proteins from other tick species. There were 446 proteins that were isoforms. It can be argued that these isoforms are the products of the unique proteins obtained from the transcriptome. Interestingly, 4 583 (91.1%) proteins were unique identified in ticks. In that sense, that would mean that they were not isoforms. Furthermore, about 3 215 (63.9%) were only unique in *R. (B.) microplus*. In this sense, the proteome resulted in over 60% coverage, and about 40% could be from other ticks or unknown *R. (B.) microplus* secreted proteins. Moreover, that proves that this is the best *R. (B.) microplus* salivary glands proteome thus far compared to other studies of similar nature that identified fewer tick proteins (Schwarz et al. 2014; Mudenda et al. 2014; Villar et al. 2014a; Gulia-Nuss et al. 2016).

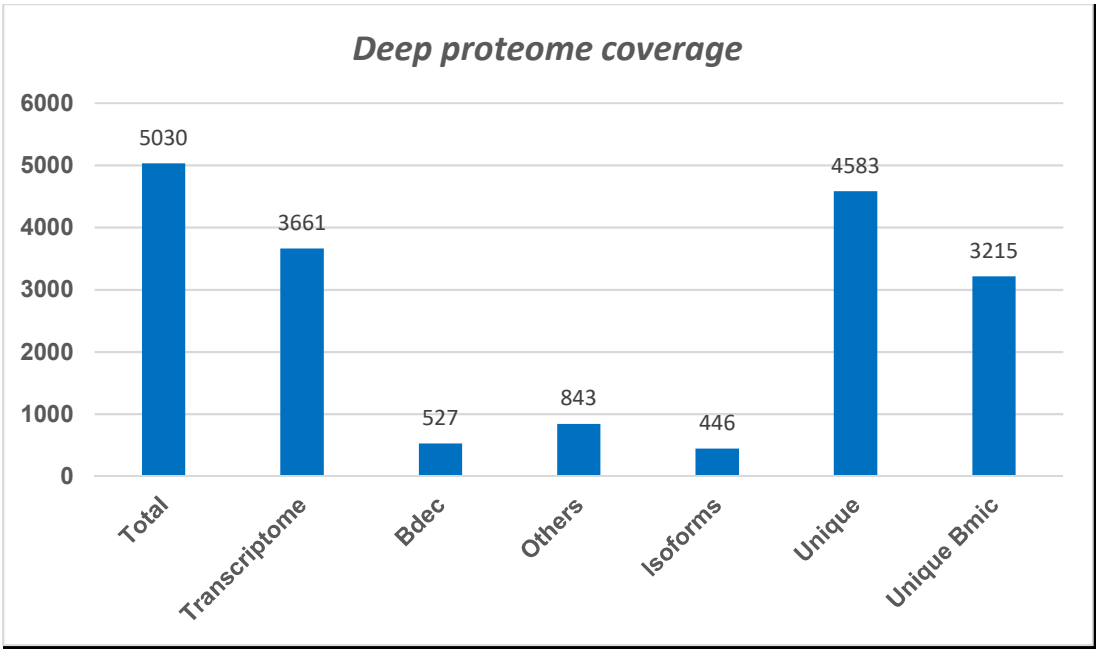

Figure 4.4: *R. (B.) microplus* salivary gland deep proteome coverage. The protein coverage was about 63.9% that was only unique in *R. (B.) microplus*. This coverage was demonstrated to be improved by the proteomic approach applied in this study.

Table 4.1: *R. (B.) microplus* salivary glands proteome protein classification

Protein group	Number of proteins	Percentage (%)
Housekeeping	3062	83.6
Secreted	342	9.3
Unknown	249	6.8
N/A	8	0.3

Unknown: tick proteins with no known or reported function.

N/A: proteins that do not belong to ticks nor have any tick related function.

4.2.1.1. Housekeeping proteins

The housekeeping proteins found in tick salivary glands may have biological importance in tick feeding. The putative housekeeping proteins, as expected, had the highest number of proteins. The housekeeping proteins were expected to be more compared to other tick protein families due to what has been published in the literature of other tick saliva/salivary gland studies. In total, there were 3661 proteins identified in the proteomics data. About 3062 (83.64%) accounted for the housekeeping protein family. Proteins responsible for translation and transport and catabolism showed high levels of expression with 499 and 259, respectively. These were followed by the identification of proteins involved in replication, and amino acid and energy metabolism. This would suggest that the majority of housekeeping proteins are responsible for protein expression and for metabolism. This pattern was also observed in other studies (Ribeiro et al. 2004; Francischetti et al. 2008a, b, 2011; Mans et al. 2008a, b; Karim et al. 2011, Garcia et al. 2014, Kim et al. 2016). These processes could be essential during tick feeding and help in successful blood meal completion (**Figure 4.5**).

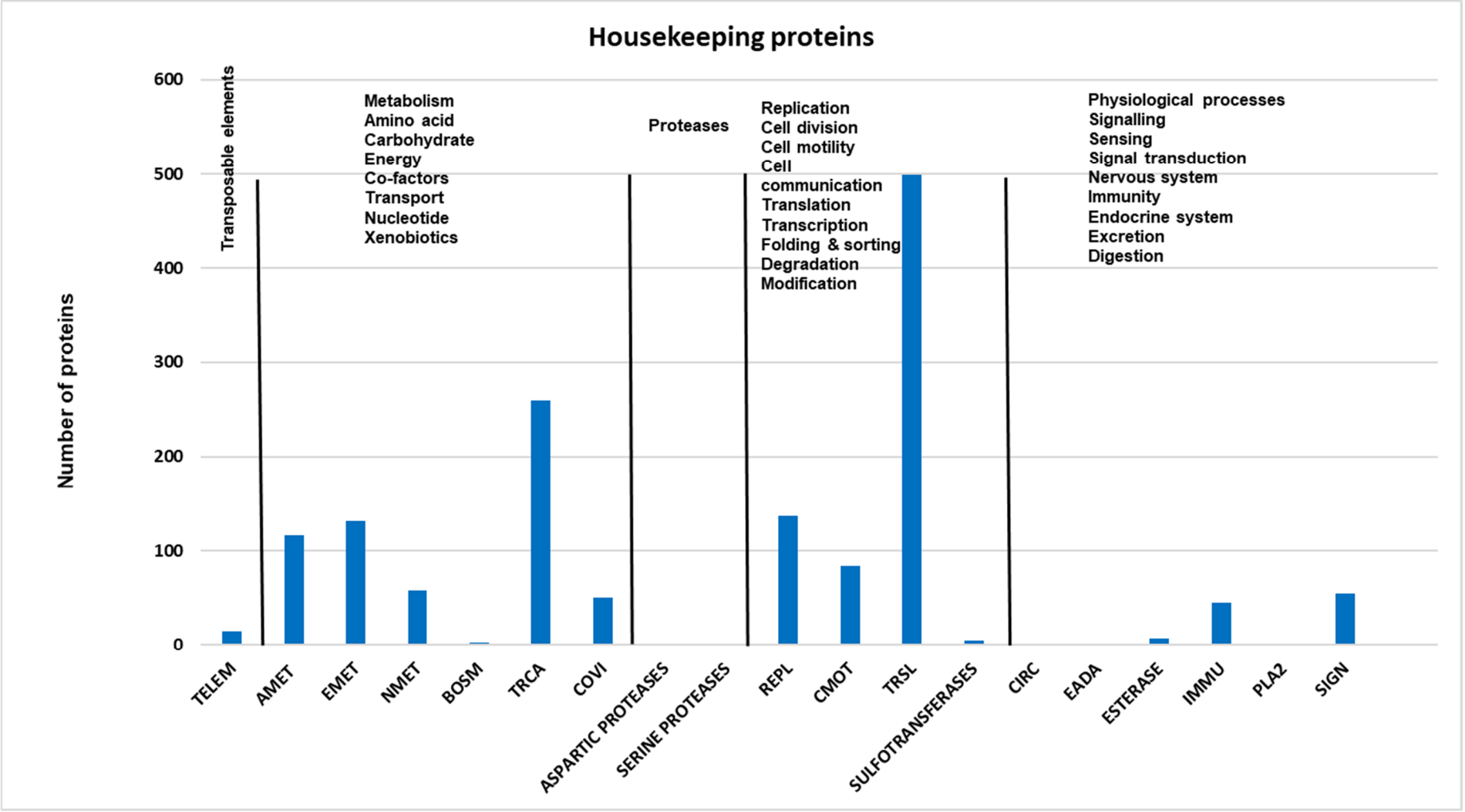

Figure 4.5: Proteomics classification and abundance of the major putative housekeeping proteins from the *R. (B.) microplus* salivary glands proteins based on LC-MS/MS analysis. Proteins associated with FSDE (folding, sorting and degradation) and metabolism were expressed at higher levels relative to other protein families. **Abbreviations:** TELEM (transposable elements), AMET (amino acid metabolism), EMET (energy metabolism), BOSM (co-factors), TRCA (transport and catabolism), CDVI (cell growth and death), REPL (replication), CMOT (cell motility), TRSL (translation), CIRC (circulation), EADA (environmental adaptation), IMMU (immunity), PLA2 (phospholipase A2), REPL (replication), SIGN (signalling).

4.2.1.2. Putative secreted proteins

The analysis of *R. (B.) microplus* salivary gland proteome produced a total number of 342 (9.34%) secreted proteins. The most abundant secretory proteins are listed in **Figure 4.6**. Available salivary gland transcriptomes show that the most commonly and abundant secreted tick salivary gland protein families are lipocalins; metalloproteases; BPTI/Kunitz; glycine-rich family; basic tail family and serpins (Francischetti et al. 2008a,b,2011; Karim et al. 2011; Garcia et al. 2014; Tirloni et al. 2014; de Castro et al. 2016). As it has previously been reported in other studies, lipocalins showed the highest number of proteins (58) followed by metalloproteases (43). The metalloproteases were represented by the gluzincin and reprolysin groups, with the latter accounting for the majority of the proteins. The BPTI showed the least number of proteins with only sixteen proteins. Interestingly, putative secreted proteins with no known functions were also identified in this study, such as Ix_8.9 kDa. These proteins were also described in other tick studies (Francischetti et al. 2011; Garcia et al. 2014), suggesting efficient throughput and thorough analysis of our data. Consequently, strengthen the confidence in our proteome in accordance and relative to other tick sialotranscriptome studies. Moreover, the proteome from this study has identified and reported the highest number of proteins compared to other studies (Valenzuela et al. 2002a; Oleaga et al. 2007; Kim et al. 2011; Tirloni et al. 2014; Radulović et al. 2014; Schwarz et al. 2014). This proves the success of this study, which is owed to the NGS approach coupled with advance bioinformatics analysis.

Table 4.2: Some of the tick proteomics studies identified major protein families

Tick species	Source	Number of proteins	Reference
I. scapularis	Saliva	87	Valenzuela et al. 2002a
O. moubata and *O. erraticus*	Salivary glands	35	Oleaga et al. 2007
R. (B.) microplus	Salivary glands	3661	Genu S, 2017 (MSc dissertation)
R. (B.) microplus	Salivary glands	187	Tirloni et al. 2014
A. americanum	Saliva	895	Radulović et al. 2014
I. ricinus	Salivary glands and midgut	1510	Schwarz et al. 2014
I. scapularis	Saliva	582	Kim et al. 2016

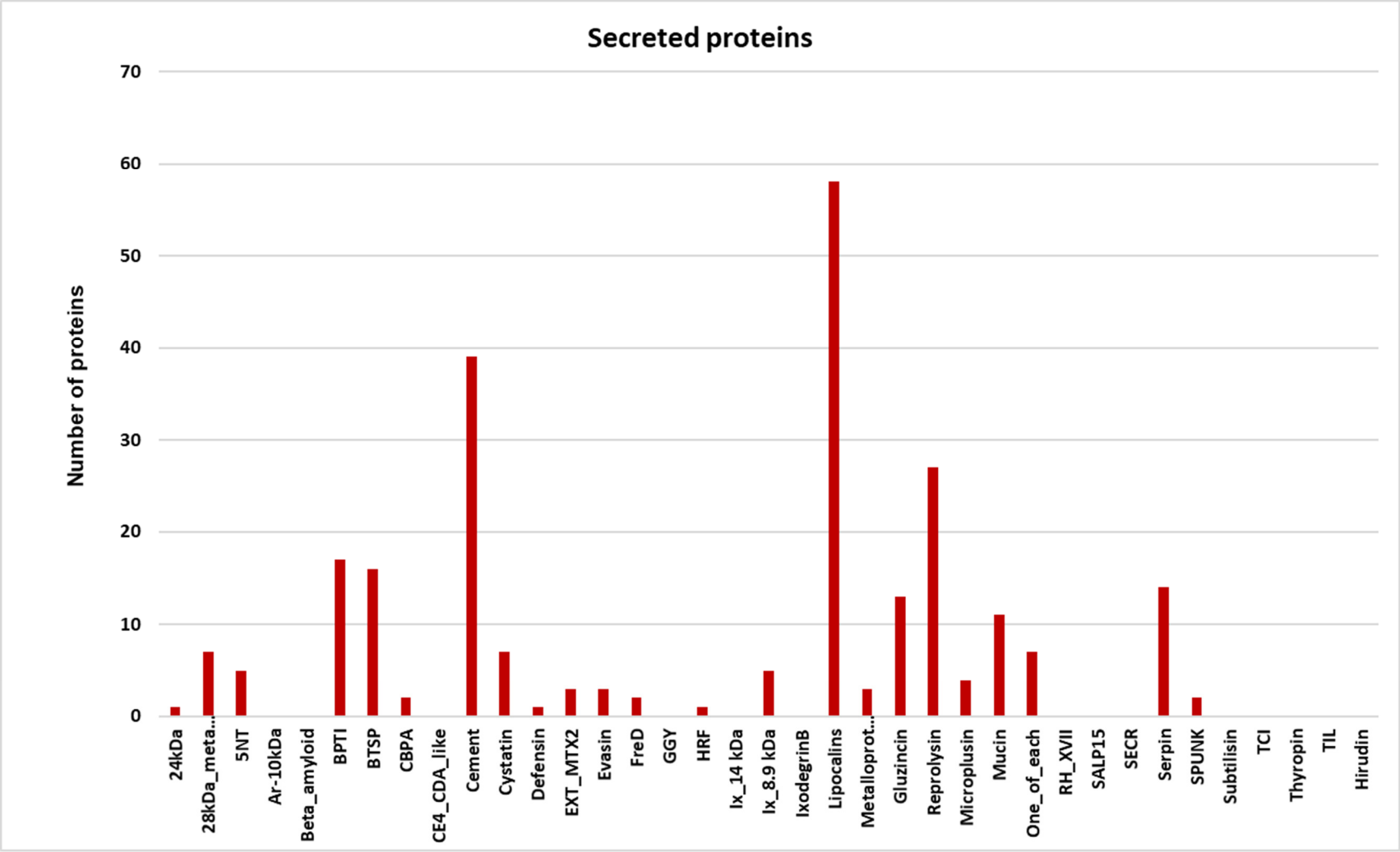

Figure 4.6: Putative secreted proteins identified in *R. (B.) microplus* (ARC-OVR) transcriptome database. That could be attributed the deeper sequencing and coverage obtained by the study design. Lipocalins, cement proteins and metalloproteases protein families were the highest identified. **Abbreviations:** 5NT (5'nucleotidase), AG5 (antigen 5), BPTI (bovine pancreatic trypsin inhibitor), BTSP (basic tail secreted protein), CBPA (choline binding protein A), FreD (fibrinogen-related domain), HRF (histamine release factor), Ix (Ixodegrin), SALP15 (tick salivary protein 15), SECR (tick secreted protein), serpin (serine protease inhibitor), TCI (tick carboxypeptidase inhibitor), TIL (trypsin inhibitor-like cysteine rich domain).

4.3. Discussion

The differences at proteome and transcriptome levels of cells or tissues could be attributed to processes such as cellular location and biological function (Ghazalpour et al. 2011), post-translational modification, half-life and translation efficiency. That information may be relevant and useful in anti-tick vaccine development (Popara et al. 2015). Our data confirm that *R. (B.) microplus* salivary glands contain a complex protein repertoire. There are also remarkable differences in protein composition among different feeding stages of ticks. *R. (B.) microplus* salivary glands are, like their transcriptome, mainly rich in housekeeping proteins and many tick secreted conserved proteins. As it has been previously described, there were abundant tick secreted proteins such as cement (glycine-rich proteins), lipocalins, metalloproteases, and BTSP as well as Kunitz/BPTI proteins.

The tick feeding process brings about some changes in the salivary glands of adult ixodid ticks. Such phenomenon is usually accompanied by significant increases in salivary gland protein profiles (Binnington, 1978; McSwain et al. 1982). Moreover, different feeding conditions determine and affect salivary gland proteins, and a number of studies were in accordance with such observations (Mudenda et al. 2014; Tirloni et al. 2014; Radulović et al. 2014; Mulenga et al. 2007; Schwarz et al. 2013, 2014; Chmelar et al. 2008). As shown in other tick saliva proteomes (Mudenda et al. 2014; Tirloni et al. 2014), protein profiles differ at different tick life stages because different proteins play different unique roles during the tick feeding process.

4.3.1. Housekeeping proteins

Housekeeping proteins are a class of proteins with essential roles in maintaining living and functions organisms. These proteins need to be constantly expressed as long as the organism is alive. In ticks, such proteins include, among others, proteins responsible for metabolism, transcription, signal transduction, protein synthesis, modification and export (Tirloni et al. 2014; Kim et al. 2016). Most studies have previously shown that the majority of housekeeping proteins expressed in all living organisms are responsible for protein synthesis, metabolism, carbohydrate metabolism, amino acid metabolism and the cytoskeleton (Francischetti et al. 2009; Schwarz et al. 2014; Tirloni et al. 2014; Kim et al. 2016). Housekeeping proteins identified in this study include those that play roles in protein synthesis (translation, replication), transport and catabolism, metabolism (energy and amino acid), immunity and transposable elements.

4.3.2. Putative secreted proteins

A number of studies have sought to characterized tick sialotranscriptomes using the proteomics approach. In other studies, proteomics methods are coupled with transcriptomics and the next generation sequencing (NGS). The main qualitative improvement brought by NGS was the ability to detect large numbers of novel transcripts by extensive transcriptomic coverage (Chmelař et al. 2017). Subsequently, NGS has provided a more complete picture of tick gene expression and confirmed the presence of major protein families across tick species. This approach holds a promise to achieve a high throughput and deep transcriptome coverages. This may allow identification of high numbers of novel proteins, especially those that were previously undiscovered by the previous studies. The well-studied tick secreted proteins include lipocalins, metalloproteases, cement proteins, Kunitz domain proteins, basic tail secreted proteins (BTSP) (Ribeiro et al. 2006; Karim et al. 2011; Schwarz et al. 2013; Garcia et al. 2014).

As expected, this study was responsible for the identification of these above-mentioned protein families as described below. Among the most identified secreted proteins were cement proteins (glycine-rich proteins), metalloproteases, lipocalins, BPTI (Kunitz) and BTSP. This trend has also been observed in a number of tick saliva/salivary gland proteome studies (Francischetti et al. 2011; Villar et al. 2014; Tirloni et al. 2014, 2015; Mudenda et al. 2015).

(a) Cement proteins

Proteins from this family have previously been identified in a number of other tick sialomes (Francischetti et al. 2009). Tick proteins containing glycine rich (Gly-rich) and proline-rich (Pro-rich) repeat motifs are known to be associated with cement functions (Bishop et al. 2002; Zhou et al. 2006). *R. (B.) microplus* saliva proteins of this superfamily were found exclusively secreted in early feeding stages (Tirloni et al. 2014). These proteins have also been identified in *O. moubata and R. sanguineus* saliva (Diaz-Martin et al. 2013; Oliveira et al. 2013). Tirloni et al. (2014) found three glycine-rich proteins that contain the [LPAE]-P-G repeats. These repeats are known as targets of proline hydroxylase (Kivirikko et al. 1972) and this characteristic is also present in cement proteins. Recently, cement cones collected in vivo and artificial fed *A. americanum* were found to contain fibrils, beta sheets and helical structures (Bullard et al. 2016). The same characteristics were described in spider silk proteins that contain collagen (Maruyama et al. 2010).

The proteomic analysis of *R. (B.) microplus* salivary gland proteome resulted in identification of 39 cement proteins. Like in other tick studies (Diaz-Martin et al. 2013; Oliveira et al. 2013; Tirloni et al. 2014), it could be suggested that these cement proteins were mainly expressed during the early feeding phase (BM1) of *R. (B.) microplus*. In turn, that proves that the proteomic data produced herein is accordance with what other previous studies discovered. However, this study identified more proteins of this superfamily and that is owed to the NGS and proteomics high-throughputs approach applied in this study. Cement proteins and proteinases identification in early feeding stages of tick feeding as described in our study is in accordance with that found in *R. (B.) microplus* (Tirloni et al. 2014). Cement proteins were detected in early feeding stage as they facilitate tick attachment on the host.

(b) Metalloproteases

Transcripts coding for metalloproteases are frequently described in tick sialotranscriptomes (Harnnoi et al. 2007; Mans et al. 2008a; Francischetti et al. 2009). Metalloproteases form part of a large number of enzymes present in the tick saliva. The roles of these enzymes include assisting blood feeding (Francischetti et al. 2009). Some metalloproteases proteins were identified in saliva of partially fed ticks (Tirloni et al. 2014). In this sense, metallopeptidases often play roles in vascular damage, tissue remodelling and degradation of serum compounds (Nagase and Woessner, 1999). During the early stages of tick blood meal, these proteins may have a role modulating host responses against ticks. As it has been observed in other ticks, these salivary metallopeptidases may be associated with fibrin(ogen)lysis (Francischetti et al. 2003); bradykinin degradation (Bastiani et al. 2002); and angiogenesis inhibition (Francischetti et al. 2005a).

The proteomic dataset from this study allowed for the identification of a total of 44 metallopeptidases majority of which belonged to the reprolysin subclass (27) followed by gluzincin subclass (14). This pattern is in accordance with results from other tick sialotranscriptome and proteome studies (Decrem et al. 2008a; Karim et al. 2011; Tirloni et al. 2014; Kim et al. 2014; Karim and Ribeiro, 2015; Xu et al. 2015; Vora et al. 2017). Moreover, this superfamily is the largest second in protein abundance. Therefore, some of the predicted metalloprotease on the *R. (B.) microplus* salivary glands proteome dataset may be involved in the modulation of host defence mechanisms during tick feeding. The higher number of these proteins could suggest their importance in successful feeding process completion.

(c) Lipocalins

Tick lipocalins are characterised based on the annotated histamine-binding proteins from other tick species. Lipocalins found in tick saliva can play a role in manipulating the host hemostatic functions (Beaufays et al. 2008; Mans et al. 2008a, c). Previously, lipocalins have been identified in several tick saliva proteomes (Diaz-Martin et al. 2013; Tirloni et al. 2014). One of those is the study where 50 lipocalins proteins were

identified in saliva of partially and full fed *R. (B.) microplus* (Tirloni et al. 2014). Our study identified, by proteomics, 58 proteins of the lipocalin family, which was the class with most identified proteins. These proteins were also shown to possess the histamine-binding domain, which strongly suggest that they are tick-secreted proteins. The number of these proteins is similar to those of other lipocalins from other tick saliva/salivary gland proteomes. The presence of high amounts of lipocalins in *R. (B.) microplus* salivary glands is in accordance with data from other tick saliva/salivary glands proteomes, where lipocalins were abundant (Diaz-Martin et al. 2013; Tirloni et al. 2014, 2015; Kim et al. 2016). All these other studies showed less identified proteins of this family, suggesting the best high-throughput of our study. Moreover, the high abundance of lipocalins in tick saliva is in accordance with their antihemostatic and immunomodulatory activities (Francischetti et al. 2009; Mans, 2011; Ribeiro, 1995; Ribeiro and Francischetti, 2003).

(d) Basic protease inhibitor (BPTI)

Proteases facilitate the initial host defense mechanism when ticks feed on them. That mechanism triggers process such as inflammatory, platelet aggregation, blood clotting, complement activation and cellular immunity (Kim et al. 2016). In turn, ticks use protease inhibitors to circumvent the actions of proteases and evade host defense system by protease inhibitors (Mulenga et al. 2000, 2001; Prevot et al. 2007; Maritz-Olivier et al. 2007). The Kunitz-type family members are characterized as inhibitors with a wide range of functionality in various serine endopeptidases (Rawlings et al. 2013). The mechanism by which these inhibitors act is through several molecules, including thrombin, factor Xa, factor XIIa, trypsin and elastase (Maritz-Olivier et al. 2007), suggesting their involvement in tick attachment and feeding. The Kunitz protein family is abundant in saliva of both soft and hard ticks (Chmelař et al. 2012).

Using the proteomics approach, only one and a few Kunitz-domain containing proteins from saliva of fully engorged female *R. (B.) microplus* and nymph, and adult females *H. longicornis*, respectively (Tirloni et al. 2014; 2015). Herein, there were 17 Kunitz proteins identified from the salivary glands of *R. (B.) microplus* using proteomics methods. This number was higher than those from other studies, for instance, studies

by Tirloni and colleagues (Tirloni et al. 2014; 2015), and others (Francischetti et al. 2002a, 2004; Mans et al. 2002a, 2008a), thus, possibly these proteins have role in tick feeding.

(e) Basic tail secreted proteins (BTSP)

The basic tail secreted proteins (BTSP) were amongst the previously known gene families (Francischetti et al. 2008a; 2011). The basic tail superfamily was first found in *I. scapularis* sialome (Valenzuela et al. 2002b; Ribeiro et al. 2006). This superfamily was later to be found to be abundant in other tick sialotranscriptomes (Karim et al. 2011). The first member of basic tail secreted protein superfamily, Salp14 was identified from *I. scapularis*. The Salp14 is an inhibitor of coagulation factor Xa that affects blood coagulation during tick feeding (Narasimhan et al. 2002).

The genes that were probably Salp14 homologs were over-represented in the salivary glands of adult *I. scapularis* female cDNA library 18-24 hours after attachment. These homologs also possessed anticoagulant activity, probably responsible for inhibition of different serine proteases (Ribeiro et al. 2006). Francischetti et al. (2008b) identified six basic tail peptides in their expected gel positions/sizes in salivary glands of *O. coriaceus* by means of proteomics. Later, *H. m. rufipes* sialotranscriptome and proteome revealed salivary secreted components including the basic tail secreted proteins (Francischetti et al. 2011). This strongly confirmed the presence of this superfamily in tick sialomes. This study also reported the identification of sixteen proteins associated with the basic tail functions from the *R. (B.) microplus* salivary gland proteome. Again, the findings of this study are in accordance with those of previously reported studies, thus giving the confidence in the quality and throughput of this study.

4.4. Conclusion

The data obtained herein showed that there was a good validation of the transcriptome using the proteome. This was strengthen by the good/strong correlation between the *R. (B.) microplus* salivary gland transcriptome and proteome in terms of protein families identified. This was observed for both the putative secreted and housekeeping proteins. In conclusion that suggest the success of our study to allow the deep coverage and correlation between tick transcriptome and proteome

Chapter 5: General conclusions and recommendations

5.1. High abundance of cement proteins in the proteome of the *R. (B.) microplus* salivary glands

(i) Tick cement proteins

In order for ticks to successfully obtain their blood meal, they need to be securely attached to their hosts throughout the feeding period (Maruyama et al. 2010). To be successfully attached on one or more hosts depends on factors such as formation of cement cones around the wound on the host skin (Maruyama et al. 2010). These proteins are secreted, within 5-30 minutes of attaching onto the host, to anchor ticks onto host skin (Sonenshine, 1993). Secretion of tick cement is by far one of the most significant biological adaptations that make ticks successful as pests and vectors. Hard ticks, like *R. (B.) microplus*, are long-term blood feeders that remain securely attached onto the host skin for long periods, 4-7 days for larva and nymphs, and 10-14 days for adult ticks with the help of an intact cement plug (Sonenshine, 1993).

The main purpose of cement proteins secreted by ticks is to keep their mouthparts embedded in the host's dermis until feeding is complete (Ullmann-Moore, 2007). The tick cement cone is a layered structure comprising two major types of cement. The first type of cement is produced minutes after tick attachment and hardens rapidly to form a rigid central core of the cone. The second type of cement is secreted later from around 24 hours after attachment and hardens gradually to form a relatively flexible outer cortex. Cement production typically continues until the third or fourth day after attachment (Binnington and Kemp, 1980).

Tick cement proteins are primarily proteinaceous and contain some carbohydrate and lipid (Kemp et al. 1982; Bishop et al. 2002; Murayama et al. 2010). The protein fraction comprised of proteins rich in glycine (glycine-rich proteins, GRPs). These GRPs have been evaluated as anti-tick vaccines (Bishop et al. 2002; Trimnell et al. 2005; Harnnoi et al. 2006; Murayama et al. 2010). The GRPs are naturally abundant and are dominant in proteins that are rich in glycine residues. A number of tick

sialomes have proved that tick saliva is rich these proteins (Francischetti et al. 2009).

(ii) Cement proteins studies on ticks

In general, several studies have been carried out on ticks as part of identifying and testing antigenicity of salivary tick proteins. This class of proteins serve as important factors for both successful tick feeding and as potential candidates as antigens for tick vaccine development. In *R. appendiculatus*, two cement proteins, RIM36 and 64P, were characterized. These proteins have previously been tested as candidate vaccines in cattle, RIM36 (Bishop et al. 2002) and Trp64 (Trimnell et al. 2005). However, the results showed that they failed to control tick infestations (unpublished data, Musoke and Bishop). A 15 kDa cement protein, 64TRP was identified from salivary glands of adult female *R. appendiculatus* (Trimnell et al. 2002; Havlíková et al. 2006). Both RIM36 and 64TRP have a high content of glycine, serine and proline, and contain two types of glycine-rich amino acid repeats, GL[G/Y/S/F/L] tripeptide and GSPLSGF septapeptide (Bishop et al. 2002).

The 64TRP was further expressed as the recombinant cement protein from salivary gland of *R. appendiculatus*. The recombinant form resembles mammalian host skin proteins and responsible for *R. appendiculatus* attachment and feeding (Trimnell et al. 2002; Havlíková et al. 2009). It is a glycine-rich protein with about 59% amino acid identity similarity with a tick cement cone protein (64P) from *R. appendiculatus* (Shapiro et al. 1989; Nuttall et al. 2006; Labuda et al. 2006). The 64TRP was tested and shown to be responsible for the death of engorged ticks when used to vaccine to target the tick-feeding site and cross-reacting with concealed midgut antigens in hamster, guinea pig and rabbit models (Trimnell et al. 2005; Havlíková et al. 2009). Moreover, a 29 kDa extracellular matrix-like salivary gland protein from *H. longicornis*. Rabbits immunized with this cement protein resulted in high mortalities of immature ticks (Norval et al. 1988). Saimo et al. (2011) also expressed and compared antigenic efficacy of two cement proteins, Ra86 and 64TRP. Herein it was showed that 64TRP has higher efficacy than Ra86 when evaluated as vaccine candidates for *R. appendiculatus*.

Our study aimed at characterizing the transcriptome as well as the proteome of *R. (B.) microplus* salivary glands. Like in other hard tick studies, cement proteins were expected to be among the most abundant protein families. Both the transcriptome and proteome revealed these proteins to be expressed in salivary glands and that strongly suggest their importance in tick feeding and host interaction. Their presence during the early feeding stages could be attributed to the fact that they have a crucial role in tick attachment. Our study showed cement protein family to be the one the most abundant protein families after lipocalins, metalloproteases and BPTI. Most tick sialotranscriptome studies have previously shown the same phenomenon (Ribeiro et al. 2006; Karim et al. 2011; Schwarz et al. 2013; Garcia et al. 2014; Tirloni et al. 2014; 2015; Kim et al. 2016). In total there were 146 genes encoded for cement proteins. The majority of genes encoding for the *R. (B.) microplus* salivary transcriptome proteins hits were from our own, *R. (B.) microplus* transcriptome database. That strongly supports deep coverage achieved by Illumina® sequencing and advanced bioinformatics tools. However, a number of tick studies have reported several proteins associated with cement functionality in ticks.

5.2. Transcriptome and proteome correlation

The current study aimed at analysing the *R. (B.) microplus* salivary gland transcriptome and to determine its correlation with the proteome. This study did not show a strong correlation between the transcriptome and proteome in terms of numbers in proteins identified at both levels. The correlation between mRNA and protein abundance depends on various biological and technical factors (Maier et al. 2009). Biological and technical factors being protein half-life or modification, and experimental methods, respectively (Hack, 2004). The mRNA analysis and protein expression data from same cells under similar conditions have failed to show a high correlation (Yeung, 2011; Ghazalpour et al. 2011). Therefore, all these could have contributed in weak or poor correlation between the proteins identified in the *R. (B.) microplus* salivary gland transcriptome and proteome.

Protein families identified as housekeeping proteins were very similar between the transcriptome and proteome. The majority of housekeeping proteins identified in the transcriptome were those with roles in metabolism and cellular processes in ticks. Both the transcriptome and proteome led to the identification of known main tick secretory protein families, such as lipocalins; metalloproteases; Kunitz/BPTI and

cement proteins. All these proteins have been shown to play their different roles in tick biology and tick feeding process. The only difference was in number of identified proteins, where that the transcriptome had higher numbers of proteins relative to the proteome. From these observations, it could be observed that cement proteins were highly expressed compared to other protein families. The same phenomenon was observed in other tick studies (Tirloni et al. 2014; Maruyama et al. 2017). This makes sense, as it is known that ticks need to secrete cement proteins in order to attach on the host during the early feeding phase. Generally, there was a strong correlation between the transcriptome and proteome in terms of protein families identified and their abundance.

5.3. Correlation between this study and other tick studies

It was important to compare findings in this study with those of other related studies. This will help to determine the correlation between this study and its counterparts. A number of studies had attempted to characterize or analyse both the salivary gland transcriptome and proteome of different tick species. The analysis of salivary transcriptome and proteome of *I. ricinus* identified housekeeping proteins involved in signal transduction, cellular protein modification process, and cytoskeleton. The secretory proteins identified in these studies were those that have previously been identified in most of tick salivary glands and in ticks in general. Both the transcriptome and proteome led to the identification of the same secretory proteins as it was observed in our study. These proteins were lipocalins, metalloproteases, Kunitz/BPTI; cement proteins (Schwarz et al. 2014; Cramaro et al. 2015; Tirloni et al. 2015). A study by Francischetti et al (2011) on *H. marginatum rufipes* expressed sequence tags (ESTs) cDNA library led to the identification of housekeeping proteins that are involved in protein synthesis machinery, metabolism/energy, conserved unknown proteins, and signal transduction. Secretory proteins such as Kunitz, lipocalins, cement, mucin and metalloproteases were also identified. This proteins expression profile fitted with that observed in this study.

5.4. Methodology and data analysis shortcomings

5.4.1. Transcriptomics

In this study two experimental approaches were applied, namely transcriptomics and proteomics. This experimental approach have both advantages and limitations/disadvantages. The transcriptomic methodology and data had a very few limitations and/or shortcomings. However, this was not the case when it comes to the proteomics methodology and data analysis.

The main challenge associated with designing experiments and data analysis for tick sialome/salivary gland transcriptome studies is that the NGS technologies typically generates shorter sequences (Wheeler et al. 2008; Bentley et al. 2008), especially in *de novo* sequencing (Hert et al. 2008). This study is one good example of the *de novo* sequencing, whose throughput and coverage could have been greatly improved by the 250bp by 250bp sequencing. However, this sequencing strategy could only cover about 50% of most genes.

5.4.2. Proteomics

The salivary gland proteins were fractionated into three subcellular parts, namely, membrane; soluble and pellet fraction. The 1D SDS-PAGE, subsequently resolved these and each fraction was further excised into ten gel slices (bands). Even though the SDS-PAGE is an inexpensive, powerful and simple way of protein fractionation, it has some drawbacks such as insufficient separation of complex samples; manual preparation; limitations in molecular weights and pI ranges and poor of separation of protein with hydrophobicity (Hellman, 2000; Granvogl et al. 2008; Albright et al. 2009). Some proteins might have been lost during the gel slicing due to their low abundance. The downstream processing such as MS/MS analysis might also pose problems when it comes to analysing the proteome of certain state, tissue or even a cell. There is a high possibility that in LC-MS/MS approaches might fail to detect proteins that are low in abundance (Kim et al. 2016).

To resolve the above-mentioned problems, alternative gel-free MS-based approach could be used to identify differentially expressed proteins. The isotope-coded affinity

tags (ICAT) method (Gygi et al. 1999) is one good example of such methods. These methods demonstrate improved sample recovery, reproducibility, throughput and mass spectrometry friendly. However, gel-free techniques typically require specialized equipment and high level of expertise (Motoyama and Yates, 2008; Chen and Pramanik, 2009). ICAT detects peptides/proteins by targeting their cysteine residues. In this approach, proteins from two populations are differentially expressed and that allows quantitation by MS. This method also offer improvement in protein identification via searching databases due to the rarity of cysteine residues (Hack, 2004).

The data analysis is one vital part of genome sequencing studies. Proper and good data analysis requires high expertise and it can require intensive computational tools. The various software packages can identify different subsets of proteins from the mass spectrometry data. Therefore, those that are identified by multiple software packages would have higher confidence than those that were identified by a single software package (Lippolis and Reinhardt, 2005). Multiple usage of software packages can be costly due to the money that goes to the purchasing and/obtaining of operating licenses. Some licenses do not allow the sharing amongst multiple users. Such limitations pose a problem due to financial implications. Another factor that hamper the proteome analysis is the lack of publicly available sequence information for most of the unsequenced and/or not full sequenced organisms, like *R. (B.) microplus* (Grossmann et al. 2005). However, in this study sequenced the *R. (B.) microplus* transcriptome and that showed a significant improvement in detection of proteins. In a number of non-model organisms' studies, a high number of proteins did not show similarity to known peptides (Hogstrand et al. 2002).

In conclusion, most of the proteins described here, like in many other tick studies, have no known function but, if secreted into their host, they should have anti-hemostatic, anti-inflammatory, anti-angiogenic or immunomodulatory function. They may also contain antimicrobial activity. For future work, potential proteins/antigens need to be selected, expressed, purified and functional characterized. Furthermore, those should be tested for their antigenicity as anti-tick vaccines in clinical trials. That can take a very long time from now, as a lot of intensive research and efforts need to be put in place in order to achieve the successful goal of producing effective tick vaccines in the near future.

References

Aguirre DH, Gaido AB, Vinabal AE, de Echaide ST and Guglielmone AA (1994). Transmission of *Anaplasma marginale* with adult *Boophilus microplus* ticks fed as nymphs on calves with different levels of rickettsaemia. *Parasites* **1**, 405-407.

Alarcon-Chaidez FJ, Sun J and Wikel SK (2007) Construction and characterization of a cDNA library from the salivary glands of *Dermacentor andersoni* stiles (Acari: Ixodidae). *Insect Biochemistry and Molecular Biology* **37**, 48–71.

Albright JC, Dassenko DJ, Mohamed EA and Beussman DJ (2009) Identifying gel separated proteins using in-gel digestion, mass spectrometry, and database searching: Consider the chemistry. *Biochemistry and Molecular Biology Education* **37**, 49–55.

Ali A, Parizi LF, Guizzo MG, Tirloni L, Seixas A, Vaz Ida S Jr and Termignoni C (2015a) Immunoprotective potential of a *Rhipicephalus (Boophilus) microplus* metalloprotease. *Veterinary Parasitology* **207**, 107-14.

Aljamali MN, Ramakrishnan VG, Weng H, Tucker JS, Sauer JR, Essenberg RC (2009) Microarray analysis of gene expression changes in feeding female and male lone star ticks, *Amblyomma americanum* (L). *Archives of Insect Biochemistry and Physiology* **71**, 236–253.

Allen J, Khalil H and Wikel S (1979a) Langerhans cells trap tick salivary antigens in tick-resistant guinea pigs. *Journal of Immunology* **122**, 563-565.

Almazán C, Kocan KM, Blouin EF and de la Fuente J (2005b) Vaccination with recombinant tick antigens for the control of *Ixodes scapularis* adult infestations. *Vaccine* **23**, 5294–5298.

Almazán C, Lagunes R, Villar M, Canales M, Rosario-Cruz R, Jongejan F and de la Fuente J (2010) Identification and characterization of *Rhipicephalus (Boophilus) microplus* candidate protective antigens for the control of cattle tick infestations. *Parasitology Research* **106**, 471–479.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–402.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Anatriello E, Ribeiro JMC, de Miranda-Santos IKF, Brandão LG, Anderson JM, Valenzuela JG, Maruyama SR, Silva JS, Ferreira BR (2010) An insight into the sialotranscriptome of the brown dog tick, *Rhipicephalus sanguineus*. *BMC Genomics* **11**, 1-17.

Andreotti R, Malavazi-Piza KC, Sasaki SD, Torquato RJ, Gomes A and Tanaka AS (2001) Serine proteinase inhibitors from eggs and larvae of tick *Boophilus microplus*: purification and biochemical characterization. *Journal of Protein Chemistry*. **20**, 337–343.

Andreotti R, Pedroso MS, Caetano AR and Martins NF (2008) Comparison of predicted binders in *Rhipicephalus (Boophilus) microplus* intestine protein variants Bm86 Campo Grande strain, Bm86 and Bm95. *Brazilian Journal of Veterinary Parasitology* **17**, 93-98.

Azhahianambi P, de la Fuente J, Suryanarayana VV and Ghosh S (2009) Cloning, expression and immuno protective efficacy of rHaa86, the homologue of the Bm86 tick vaccine antigen, from *Hyalomma anatolicum anatolicum*. *Parasite Immunology* **31**, 111–122.

Balashov YS (1972) Bloodsucking ticks (*Ixodidae*) - vectors of disease of man and animals. *Miscellaneous publications of the Entomological Society of America* **8**, 161–376.

Barker SC and Murrell A (2002) Phylogeny, evolution and historical zoogeography of ticks: a review of recent progress. *Experimental and Applied Acarology* **28**, 55–68.

Barnard AC, Nijhof AM, Gaspar AR, Neitz AW, Jongejan F and Maritz-Olivier C (2012) Expression profiling, gene silencing and transcriptional networking of metzincin metalloproteases in the cattle tick, *Rhipicephalus (Boophilus) microplus*. *Veterinary Parasitology* **186**, 403-414.

Bastiani M, Hillebrand S, Horn F, Kista TBL, Guimarãesa JA and Termignonia C (2002) Cattle tick *Boophilus microplus* salivary gland contains a thiol-activated metalloendopeptidase displaying kininase activity. *Insect Biochemistry and Molecular Biology* **32**, 1439–1446.

Bastos RG, Ueti MW**,** Knowles DP and Scoles GA (2010) *The Rhipicephalus (Boophilus) microplus* Bm86 gene plays a critical role in the fitness of ticks fed on cattle during acute *Babesia bovis* infection. *Parasites and Vectors* **3,** 1-11.

Batista I.C, Ramos OHP, Ventura JS, Junqueira-de-Azevedo ILM, Ho P and Chudzinski-Tavassi AM (2010) A new Factor Xa inhibitor from *Amblyomma cajennense* with a unique domain composition. *Archives of Biochemistry and Biophysics* **493**, 151–156.

Beaufays J, Adam B, Menten-Dedoyart C, Fievez L, Grosjean A, Decrem Y, Prévôt PP, Santini S, Brasseur R, Brossard M, Vanhaeverbeek M, Bureau F, Heinen E, Lins L, Vanhamme L and Godfroid E (2008) Ir-LBP, an *Ixodes ricinus* tick salivary LTB4-binding lipocalin, interferes with host neutrophil function. *PLoS ONE* **3**, e3987.

Bedford GAH (1931*) Nuttalliella namaqua*, a new genus and species of tick. *Parasitology* **23**, 230–232.

Belkaid Y, Kamhawi S, Modi G, Valenzuela J, Noben-Trauth N, Rowton E, Ribeiro J and Sacks D (1998). Development of a natural model of cutaneous *leishmaniasis*: powerful effects of vector saliva and saliva pre-exposure on the long-term outcome

of *Leishmania* major infection in the mouse ear dermis. *Journal of Experimental Medicine* **188,** 1941–1953.

Bellgard MI, Moolhuijzen PM, Guerrero FD, Schibeci D, Rodriguez-Valle M, Peterson DG, Dowd SE, Barrero R, Hunter A, Miller RJ and Lew-Tabor AE (2012) Cattle Tick Base: An integrated Internet-based bioinformatics resource for *Rhipicephalus (Boophilus) microplus*. *International Journal of Parasitology* **42**, 161 - 169.

Ben Said M , Galai Y, Mhadhbi M, Jedidi M , de la Fuente J and Darghouth MA (2012) Molecular characterization of Bm86 gene orthologs from *Hyalomma excavatum, Hyalomma dromedarii* and *Hyalomma marginatum marginatum* and comparison with a vaccine candidate *from Hyalomma scupense*. *Veterinary Parasitology* **190**, 230-240.

Bentley DR (2006) Whole-genome re-sequencing. *Current Opinion in Genetics and Development* **16,** 545–552.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara E Catenazzi M, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R and Smith AJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59.

Bertone P and Snyder M (2005) Prospects and Challenges in Proteomics. *Plant Physiology* **138**, 560–562.

Binnington KC (1978) Sequential changes in salivary gland structure during attachment and feeding of the cattle tick, *Boophilus microplus*. *International Journal of Parasitology* **8**, 97–115.

Binnington KC and Kemp DH (1980) Role of the tick salivary glands in feeding and disease transmission. *Advances in Parasitology* **18**, 316–340.

Bishop R, Lambson B, Wells C, Pandit P, Osaso J, Nkonge C, Morzaria SP, Musoke A and Nene V (2002) A cement protein of the tick *Rhipicephalus appendiculatus*, located in the secretory e cell granules of the type 3 salivary gland acini, induces strong antibody responses in cattle. *International Journal of Parasitology* **32**, 833–842.

Bodzon-Kulakowska A, Bierczynska-Krzysik A, Dylag T, Drabik A, Suder P, Noga M, Jarzebinska J and Silberring J (2007) Methods for samples preparation in proteomic research. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Science* **849**, 1–31.

Bowman AS and Sauer JR (2004) Tick salivary glands: function, physiology and future. *Parasitology* **129**, S67–S81.

Brake DK, Wikel SK, Tidwell JP, Pérez de León AA (2010) *Rhipicephalus microplus* salivary gland molecules induce differential CD 86 expression in murine macrophages. *Parasites and Vectors* **3**, 103-110.

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, Du Bridge RB, Kirchner J, Fearon F, Mao J and Corcoran K (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nature Biotechnology* **18,** 630-634.

Brewer MS, Cotoras DD, Peter Croucher PJP and Gillespie RG (2014) New sequencing technologies, the development of genomics tools, and their applications in evolutionary arachnology. *The Journal of Arachnology* **42**, 1–15.

Brock AR, Wang Y, Berger S, Renkawitz-Pohl R, Han VC, Wu Y and Galko MJ (2012) Transcriptional regulation of profilin during wound closure in *Drosophila* larvae. *Journal of Cell Science* **125**, 5667–5676.

Brossard M and Wikel S (2008) Tick immunobiology. In, *Ticks: Biology, Disease and Control* (ed. Bowman AS and Nuttall PA), pp. 186-204. Cambridge University Press, Cambridge, U.K.

Brossard M and Wikel SK (2004) Tick immunobiology. *Parasitology* **129**, 161-176.

Bullard RL (2016) Characterization of glycine rich proteins from the salivary glands of the lone star tick *Amblyomma americanum* (PhD thesis, *University of Southern Mississippi*, Mississippi). Retrieved from http://aquila.usm.edu/dissertations/339/.

Canales M, Almazán C, Naranjo V, Jongejan F and de la Fuente J (2009a) Vaccination with recombinant *Boophilus annulatus* Bm86 orthologue protein, Ba86, protects cattle against *B. annulatus and B. microplus* infestations. *BMC Biotechnology* **9**, 1-8.

Canales M, Enríquez A, Ramos E, Cabrera D, Dandie H and Soto A (1997) Large-scale production in *Pichia pastoris* of the recombinant vaccine Gavac™ against cattle tick. *Vaccine* **15,** 414–422.

Canales M, Labruna M, Soares J, Prudencio C and de la Fuente J (2009b) Protective efficacy of bacterial membranes containing surface exposed BM95 antigenic peptides for the control of cattle tick infestations. *Vaccine* **27,** 7244–7248.

Canales M, Pérez de la Lastra JM, Naranjo V, Nijhof AM, Hope M, Jongejan F and de la Fuente J (2008). Expression of recombinant *Rhipicephalus (Boophilus) microplus, R. annulatus and R. decoloratus* Bm86 orthologs as secreted proteins in *Pichia patoris*. *BMC Biotechnology* **8,** 1-12.

Carr AL, Mitchell III RD , Dhammi A , Bissinger BW , Sonenshine DE and Roe RM (2017) Tick Haller's Organ, a New Paradigm for Arthropod Olfaction: How Ticks Differ from Insects. *International Journal of Molecular Sciences* **18**, 1-35.

Chen G and Pramanik BN (2009) Application of LC/MS to proteomics studies: current status and future prospects.*Drug Discovery Today* **14**, 465–471.

Chmelař J, Anderson JM, Mu J, Jochim RC, Valenzuela JG and Kopecký J (2008) Insight into the sialome of the castor bean tick, *Ixodes ricinus*. *BMC Genomics* **9**, 233.

Chmelař J, Calvo E, Pedra JH, Francischetti IM and Kotsyfakis M (2012) Tick salivary secretion as a source of antihemostatics. *Journal of Proteomics* **75**, 3842–3854.

Chmelař J, Oliveira CJ, Rezacova P, Francischetti IM, Kovarova Z, Pejler G, Kopacek P, Ribeiro JM, Mares M, Kopecky J and Kotsyfakis M (2011) A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. Blood **117**, 736–744.

Chmelař J, Kotál J, Karim S, Kopacek P, Francischetti IMB, Pedra JHF and Kotsyfakis M (2016) Sialomes and mialomes: a systems biology view of tick tissues and tick-host interactions. *Trends in Parasitology* **32**, 242-254.

Chmelař J, Kotál J, Langhansová H and Kotsyfakis M (2017) Protease Inhibitors in Tick Saliva: The Role of Serpins and Cystatins in Tick-host-Pathogen Interaction. *Frontiers in Cellular and Infection Microbiology* **7**, 216.

Ciprandi A, de Oliveira SK, Masuda A, Horn F and Termignoni C (2006) *Boophilus microplus*: its saliva contains microphilin, a small thrombin inhibitor. *Experimental Parasitology* **114**, 40–46.

Cohen AS, Najarian DR, Paulus A, Guttman A, Smith JA and Karger BL (1988) Rapid separation and purification of oligonucleotides by high-performance capillary gel electrophoresis. *Proceedings of the National Academy of Sciences of the United States of America* **85,** 9660–9663.

Coons LB, Rosell-Davis R and Tarnowski BI (1986) Bloodmeal digestion in ticks, in: Sauer, J.A., Hair, J.A. (Eds.), Morphology, physiology, and behavioral biology of ticks. Ellis Horwood Lmtd, Chichester, pp 248–279.

Cotté V, Sabatier L, Schnell G, Carmi-Leroy A, Rousselle JC, Arsène-Ploetze F, Malandrin L, Sertour N, Namane A and Choumet V (2014) Differential expression of *Ixodes ricinus* salivary gland proteins in the presence of the *Borrelia burgdorferi* sensu lato complex. *Journal of Proteomics* **96**, 29–43.

Cowin AJ (2006) Role of the actin cytoskeleton in wound healing and scar formation. *Primary Intention* **14**, 39-42.

Cramaro WJ, Hunewald OE, Bell-Sakyi L and Muller CP (2017) Genome scaffolding and annotation for the pathogen vector *Ixodes ricinus* by ultra-long single molecule sequencing. *Parasites and Vectors* **10**, 1-9.

Cramaro WJ, Revets D, Hunewald OE, Sinner R, Reye AL and Muller CP (2015) Integration of *Ixodes ricinus* genome sequencing with transcriptome and proteome annotation of the naïve midgut. *BMC Genomics* **16**, 871.

Dalgliesh RJ and Stewart NP (1982) Infectivity of *Anaplasma marginale* extracted from *Boophilus microplus* ticks exposed to certain temperatures. *Australian Veterinary Journal* **58**, 24-26.

de Castro JJ (1997) Sustainable tick and tick-borne disease control in livestock improvement in developing countries. *Veterinary Parasitology* **71,** 77-97.

de Castro MH, de Klerk D, Pienaar R, Latif AA, Rees DJ and Mans BJ (2016) *De novo* assembly and annotation of the salivary gland transcriptome of *Rhipicephalus appendiculatus* male and female ticks during blood feeding. *Ticks and Tick Borne Diseases* **7**, 536–548.

de Graaf DC, Aerts M, Brunain M, Desjardins CA, Jacobs FJ, Werren JH and Devreese B (2010) Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies. *Insect Molecular Biology* **19**, 11–26.

de la Fuente J and Kocan KM (2003) Advances in the identification and characterization of protective antigens for development of recombinant vaccines against tick infestations. *Expert Review of Vaccines* **2,** 583-593.

de La Fuente J and Kocan KM (2006) Strategies for development of vaccines for control of Ixodid tick species. *Parasite Immunology* **28,** 275–283.

de la Fuente J, Kocan KM and Blouin EF (2007b) Tick vaccines and the transmission of tick borne pathogens. *Veterinary Research Communications* **3,** 85–90.

de la Fuente J, Kocan KM, Blouin EF, Zivkovic Z, Naranjo V, Almazán C, Esteves E, Jongejan J, Daffre S and Mangold AJ (2010) Functional genomics and evolution of tick-*Anaplasma* interactions and vaccine development. *Veterinary Parasitology* **167**, 175–186.

de la Fuente J, Rodriguez M, Montero C, Redondo M, García-García JC, and Méndez L (1999) Vaccination against ticks (*Boophilus spp*.): the experience with the Bm86-based vaccine Gavac. *Genetic Analysis* **15,** 143–148.

de la Fuente J, Rodriguez M, Redondo M, Montero C, García-García JC and Méndez L (1998) Field studies and cost-effectiveness analysis of vaccination with Gavac™ against the cattle tick *Boophilus microplus*. *Vaccine* **16,** 366–373.

de Muinck EJ, Trosvik P, Gilfillan GD, Hov JR and Sundaram AYM (2017) A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome* **5**, 1-15.

Decrem Y, Mariller M, Lahaye K, Blasioli V, Beaufays J, Zouaoui Boudjeltia K, Vanhaeverbeek M, Cérutti M, Brossard M, Vanhamme L and Godfroid E (2008b) The impact of gene knock-down and vaccination against salivary metalloproteases on blood feeding and egg laying by *Ixodes ricinus*. *International Journal of Parasitology* **38**, 549–560.

Decrem Y, Beaufays J, Blasioli V, Lahaye K, Brossard M, Vanhamme L and Godfroid E (2008a). A family of putative metalloproteases in the salivary glands of the tick *Ixodes ricinus*. *FEBS Journal* **275**, 1485–1499.

Deracinois B, Flahaut C, Duban-Deweer S and Karamanos Y (2013) Comparative and Quantitative Global Proteomics Approaches: An Overview. *Proteomes* **1***,* 180-218.

Di Venere M, Fumagalli M, Cafiso A, De Marco L, Epis S, Plantard O, Bardoni A, Salvini R, Viglio S, Bazzocchi C, Iadarola P and Sassera D (2015) *Ixodes ricinus* and Its Endosymbiont *Midichloria mitochondrii*: A Comparative Proteomic Analysis of Salivary Glands and Ovaries. *PLoS ONE* **10**, e0138842.

Diaz-Martin V, Manzano-Roman R, Oleaga A, Encinas-Grandes A and Pérez-Sánchez R (2013a) Cloning and characterization of a plasminogen-binding enolase from the saliva of the argasid tick *Ornithodoros moubata*. *Veterinary Parasitology* **191**, 301–314.

Díaz-Martín V, Manzano-Román R, Valero L, Oleaga A, Encinas-Grandes A and Pérez-Sánchez R (2013b) An insight into the proteome of the saliva of the argasid tick *Ornithodoros moubata* reveals important differences in saliva protein composition between the sexes. *Journal of Proteomics* **80**, 216-235.

Doillon CJ, Hembry RM, Ehrlich HP and Burke JF (1987) Actin filaments in normal dermis and during wound healing. *American Journal of Pathology* **126**, 164-170.

Domingos A, Antunes S, Borges L and do Rosário VE (2013) Approaches towards tick and tick-borne diseases control. *Revista da Sociedade Brasileira de Medicina Tropical* **46,** 265-269.

Domon B and Aebersold R (2006) Mass spectrometry and protein analysis. *Science* **312**, 212-217.

Drummond RO (1983) Tick-borne livestock diseases and their vectors. Chemical control of ticks. *World Animal Reviews (FAO)* **36**, 28-33.

Egekwu N, Sonenshine DE, Bissinger BW and Roe RM (2014) Transcriptome of the female synganglion of the black-legged tick *Ixodes scapularis* (Acari: Ixodidae) with comparison between Illumina and 454 systems. *PLoS One* **9**, e102667.

Estrada-Pena A, Bouattour A, Camicas JL, Guglielmone A, Horak IG, Jongejan F, Latif AA, Pegram R and Walker AR (2006) The known distribution and ecological preferences of the tick subgenus *Boophilus* (Acari: *Ixodidae*) in Africa and Latin America. *Experimental and Applied Acarology* **38,** 219–235.

Ferreira CAS, da Silva Vaz I, Sda Silva SS, Haag KL, Valenzuela JG, and Masuda A (2002) Cloning and partial characterization of a *Boophilus microplus* (Acari: *Ixodidae*) calreticulin. *Experimental Parasitology* **101**, 25–34.

Ferreira CAS, Da Silva Vaz Jr I, Da Silva SS, Haag KL, Valenzuela JG and Masuda A (2002) Cloning and partial characterization of a *Boophilus microplus* (Acari*: Ixodidae*) calreticulin. *Experimental Parasitology* **101,** 25–34.

Fontaine A, Diouf I, Bakkali N, Missé D, Pagès F, Fusai T, Rogier C and Almeras L (2011) Implication of haematophagous arthropod salivary proteins in host-vector interactions. *Parasites and Vectors* **4**, 1-17.

Francischetti IM, Anderson JM, Manoukis N, Pham VM and Ribeiro JM (2011) An insight into the sialotranscriptome and proteome of the coarse bontlegged tick, *Hyalomma marginatum rufipes*. *Journal of Proteomics* **74**, 2892–2908.

Francischetti IM, Mans BJ, Meng Z, Gudderra N, Veenstra TD, Pham VM and Ribeiro JM (2008a) An insight into the sialome of the soft tick, *Ornithodorus parkeri. Insect Biochemistry and Molecular Biology* **38**, 1-21.

Francischetti IM, Mather TN and Ribeiro JM (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*. *Biochemical and Biophysical Research Communications* **305**, 869–875.

Francischetti IM, Mather TN and Ribeiro JMC (2005a) Tick saliva is a potent inhibitor of endothelial cell proliferation and angiogenesis. *Thrombosis and. Haemostasis* **94**, 167-74.

Francischetti IM, Pan VM, Mans BJ, Andersen JF, Mather TN, Lane RS and Ribeiro JMC (2005b) The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). *Insect Biochemistry and Molecular Biology* **35**, 1142–1161.

Francischetti IM, Sa-Nunes A, Mans BJ, Santos IM and Ribeiro JM (2009) The role of saliva in tick feeding. *Frontiers Bioscience* **14**, 2051-2088.

Francischetti IMB, Mather TN and Ribeiro JMC (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*. *Thrombosis and Haemostasis* **91**, 886–898.

Francischetti IMB, Mather TN and Ribeiro JMC (2005a) Tick saliva is a potent inhibitor of endothelial cell proliferation and angiogenesis. *Thrombosis and Haemostasis* **94**, 167–174.

Francischetti IMB, Meng Z, Mans BJ, Gudderra N, Hall M, Veenstra TD, Pham VM, Kotsyfakis M and Ribeiro JMC (2008b) An insight into the salivary transcriptome and proteome of the soft tick and vector of epizootic bovine abortion, *Ornithodoros coriaceus*. *Journal of Proteomics* **71**, 493–512.

Francischetti IMB, Pham VM, Mans BJ, Andersen JF, Mather TN, Lane RS and Ribeiro JMC (2005b) The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (*Acari: Ixodidae*). *Insect Biochemistry and Molecular Biology* **35**, 1142–1161.

Francischetti IMB, Valenzuela, JG, Andersen JF, Mather TN and Ribeiro JMC (2002) 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.

Frans J (2000) *Final Report, Integrated Control of Ticks and Tick-Born Diseases (ICTTD)* p. 4.

Freeman JM, Davey RB , Kappmeyer LS, Kammlah DM and Olafson PU (2010) Bm86 midgut protein sequence variation in South Texas cattle fever ticks. *Parasites and Vectors* **3**, 1-8.

Freitas DRJ, Rosa RM, Moraes J, Campos E, Logullo C, Da Silva Vaz I Jr and Masuda A (2007) Relationship between glutathione S-transferase, catalase, oxygen consumption, lipid peroxidation and oxidative stress in eggs and larvae of *Boophilus microplus* (Acarina: Ixodidae). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **146**, 688–694.

Frisch JE (1999) Towards a permanent solution for controlling cattle ticks. *International Journal of Parasitology* **29**, 57–71.

Gaede K and Knülle W (1997) On the mechanism of water vapour sorption from unsaturated atmospheres by ticks. *Journal of Experimental Biology* **200**, 1491-1498.

Garcia G, Gardinassi L, Ribeiro J, Anatriello E, Ferreira B, Moreira HN, Mafra C, Martins M, Szabó MP, de Miranda-Santos IK and Maruyama S (2014) The sialotranscriptome of *Amblyomma triste, Amblyomma parvum* and *Amblyomma cajennense* ticks, uncovered by 454-based RNA-seq. *Parasites & Vectors* **7**, 1-18.

Garcia-Garcia JC, Montero C, Redondo M, Vargas M, Canales M, Boue O, Rodriguez M, Joglar M, Machado H, Gonzalez IL, Valdes M, Mendez L and de la Fuente J (2000) Control of ticks resistant to immunization with Bm86 in cattle

vaccinated with the recombinant antigen Bm95 isolated from the cattle tick, *Boophilus microplus. Vaccine* **18**, 2275-2287.

García-García JC, Montero C, Redondo M, Vargas M, Canales M, Boue O, Rodríguez M, Joglar M, Machado H, González IL, Valdés M, Méndez L and de la Fuente J (2000) Control of ticks resistant to immunization with Bm86 in cattle vaccinated with the recombinant antigen Bm95 isolated from the cattle tick, *Boophilus microplus*. *Vaccine* **18**, 2275-2287.

Gerhard DS, Wagner L, Feingold EA, Shenmen CM, Grouse LH, Schuler G, Klein SL, Old S, Rasooly R, Good P, Guyer M, Peck AM, Derge JG, Lipman D, Collins FS, Jang W, Sherry S, Feolo M, Misquitta L, Lee E, Rotmistrovsky K, Greenhut SF, Schaefer CF, Buetow K, Bonner TI, Haussler D, Kent J, Kiekhaus M, Furey T, Brent M, Prange C, Schreiber K, Shapiro N, Bhat NK, Hopkins RF, Hsie F, Driscoll T, Soares MB, Casavant TL, Scheetz TE, Brown-stein MJ, Usdin TB, Toshiyuki S, Carninci P, Piao Y, Dudekula DB, Ko MS, Kawakami K, Suzuki Y, Sugano S, Gruber CE, Smith MR, Simmons B, Moore T, Waterman R, Johnson SL, Ruan Y, Wei CL, Mathavan S, Gunaratne PH, Wu J, Garcia AM, Hulyk SW, Fuh E, Yuan Y, Sneed A, Kowis C, Hodgson A, Muzny DM, McPherson J, Gibbs RA, Fahey J, Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madari A, Young AC, Wetherby KD, Granite SJ, Kwong PN, Brinkley CP, Pearson RL, Bouffard GG, Blakesly RW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Griffith M, Griffith OL, Krzywinski MI, Liao N, Morin R, Palmquist D, Petrescu AS, Skalska U, Smailus DE, Stott JM, Schnerch A, Schein JE, Jones SJ, Holt RA, Baross A, Marra MA, Clifton S, Makowski KA, Bosak S and Malek J (2004) The status, quality, and expansion of the NIH full-length cDNA project: The Mammalian Gene Collection (MGC). *Genome Research* **14**, 2121–2127.

Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, Park CC, Wen PZ, Brewer H, Karl Weitz, Camp II DG, Pan C, Yordanova R, Neuhaus I, Tilford C, Siemers N, Gargalovic P, Eskin E, Kirchgessner T, Smith DJ, Smith RD and Lusis AJ (2011) Comparative Analysis of Proteome and Transcriptome Variation in Mouse. *PLoS Genetics* **7**, e1001393.

Ghosh S, Azhahianambi P and Yadav MP (2007) Upcoming and future strategies of tick control: A review. *Journal of Vector Borne Diseases* **44,** 79–89.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N and Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**, 644–652.

Graf JF, Gogolewski R, Leach-Bing N, Sabatini GA, Molento MB, Bordin EL and Arantes GJ (2004) Tick control: an industry point of view. *Parasitology* **129,** 427–442.

Graham KM, Sparagano OA and Finn RD (2016) Isolation of the monooxygenase complex from *Rhipicephalus (Boophilus) microplus* - clues to understanding acaricide resistance. *Ticks and Tick Borne Diseases* **7**, 614-623.

Granvogl B, Gruber P and Eichacker LA (2008) Standardisation of rapid in-gel digestion by mass spectrometry. *Proteomics* **7**, 642–654.

Grossmann J, Roos FF, Cieliebak M, Liptak Z, Mathis LK, Muller M, Gruissem W and Baginsky S (2005) AUDENS: A tool for automated peptide de novo sequencing. *Journal Proteome Research* **4**, 1768–1774.

Guerrero FD, Andreotti R, Bendele KG, Cunha RC, Miller RJ, Yeater K and de Leõn AAP (2014) *Rhipicephalus (Boophilus) microplus* aquaporin as an effective vaccine antigen to protect against cattle tick infestations. *Parasites and Vectors* **7**, 1-12.

Guerrero FD, Miller RJ and Pérezde León AA (2012) Cattle tick vaccines: many candidate antigens, but will a commercially viable product emerge? *International Journal of Parasitology* **42,** 421–427.

Guerrero FD, Miller RJ, Rousseau ME, Sunkara S, Quackenbush J, Lee Y and Nene V (2005) BmiGI: a database of cDNAs expressed in *Boophilus microplus*, the tropical/southern cattle tick. *Insect Biochemistry and Molecular Biology* **35,** 585-595.

Guerrero FD, Nene VM, George JE, Barker SC, and Willadsen P (2006) Sequencing a New Target Genome: The *Boophilus microplus* (Acari: *Ixodidae*) Genome Project. *Journal of Medical Entomology* **43,** 9-16.

Guglielmone AA, Robbins RG, Apanaskevich DA, Petney TN, Estrada-Pena A, Horak IG, Shao R and Barker SC (2010) The *Argasidae, Ixodidae* and *Nuttalliellidae* (Acari: Ixodida) of the world: a list of valid species names. *Zootaxa* **2528**, 1–28.

Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, Sattelle DB, de la Fuente J, Ribeiro JM, Megy K, Thimmapuram J, Miller JR, Walenz BP, Koren S, Hostetler JB, Thiagarajan M, Joardar VS, Hannick LI, Bidwell S, Hammond MP, Young S, Zeng Q, Abrudan JL, Almeida FC, Ayllón N, Bhide K, Bissinger BW, Bonzon-Kulichenko E, Buckingham SD, Caffrey DR, Caimano MJ, Croset V, Driscoll T, Gilbert D, Gillespie JJ, Giraldo-Calderón GI, Grabowski JM, Jiang D, Khalil SM, Kim D, Kocan KM, Koči J, Kuhn RJ, Kurtti TJ, Lees K, Lang EG, Kennedy RC, Kwon H, Perera R, Qi Y, Radolf JD, Sakamoto JM, Sánchez-Gracia A, Severo MS, Silverman N, Šimo L, Tojo M, Tornador C, Van Zee JP, Vázquez J, Vieira FG, Villar M, Wespiser AR, Yang Y, Zhu J, Arensburger P, Pietrantonio PV, Barker SC, Shao R, Zdobnov EM, Hauser F, Grimmelikhuijzen CJ, Park Y, Rozas J, Benton R, Pedra JH, Nelson DR, Unger MF, Tubio JM, Tu Z, Robertson HM, Shumway M, Sutton G, Wortman JR, Lawson D, Wikel SK, Nene VM, Fraser CM, Collins FH, Birren B, Nelson KE, Caler E and Hill CA (2016) Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nature Communications* **7**, 10507.

Gupta AK and Gupta UD (2014) *Next Generation Sequencing and Its Applications*, In: Verma, A. and Singh, A. (eds), Animal Biotechnology: Models In Discovery and Translation, Elsevier Inc., 225 Wyman Street, Waltham, MA 02541, USA, pp. 345-367.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH and Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags". *Nature Biotechnology* **17**, 994–999.

Hack CJ (2004) Integrated transcriptome and proteome data: The challenges ahead. *Briefings in Functional Genomics and Proteomics* **3**, 212–219.

Hajnická V, Vancová-Štibrániová I, Slovák M, Kocáková P and Nuttall PA (2011) Ixodid tick salivary gland products target host wound healing growth factors. *International Journal of Parasitology* **41**, 213–223.

Harnnoi T, Sakaguchi T, Nishikawa Y, Xuan X, Fujisaki K (2007) Molecular characterization and comparative study of 6 salivary gland metalloproteases from the hard tick, *Haemaphysalis longicornis*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **147**, 93–101.

Havlíková S, Roller L, Koči J, Trimnell AR, Kazimírová M, Klempa B and Nuttall PA (2009) Functional role of 64P, the candidate transmission-blocking vaccine antigen from the tick, *Rhipicephalus appendiculatus*. *International Journal of Parasitology* **13**, 1485-94.

Hedin M, Starrett J, Akhter S, Schönhofer AL and Shultz JW (2012) Phylogenomic Resolution of Paleozoic Divergences in Harvestmen (Arachnida, Opiliones) via Analysis of Next-Generation Transcriptome Data. *PLoS ONE* **7,** 1-13.

Hellman U (2000) Sample preparation by SDS/PAGE and in-gel digestion. *Experientia Supplementum* **88**, 43–54.

Hert DG, Fredlake CP and Barron AE (2008) Advantages and limitations of next-generation sequencing technologies: A comparison of electrophoresis and non-electrophoresis methods. *Nucleic Acids* **29**, 4618–4626.

Ho S, Marcal H and Foster LJ (2014) Towards scarless wound healing: A comparison of protein expression between human, adult and foetal fibroblasts. *BioMed Research International* 676493.

Hogstrand C, Balesaria S and Glover CN (2002) Application of genomics and proteomics for study of the integrated response to zinc exposure in a non-model fish species, the rainbow trout. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **133**, 523–535.

Horak IG, Camicas JL and Keirans JE (2002) The *Argasidae, Ixodidae* and *Nuttalliellidae* (Acari: I*xodida*): a world list of valid tick names. *Experimental and Applied Acarology* **28,** 27–54.

Horn F, dos Santos PC and Termignoni C (2000) *Boophilus microplus* anticoagulant protein: an antithrombin inhibitor isolated from the cattle tick saliva. *Archives of Biochemistry and Biophysics* **384**, 68–73.

Hoy MA, Waterhouse RM, Wu K , Estep AS, Ioannidis P, Palmer WJ, Pomerantz AF, Simaõ FA, Thomas J, Jiggins FM, Murphy TD, Pritham EJ, Robertson HM, Zdobnov EM, Gibbs RA and Richards S (2016) Genome Sequencing of the

Phytoseiid Predatory Mite *Metaseiulus occidentalis* Reveals Completely Atomized Hox Genes and Super dynamic Intron Evolution. *Genome Biology and Evolution* **8**, 1762–1775.

Hoy MA, Waterhouse RM, Wu K, Estep AS, Ioannidis P, Palmer WJ, Pomerantz AF, Simão FA, Thomas J, Jiggins FM, Murphy TD, Pritham EJ, Robertson HM, Zdobnov EM, Gibbs RA and Richards S (2016) Genome sequencing of the phytoseiid predatory mite *Metaseiulus occidentalis* reveals completely atomized hox genes and superdynamic intron evolution. *Genome Biology and Evolution* **8**, 1762–1775.

Hutchison CA (2007) DNA sequencing: bench to bedside and beyond. *Nucleic Acids Research* **35**, 6227–6237.

Ibelli AM, Kim TK, Hill CC, Lewis LA, Bakshi M, Miller S, Porter L and Mulenga A (2014) A blood meal-induced *Ixodes scapularis* tick saliva serpin inhibits trypsin and thrombin, and interferes with platelet aggregation and blood clotting. *International Journal of Parasitology* **44**, 369–79.

Imamura S, da Silva Vaz I Jr, Konnai S, Yamada S, Nakajima C, Onuma M and Ohashi K (2009) Effect of vaccination with a recombinant metalloprotease from *Haemaphysalis longicornis*. *Experimental and Applied Acarology* **48**, 345-358.

Imamura S, Da Silva Vaz Junior I, Sugino M, Ohashi K, Onuma M (2005) A serine protease inhibitor (serpin) from *Haemaphysalis longicornis* as an anti-tick vaccine. *Vaccine* **23**, 1301–1311.

Imamura S, Konnai S, Vaz Ida S, Yamada S, Nakajima C, Ito Y, Tajima T, Yasuda J, Simuunza M, Onuma M and Ohashi K (2008) Effects of anti-tick cocktail vaccine against *Rhipicephalus appendiculatus*. *Japanese Journal of Veterinary Research* **56**, 85–98.

Jaworski DC, Simmen FA, Lamoreaux W, Coons LB, Muller MT and Needham GR (1995) A secreted calreticulin protein in Ixodid tick (*Amblyomma americanum*) saliva. *Journal of Insect Physiology* **41,** 369–375.

Jeyabal L, Azhahianambi P, Susitha K, Ray DD, Chaudhuri P and Vanlahmuaka (2010). Efficacy of rHaa86, an orthologue of Bm86, against challenge infestations *of Hyalomma anatolicum anatolicum. Transboundary and Emerging Diseases* **57,** 96–102.

Jongejan F and Uilenberg G (2004) The global importance of ticks. *Parasitology* **129,** 3-14.

Jutel M and Akdis CA (2001b) Immunological mechanisms of allergen-specific immunotherapy. *Allergy* **66**, 725–732.

Käll L, Krogh A and Sonnhammer EL (2004) A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology* **338**, 1027-1036.

Kanehisa M, Sato Y, Kawashima M, Furumichi M and Tanaba M (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44**, D457-D462.

Karim S and Ribeiro JM (2015) An Insight into the Sialome of the Lone Star Tick, *Amblyomma americanum*, with a glimpse on its time dependent gene expression. *PLoS ONE* **10**, e0131292.

Karim S, Singh P and Ribeiro JM (2011) A deep insight into the sialotranscriptome of the gulf coast tick, *Amblyomma maculatum*. *PLoS One* **6**, e28525.

Kaufman W (1976) The influence of various factors on fluid secretion by in vitro salivary glands of ixodid ticks. *Journal of Experimental Biology* **64**, 727-742.

Kaufman WR (1983) The function of tick salivary glands. In: K.F. Harris (Editor), Current Topics in Vector Research, Vol. 1. Praeger Scientific, New York, pp. 215–247.

Kaufman WR and Phillips JE (1973a) Ion and water-balance in ixodid tick *Dermacentor andersoni*. I. Routes of ion and water excretion. *Journal of Experimental Biology* **58**, 523-536.

Kaufman WR and Phillips JE (1973b) Ion and water-balance in ixodid tick *Dermacentor andersoni*. II. Mechanism and control of salivary secretion. *Journal of Experimental Biology* **58**, 537-547.

Kaufman WR and Wong DL (1983) Evidence for multiple receptors mediating fluid secretion in salivary glands of ticks. *European Journal of Pharmacology* **87**, 43–52.

Kaufman WR, Sloley BD, Tatchell RJ, Zbitnew GL, Diefenbach TJ and Goldberg JI (1999) Quantification and cellular localization of dopamine in the salivary gland of the ixodid tick *Amblyomma hebraeum*. *Experimental and Applied Acarology* **23**, 251–265.

Kazimírová M (2008) Pharmacologically active compounds in tick salivary glands. *Monographs* **12,** 281-296.

Kazimírová M and Štibrániová I (2013) Tick salivary compounds: their role in modulation of host defences and pathogen transmission. *Frontiers in Cellular and Infection Microbiology* **3**, 1-19.

Kellner R (2000) Proteomics. Concepts and perspectives. *Fresenius Journal of Analytical Chemistry* **366**, 517-524.

Kemp DH, Stone BF and Binnington KC: Tick attachment and feeding: Role of mouthparts, feeding apparatus, salivary gland secretions and the host response. Physiology of ticks. Edited by: Obenchain FD, Galun R. 1982, *Pergamon Press*, Ltd, Oxford, 119-168.

Kemp DH and Bourne A (1980) *Boophilus microplus*: the effect of histamine on the attachment of cattle-larvae studies *in vivo* and *in vitro. Journal of Parasitology* **80**, 487-496.

Kim TK, Tirloni L, Pinto AFM, Moresco J, Yates JR III, da Silva Vaz I Jr and Mulenga A (2016) *Ixodes scapularis* Tick Saliva Proteins Sequentially Secreted Every 24 h during Blood Feeding. *PLoS Neglected Tropical Diseases* **10**, e0004323.

Kim YH, Slam MS and You MJ (2015) Proteomic screening of antigenic proteins from the hard tick, *Haemaphysalis longicornis* (Acari: Ixodidae). *Korean Journal of Parasitology* **53**, 85-93.

Kivirikko KI, Kishida Y, Sakakibara S and Prockop DJ (1972) Hydroxylation of (X–Pro–Gly)n by protocollagen proline hydroxylase. Effect of chain length, helical conformation and amino acid sequence in the substrate. *Biochimica et Biophysica Acta* **271**, 347–356.

Knowles DP, Lowell DP, Kappmeyer S, Stiller D, Hennager SG and Perryman LE (1992) Antibody to a recombinant merozoite protein epitope identifies horses infected with *Babesia equi*. *Journal of Clinical Microbiology* **30**, 3122-3126.

Koči J, Šimo L and Park Y (2014) Autocrine/paracrine dopamine in the salivary glands of the blacklegged tick *Ixodes scapularis*. *Journal of Insect Physiology* **62**, 39–45.

Kongsuwan K, Josh P, Zhu Y, Pearson R, Gough J and Colgrave ML (2010) Exploring the midgut proteome of partially fed female cattle tick (*Rhipicephalus* (*Boophilus*) *microplus*). *Journal of Insect Physiology* **56**, 212–226.

Kotsyfakis M, Karim S, Andersen JF, Mather TN and Ribeiro JM (2007) Selective cysteine protease inhibition contributes to blood-feeding success of the tick Ixodes scapularis. *Journal of Biological Chemistry* **282**, 29256–29263.

Kotsyfakis M, Kopacek P, Franta Z, Pedra JH, Ribeiro JM (2015b) Deep Sequencing Analysis of the *Ixodes ricinus* Haemocytome. *PLOS Neglected Tropical Diseases* 9:e0003754.

Kotsyfakis M, Kopáček P, Franta Z, Pedra JHF and Ribeiro JMC (2015b) Deep Sequencing Analysis of the *Ixodes ricinus* Haemocytome. *PLoS Neglected Tropical Diseases* **9**, e000375.

Kotsyfakis M, Schwarz A, Erhart J and Ribeiro JM (2015a) Tissue- and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Scientific Reports* **5**, 9103.

Kotsyfakis M, Schwarz A, Erhart J and Ribeiro JMC (2015a) Tissue- and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Scientific Reports* **5**, 9103.

Krasnov BR, Khokhlova IS, Shenbrot GI and Poulin R (2008) How are the host spectra of hematophagous parasites shaped over evolutionary time? Random choice vs selection of a phylogenetic lineage. *Parasitology Research* **102**, 1157-1164.

Krogh A, Larsson B, von Heijne G and Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567-580.

Labuda M, Trimnell AR, Ličková M, Kazimírová M, Davies GM, Lissina O, Hails RS and Nuttall PA (2006) An Antivector Vaccine Protects against a Lethal Vector-Borne Pathogen. *PLoS Pathogens* **2**, e27.

Lambertz C, Chongkasikit N, Jittapalapong S and Gauly M (2012) Immune Response of *Bos indicus* Cattle against the Anti-Tick Antigen Bm91 Derived from Local *Rhipicephalus (Boophilus) microplus* Ticks and Its Effect on Tick Reproduction under Natural Infestation. *Journal of Parasitology Research Article* **2012**, 1-9.

Lang F, Föller M, Lang KS, Lang PA, Ritter M, Gulbins E, Vereninov A and Huber SM (2005). Ion channels in cell proliferation and apoptotic cell death. *Journal of Membrane Biology* **205**, 147–157.

Larregina AT and Falo LD (2005) Changing paradigms in cutaneous immunology: adapting with dendritic cells. *Journal of Investigative Dermatology* **124,** 1–12.

Latif AA and Walker AR (2004) An introduction to the biology and control of ticks in Africa. *ICTTD-2 Project*, page 1-29.

Latif AA, Putterill JF, de Klerk DG, Pienaar R and Mans BJ (2012) *Nuttalliella namaqua* (*Ixodoidea: Nuttalliellidae*): First Description of the Male, Immature Stages and Re-Description of the Female. *PLoS ONE* **7**, 1-9.

Leboulle G, Crippa M, Decrem Y, Mejri N, Brossard M, Bollen A and Godfroid E (2002). Characterization of a novel salivary immunosuppressive protein from *Ixodes ricinus* ticks. *Journal of Biological Chemistry* **277**, 10083–10089.

Lee Y, Tsai J, Sunkara S, Karamycheva S, Pertea G, Sultana R, Antonescu V, Chan A, Cheung F and Quackenbush J (2005) The TIGR Gene Indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. *Nucleic Acids Research* **33**, D71–D74,

Leinonen R, Sugawara H and Shumway M (2011) The Sequence Read Archive. *Nucleic Acids Res* **39**, D19–D21.

Levy SE and Myers RM (2016) Advancements in Next-Generation Sequencing. *Annual Review of Genomics and Human Genetics* **17**, 95-115.

Li W and Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659.

Lima CA, Torquato RJS, Sasaki SD, Justo GZ and Tanaka AS (2010) Biochemical characterization of a Kunitz-type inhibitor similar to dendrotoxins produced by *Rhipicephalus (Boophilus) microplus* (Acari: Ixodidae) hemocytes. Veterinary Parasitology **167**, 279–287.

Lindsay PJ and Kaufman WR (1986) Potentiation of salivary fluid secretion in ixodid ticks: a new receptor system for gamma-aminobutyric acid. *Canadian Journal of Physiology and Pharmacology* **64**, 1119–1126.

Lippolis JD and Reinhardt TA (2005) Proteomic survey of bovine neutrophils. *Veterinary Immunology and Immunopathology* **103**, 53–65.

Liu X and Bonnet S (2014) Hard tick factors implicated in pathogen transmission. *PLoS Neglected Tropical Diseases* **8**, e2566.

Liyou N, Hamilton S, Elvin, C and Willadsen P (1999) Cloning and expression of ecto 5-nucleotidase from the cattle tick *Boophilus microplus*. *Insect Molecular Biology* **8**, 257–266.

Louw E, van der Merwe NA, Neitz AW and Maritz-Olivier C (2013) Evolution of the tissue factor pathway inhibitor-like Kunitz domain-containing protein family in *Rhipicephalus microplus*. *International Journal of Parasitology* **43**, 81-94.

Lundberg E, Fagerberg L, Klevebring D, Matic I, Geiger T, Cox J, Algenas C, Lundeberg J, Mann M and Uhlen M (2010) Defining the transcriptome and proteome in three functionally different human cell lines. *Molecular Systems Biology* **6**, 1-10.

Lynen G, Zeman P, Bakuname C, Di Giulio G, Mtui P, Sanka P and Jongejan F (2008) Shifts in the distributional ranges of *Boophilus* ticks in Tanzania: evidence that a parapatric boundary between *Boophilus microplus* and *Boophilus decoloratus* follows climate gradients. *Experimental and Applied Acarology* **44,** 147-64.

Macedo-Ribeiro S, Almeida C, Calisto BM, Friedrich T, Mentele R, Stürzebecher J, Fuentes-Prior P and Pereira PJB (2008) Isolation, Cloning and Structural Characterisation of Boophilin, a Multifunctional Kunitz-Type Proteinase Inhibitor from the Cattle Tick. *PLoS One*, **3**, e1624.

Madden RD, Sauer JR and Dillwith JW (2004) A proteomics approach to characterizing tick salivary secretions. *Experimental and Applied Acarology* **32**, 77-87.

Madder M, Adehan S, De Deken R, Adehan R and Lokossou R (2012a). New foci of *Rhipicephalus microplus* in West Africa. *Experimental and Applied Acarology* **56**, 385–390.

Madder M, Thys E, Achi L, Toure´ A and De Deken R (2011) *Rhipicephalus (Boophilus) microplus*: a most successful invasive tick species in West-Africa. *Experimental and Applied Acarology* **53,** 139–145.

Maier T, Güell M and Serrano L (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Letters* **583**, 3966-3973.

Mane SD (1985) Cyclic AMP-depended protein kinase from the salivary glands of the tick, *Amblyomma americanum* - partial purification and properties. *Insect Biochemistry* **15**, 777–787.

Mane SD, Sauer JR and Essenberg RC (1988) Molecular forms and free cAMP receptors of the cAMP dependent protein kinase from tick salivary glands. *Comparative Biochemistry and Physiology* **91B**, 117-124.

Mans BJ and Neitz AWH (2004) Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochemistry and Molecular Biology* **34**, 1–17.

Mans BJ and Ribeiro JM (2008a) A novel clade of cysteinyl leukotriene scavengers in soft ticks. *Insect Biochemistry and Molecular Biology* **38**, 862-870.

Mans BJ and Ribeiro JM (2008b) Function, mechanism and evolution of the moubatin-clade of soft tick lipocalins. *Insect Biochemistry and Molecular Biology* **38**, 841-852.

Mans BJ, Andersen JF, Francischetti IM, Valenzuela JG, Schwan TG, Pham VM, Garfield MK, Hammer CH and Ribeiro JM (2008b) Comparative sialomics between hard and soft ticks: implications for the evolution of blood-feeding behavior. *Insect Biochemistry and Molecular Biology* **38**, 42-58.

Mans BJ, Andersen JF, Francischetti IMB, Valenzuela JG, Schwan TG, Pham VM, Garfield MK, Hammer CH and Ribeiro JMC (2008a) Comparative sialomics between hard and soft ticks: implications for the evolution of blood-feeding.

Mans BJ, Andersen JF, Schwan TG and Ribeiro JM (2008a) Characterization of anti-hemostatic factors in the argasid, *Argas monolakensis*: implications for the evolution of blood-feeding in the soft tick family. *Insect Biochemistry and Molecular Biology* **38**, 22- 41.

Mans BJ, de Castro MH, Pienaar R, de Klerk D, Gaven P, Genus S and Latif AA (2016) Ancestral reconstruction of tick lineages. *Ticks and Tick Borne Diseases* **7**, 509-535.

Mans BJ, de Klerk D, Pienaar R, de Castro MH and Latif AA (2012) The mitochondrial genomes of *Nuttalliella namaqua* (Ixodoidea: *Nuttalliellidae*) and *Argas africolumbae* (Ixodoidae: *Argasidae*): estimation of divergence dates for the major tick lineages and reconstruction of ancestral blood-feeding characters. *PLoS One* **7**, e49461.

Mans BJ, de Klerk D, Pienaar R, Latif AA (2011) *Nuttalliella namaqua*: a living fossil and closest relative to the ancestral tick lineage: implications for the evolution of blood-feeding in ticks. *PLoS One* **6**, 1-11.

Mans BJ, de Klerk DG, Pienaar R and Latif AA (2014) The host preferences of *Nuttalliella namaqua* (Ixodoidea: *Nuttalliellidae*): a generalist approach to surviving multiple host-switches. *Experimental and Applied Acarology* **62**, 233-240.

Mans BJ, Louw AI and Neitz AW (2002b). Amino acid sequence and structure modeling of savignin, a thrombin inhibitor from the tick, *Ornithodoros savignyi*. *Insect Biochemistry and Molecular Biology* **32**, 821–828.

Mans BJ, Louw AI and Neitz AW (2003) The major tick salivary gland proteins and toxins from the soft tick, *Ornithodoros savignyi*, are part of the tick Lipocalin family: implications for the origins of tick toxicoses. *Molecular Biology and Evolution* **20**, 1158–1167.

Mans BJ, Louw AI and Neitz AWH (2002a) Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. *Molecular Biology and Evolution* **19**, 1695–1705.

Mans BJ, Ribeiro JM and Andersen, JF (2008c) Structure, function, and evolution of biogenic amine-binding proteins in soft ticks. *J. Biol. Chem.* **283**, 18721-18733.

Marcelino I, Almeida AM, Ventosa M, Pruneau L, Meyer DF, Martinez D, Lefrançoise T, Vachiérye N and Coelhob AV (2012) Tick-borne diseases in cattle: Applications of proteomics to develop new generation vaccines. *Journal of Proteomics* **75**, 4232-4250.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF and Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 1-14.

Maritz-Olivier C, Stutzer C, Jongejan F, Neitz AW and Gaspar AR (2007) Tick anti-hemostatics: targets for future vaccines and therapeutics. *Trends in Parasitology* **23**, 397–407.

Martin JA and Wang Z (2011) Next-generation transcriptome assembly, *Nature Reviews Genetics* **12**, 671-682.

Martin P and Lewis J (1992) Actin cables and epidermal movement in embryonic wound healing. *Nature* **360,**179–183.

Martins JR, Corrêa BL, Ceresér VH and Arteche CCP (1995) *A Situation Report on Resistance to Acaricides by the Cattle Tick Boophilus microplus in the State of Rio Grande do Sul, Southern Brazil*, Rodriguez CS and Fragoso SH (eds), Animal Acapulco, Mexico, pp. 1-8.

Maruyama SR, Anatriello E, Anderson JM, Ribeiro JM, Brandão LG, Valenzuela JG, Ferreira BR, Garcia GR, Szabó MP, Patel S, Bishop R and de Miranda-Santos IK (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.

McSwain JL, Essenberg RC and Sauer JR (1982) Protein changes in the salivary glands of the female lone star tick, *Amblyomma americanum*, during feeding. *Journal of Parasitology* **68**, 100–106.

Merino O, Alberdi P, la Lastra JMP and de la Fuente J (2013) Tick vaccines and the control of tick-borne pathogens. *Frontiers in Cellular and Infection Microbiology* **3**, 1-10.

Merino O, Almazán C, Canales M, Villar M, Moreno-Cid JA and Estrada-Peña A, Kocan KM and de la Fuente J (2011a) Control of *Rhipicephalus (Boophilus) microplus* infestations by the combination of subolesin vaccination and tick autocidal control after subolesin gene knockdown in ticks fed on cattle. *Vaccine* **29,** 2248–2254.

Merino O, Almazán C, Canales M, Villar M, Moreno-Cid JA, Galindo RC and de la Fuente J (2011b) Targeting the tick protective antigen subolesin reduces vector infestations and pathogen infection by *Anaplasma marginale* and *Babesia bigemina*. *Vaccine* **29**, 8575–8579.

Metzker ML (2010) Sequencing technologies – the next generation. *Nature Reviews* **11,** 31–46.

Meyer E, Aglyamova GV, Wang S, Buchanan-Carter J, Abrego D, Colbourne JK, Willis BL and Matz MV (2009) Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics* **10,** 1-18.

Meyer JM, Ejendal KFK, Watts VJ and Hill AC (2011) Molecular and pharmacological characterization of two D1-like dopamine receptors in the Lyme disease vector, *Ixodes scapularis*. *Insect Biochemistry and Molecular Biology* **41**, 563–571.

Mitra A, Skrzypczak M, Ginalski K and Rowicka M (2015) Strategies for Achieving High Sequencing Accuracy for Low Diversity Samples and Avoiding Sample Bleeding Using Illumina Platform. *PLoS One* **10**, e0120520.

Molloy MP, Herbert BR, Slade MB, Rabilloud T, Nouwens AS, Williams KL and Gooley AA (2000) Proteomic analysis of the *Escherichia coli* outer membrane. *European Journal of Biochemistry* **267**, 2871-81.

Morey M, Fernandez-Marmiesse A, Castineiras D, Fraga JM, Couce ML and Cocho JA (2013) A glimpse into past, present, and future DNA sequencing. *Molecular Genetics and Metabolism* **110**, 3–24.

Morozova O and Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**, 255-264.

Motoyama A and Yates JR (2008) Multidimensional LC Separations in Shotgun Proteomics. *Analytical Chemistry* **80**, 7187–7193.

Mount, D.W., 2007. Using the basic local alignment search tool (BLAST). *Cold Spring Harbor Protocols*, *2007*(7), pp.pdb-top17.

Mudenda L, Pierle SA, Turse JE, Scoles GA, Purvine SO, Nicora CD, Clauss TR, Ueti MW, Brown WC and Brayton KA (2014) Proteomics informed by transcriptomics identifies novel secreted proteins in *Dermacentor andersoni* saliva. *International Journal of Parasitology* **44**, 1029–1037.

Mulenga A, Blandon M and Khumthong R (2007) The molecular basis of the *Amblyomma americanum* tick attachment phase. *Experimental and Applied Acarology* **41**, 267–287.

Mulenga A, Sugimoto C and Onuma M (2000) Issues in tick vaccine development: identification and characterization of potential antigens. *Microbes and Infection* **2**, 1353–1361.

Mulenga A, Sugimoto C, Sako Y, Ohashi K, Musoke A, Shubash M and Onuma M (1999) Molecular characterisation of a *Haemaphysalis longicornis* tick salivary gland associated 29-Kilodalton protein and its effect as a vaccine against tick infestation in rabbits. *Infections and Immunology* **67,** 1652-1658.

Mulenga A, Sugino M, Nakajima M, Sugimoto C and Onuma M (2001) Tick-Encoded serine proteinase inhibitors (serpins); potential target antigens for tick vaccine development. *Journal of Veterinary Medical Science* **63**, 1063–1069.

Myllykangas S, Buenrostro J and Ji HP (2012) Overview of sequencing technology platforms. *Bioinformatics for High Throughput Sequencing*, 11–25.

Nagase H and Woessner JF Jr (1999) Matrix metalloproteinases. *Journal of Biological Chemistry* **274**, 21491–21494.

Narasimhan S, Koski RA, Beaulieu B, Anderson JF, Ramamoorthi N, Kantor F, Cappello M and Fikrig E (2002) A novel family of anticoagulants from the saliva of *Ixodes scapularis*. *Insect Molecular Biology* **11**, 641–650.

Needham GR and Pannabecker TL (1983) Effects of octopamine, chlordimeform, and demethylchlordimeform on amine-controlled tick salivary glands isolated from feeding *Amblyomma americanum*. *Pesticide Biochemistry and Physiology* **19**, 133–140.

Needham GR and Sauer JR (1979) Involvement of calcium and cyclic AMP in controlling ixodid tick salivary fluid secretion. *Journal of Parasitology* **65**, 531-542.

Nene V, Lee D, Quackenbush J, Skilton R, Mwaura S, Gardner MJ and Bishop R (2002) AvGI, an index of genes transcribed in the salivary glands of the ixodid tick *Amblyomma variegatum*. *International Journal of Parasitology* **32**, 1447–1456.

Nesvizhskii AI, Keller A, Kolker E and Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry* **75**, 4646–4658.

Nijhof AM, Balk JA, Postigo M, Rhebergen AM, Taoufik A and Jongejan F (2010). Bm86 homologues and novel ATAQ proteins with multiple epidermal growth factor (EGF)-like domains from hard and soft ticks. *International Journal of Parasitology* **40,** 1587–1597.

Nithiuthai S and Allen JR (1985) Langerhans cells present tick antigens to lymph node cells from tick-sensitised guinea pigs. *Immunology* **55,** 157–163.

Nunn MA, Sharma A, Paesen GC, Adamson S, Lissina O, Willis AC and Nuttall PA (2005) Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. *Journal of Immunology* **174**, 2084–2091.

Nutt LK, Pataer A, Pahler J, Fang B, Roth J, McConkey DJ and Swisher SG (2002) Bax and Bak promote apoptosis by modulating endoplasmic reticular and mitochondrial Ca^{2+} stores. *Journal of Biological Chemistry* **277**, 9219–9225.

Nuttall P A and Jones LD (1991) Non-viraemic tick borne virus transmission: mechanism and significance. In *Modern Acarology* (ed. Dusbabek F and Bukva V), Academia and SPB Academic Publishing, Prague and The Hague, Amsterdam, Netherlands, pp. 3–6.

Nuttall PA and Labuda M (2004) Tick–host interactions: saliva-activated transmission. *Parasitology* **129,** 177–189.

Nuttall PA and Labuda M (2008) Saliva-assisted transmission of tick-borne pathogens. In *Ticks: Biology, Disease and Control* (ed. Bowman AS and Nuttall PA), pp. 205–219. Cambridge University Press, Cambridge.

Nuttall PA, Paesen GC, Lawrie CH and Wang H (2000) Vector-Host Interactions in Disease Transmission. *Journal of Molecular Microbiology and Biotechnology* **2,** 381-386.

Nuttall PA, Trimnell RA, Kazimírová M and Labuda M (2006) Exposed and concealed antigens as vaccine targets for controlling ticks and tick-borne diseases. *Parasite Immunology* **28**, 155–163.

Odongo D, Kamau L, Skilton R, Mwaura S, Nitsch C, Musoke A, Taracha E, Daubenberger C, Bishop R (2007) Vaccination of cattle with TickGARD induces cross-reactive antibodies binding to conserved linear peptides of Bm86 homologues in *Boophilus decoloratus*. *Vaccine* **25**, 1287–1296.

Oleaga A, Escudero-Población A, Camafeita E and Pérez-Sánchez R (2007) A proteomic approach to the identification of salivary proteins from the argasid ticks *Ornithodoros moubata* and *Ornithodoros erraticus*. *Insect Biochemistry and Molecular Biology* **37**, 1149-1159.

Oliveira CJ, Anatriello E, de Miranda-Santos IK, Francischetti IM, Sa-Nunes A, Ferreira BR and Ribeiro JM (2013) Proteome of *Rhipicephalus sanguineus* tick saliva induced by the secretagogues pilocarpine and dopamine. *Ticks and Tick-borne Diseases* **4**, 469–477.

Oliveira CJ, Sa-Nunes A, Francischetti IM, Carregaro V, Anatriello E, Silva JS, Santos IK, Ribeiro JM and Ferreira BR (2011) Deconstructing tick saliva: non-protein molecules with potent immunomodulatory properties. *Journal of Biological Chemistry* **286**, 10960–10969.

Olwoch JM, Van Jaarsveld AS, Scholtz CH and Horak IG (2007) Climate change and the genus *Rhipicephalus* (Acari: Ixodidae) in Africa. *Onderstepoort Journal of Veterinary Research* **74**, 45–72.

Osorio JE, Godsey MS, Defolliart GR and Yuill M (1996) La Crosse viremias in white-tailed deer and chipmunks exposed by injection or mosquito bite. *American Journal of Tropical Medicine and Hygiene* **54,** 338–342.

Paesen GC, Adams PL, Harlos K, Nuttall PA and Stuart DI (1999) Tick histamine binding proteins: isolation, cloning, and three-dimensional structure. *Molecular Cell* **3**, 661.

Paine SH, Kemp DH and Allen JR (1983) *In vitro* feeding of *Dermacentor andersoni* (stiles): effects of histamine and other mediators. *Parasitology* **86**, 419-428.

Palmer GH, Oberle SM, Barbet AF, Goff WL , Davis WC and McGuire TC (1988) Immunization of cattle with a 36-kilodalton surface protein induces protection against homologous and heterologous *Anaplasma marginale* challenge. *Infection and Immunity* **56**, 1526-1531.

Pandey A and Mann M (2000) Proteomics to study genes and genomes. *Nature* **405**, 837-846.

Pannabecker T and Needham GR (1982) Effects of octopamine on fluid secretion by isolated salivary glands of a feeding ixodid tick. *Archives of Insect Biochemistry and Physiology* **2**, 217–226.

Pareek CS, Smoczynski R and Tretyn A (2011) Sequencing technologies and genome sequencing. *Journal of Applied Genetics* **52,** 413–435.

Parizi LF, Githaka NW, Logullo C, Konnai S, Masuda A and Ohashi K and da Silva VIJ (2012a) The quest for a universal vaccine against ticks: cross immunity insights. *The Veterinary Journal* **194,** 158–165.

Parizi LF, Pohl PC, Masuda A and da Silva VIJ (2009) New approaches toward anti-*Rhipicephalus (Boophilus) microplus* tick vaccine. *Revista Brasileira de Parasitologia Veterinária* **18**, 1-7.

Parizi LF, Reck J, Oldiges DP, Guizzo MG, Seixas A, Logullo C, de Oliveira PL, Termignoni C, Martins JR and Vaz Ida S Jr (2012b) Multi-antigenic vaccine against the cattle tick *Rhipicephalus (Boophilus) microplus*: a field evaluation. *Vaccine* **30**, 6912–6917.

Park OK (2004) Proteomic studies in plants. *Journal of Biochemistry and Molecular Biology* **37**, 133–138.

Patramool S, Choumet V, Surasombatpattana P, Sabatier L, Thomas F, Thongrungkiat S, Rabilloud T, Boulanger N, Biron DG and Misse D (2012) Update on the proteomics of major arthropod vectors of human and animal pathogens. *Proteomics* **12**, 3510–3523.

Penichet ML, Rodríguez M, Castellano O, Mandado S, Rojas Y, Lleonart R and de la Fuente J (1994) Detection of Bm86 antigen in different strains of *Boophilus microplus* and effectiveness of immunization with recombinant Bm86. *Parasite Immunology* **16**, 493–500.

Peter RJ, van den Bossche P, Penzhorn BL and Sharp B (2005) Tick, fly, and mosquito control—Lessons from the past, solutions for the future. *Veterinary Parasitology* **132,** 205–215.

Petersen TN, Brunak S, von Heijne G and Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**,785-786.

Pichu S, Yalcin EB, Ribeiro JM, King RS and Mather TN (2011) Molecular characterization of novel sulfotransferases from the tick, *Ixodes scapularis*. *BMC Biochemistry* **12**, 1-11.

Popara M, Villar M and de la Fuente J (2015) Proteomics Characterization of Tick-Host-Pathogen Interactions. *Methods in Molecular Biology* **1247**, 513-527.

Porreca GJ, Zhang K, Li JB, Xie B, Austin D, Vassallo SL, LeProust EM, Peck BJ, Emig CJ, Dahl F, Gao Y, Church GM and Shendure J (2007) Multiplex amplification of large sets of human exons. *Nature Methods* **4**, 931 – 936.

Prevot PP, Couvreur B, Denis V, Brossard M, Vanhamme L and Godfroid E (2007) Protective immunity against *Ixodes ricinus* induced by a salivary serpin. *Vaccine* **25**, 3284–3292.

Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA and Baumeister K (1987) A system for rapid DNA sequencing with fluorescent chain terminating dideoxynucleotides. *Science* **238,** 336–341.

Rachinsky A, Guerro FD and Scoles GA (2007) Differential protein expression in ovaries of uninfected and Babesia-infected southern cattle ticks, *Rhipicephalus (Boophilus) microplus*. *Insect Biochemistry and Molecular Biology* **37**, 1291-1308.

Radulović ZM, Kim TK, Porter LM, Sze SH, Lewis L and Mulenga A (2014) A 24–48 h fed *Amblyomma americanum* tick saliva immuno-proteome. *BMC Genomics* **15**, 1-30.

Rajput ZI, Hu S, Chen W, Arijo A and Xiao C (2006) Importance of ticks and their chemical and immunological control in livestock. *Journal of Zhejiang University Science B* **7,** 912-921.

Rawlings ND, Waller M, Barrett AJ and Bateman A (2013) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* **40**, D343–D350.

Reuter JA, Spacek D and Snyder MP (2015) High-Throughput Sequencing Technologies. *Molecular Cell* **58**, 586–597.

Ribeiro JM (1995) Blood-feeding arthropods: live syringes or invertebrate pharmacologists? *Infectious Agents and Disease* **4**, 143–152.

Ribeiro JM and Francischetti IM (2003) Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annual Review of Entomology* **48**, 73–88.

Ribeiro JM, Alarcon-Chaidez F, Francischetti IM, Mans BJ, Mather TN, Valenzuela JG and Wikel SK (2006) An annotated catalog of salivary gland transcripts from Ixodes scapularis ticks. *Insect Biochemistry and Molecular Biology* **36**, 111–129.

Ribeiro JM, Anderson JM, Manoukis NC, Meng Z and Francischetti IM (2011) A further insight into the sialome of the tropical bont tick, *Amblyomma variegatum. BMC Genomics* **12**, 1-11.

Ribeiro JM, Labruna MB, Mans BJ, Maruyama SR, Francischetti IM, Barizon GC and de Miranda Santos IK (2012) The sialotranscriptome of *Antricola delacruzi* female ticks is compatible with non-hematophagous behavior and an alternative source of food. *Insect Biochemistry and Molecular Biology* **42**, 332-342.

Ribeiro JMC (1995) How ticks make a living. *Parasitology Today* **11,** 91–93.

Rinker DC, Pitts RJ and Zwiebel LJ (2016) Disease vectors in the era of next generation sequencing. *Genome Biology* **17**, 1-11.

Rockey DC, Weymouth N and Shi Z (2013) Smooth muscle alpha actin (Acta2) and myofibroblast function during hepatic wound healing. *PLoS One* **8**, e77166.

Rodríguez M, Massard CL, da Fonseca AH, Ramos NF, Machado H and Labarta V (1995) Effect of vaccination with a recombinant Bm86 antigen preparation on natural infestations of *Boophilus microplus* in grazing dairy and beef pure and cross-bred cattle in Brazil. *Vaccine* **13,** 1804–1808.

Rodriguez M, Massard CL, da Fonseca AH, Ramos NF, Machado H, Labarta V and de la Fuente J (1999a) Effect of vaccination with a recombinant Bm86 antigen preparation on natural infestations of *Boophilus microplus* in grazing dairy and beef pure and cross breed cattle in Brazil. *Vaccine* **13**, 1804-1808.

Rodriguez M, Penichet ML, Mouris AE, Labarta V, Luaces LL, Rubiera R, Cordoves C, Sachez PA, Ramos E, Soto A, Canales M, Pelenzuela D, Triguero A, Lleonart R, Herrera L and de la Fuente (1995b) Control of *Boophilus microplus* populations in grazing cattle vaccinated with a recombinant Bm86 antigen preparation. *Veterinary Parasitology* **57**, 339–349.

Rodríguez M, Rubiera R, Penichet M, Montesino R, Cremeta J, Falcon V, Sanchez G, Bringas R, Cordoves C, Valdez M, Lleonart R, Herrera L and de la Fuente J (1994) High level expression of the *B. microplus* Bm86 antigen in the yeast *Pichia pastoris* forming highly immunogenic particles for cattle. *Journal of Biotechnology* **33**, 135–146.

Rodriguez-Valle M, Moolhuijzen P, Piper EK, Weiss O, Vance M, Bellgard M and Lew-Tabor A (2013) *Rhipicephalus microplus* lipocalins (LRMs): genomic identification and analysis of the bovine immune response using in silico predicted B and T cell epitopes. *International Journal of Parasitology* **43**, 739–752.

Rodriguez-Valle M, Xu T, Kurscheid S and Lew-Tabor AE (2015) *Rhipicephalus microplus* serine protease inhibitor family: annotation, expression and functional characterisation assessment. *Parasitology Vectors* **8**, 1-9.

Ronaghi M, Karamohamed S, Pettersson B, Uhlén M and Nyrén P (1996) Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. *Analytical Biochemistry* **242**, 84-89.

Rosendale AJ, Romick-Rosendale LE, Watanabe M, Dunlevy ME and Benoit JB (2016) Mechanistic underpinnings of dehydration stress in the American dog tick revealed through RNA-Seq and metabolomics. *Journal of Experimental Biology* **219**, 1808-1819.

Rutti B, Schlunegger B, Kaufman W and Aeschlimann A. (1980) Properties of Na, K-ATPase from the salivary glands of the ixodid tick *Amblyomma hebraeum*. *Canadian Journal of Zoology* **58**, 1052–1059.

Sangamnatdej S, Paesen GC, Slovak M and PA Nuttall (2002) A high affinity serotonin- and histamine-binding lipocalin from tick saliva. *Insect Biochemistry and Molecular Biology* **11**, 79–86.

Sanger F, Nicklen S and Coulson R (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-5467.

Santos IK, Valenzuela JG, Ribeiro JM, de Castro M, Costa JN, Costa AM, da Silva ER, Neto OB, Rocha C, Daffre S, Ferreira BR, da Silva JS, Szabó MP and Bechara GH (2004) Gene discovery in *Boophilus microplus*, the cattle tick: the transcriptomes of ovaries, salivary glands, and hemocytes. *Annals of the New York Academy of Sciences* 1**026**, 242–246.

Sasaki SD, Azzolini SS, Hirata IY, Andreotti R and Tanaka AS (2004) *Boophilus microplus* tick larvae, a rich source of Kunitz-type serine proteinase inhibitors. *Biochimie* **86**, 643–649.

Sauer JR and Hair JA (1971) Water balance in the lone star tick (Acarina: Ixodidae): the effects of relative humidity and temperature on weight changes and total water content. *Journal of Medical Entomology* **8**, 479–485.

Sauer JR, Essenberg RC and Bowman AS (2000) Salivary glands in ixodid ticks: control and mechanism of secretion. *Journal of Insect Physiology* **46**, 1069-1078.

Sauer JR, McSwain JL, Bowman AS and Essenberg RC (1995) Tick salivary gland physiology. *Annual Review of Entomology* **40**, 245–267.

Schmidt A, Forne I and Imhof A (2014) Bioinformatic analysis of proteomics data. *BMC Systems Biology* **8**, S3.

Schwarz A, Tenzer S, Hackenberg M, Erhart J, Gerhold-Ay A, Mazur J, Kuharev J, Ribeiro JM and Kotsyfakis M (2014) A systems level analysis reveals transcriptomic and proteomic complexity in *Ixodes ricinus* midgut and salivary glands during early attachment and feeding. *Molecular & Cellular Proteomics* **13**, 2725–2735.

Schwarz A, von Reumont BM, Erhart J, Chagas AC, Ribeiro JM and Kotsyfakis M (2013) *De novo Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB Journal* **27**, 4745–4756.

Shao C, Liu Y, Ruan H, Li Y, Wang H, Kohl F, Goropashnaya AV, Fedorov VB, Zeng R, Barnes BM and Yan J (2010) Shotgun proteomics analysis of hibernating arctic ground squirrels. *Molecular and Cellular Proteomics* **9**, 313-326.

Shapiro SZ, Voigt WP and Ellis JA (1989) Acquired resistance to ixodid ticks induced by tick cement antigen. *Experimental and Applied Acarology* **7**, 33-41.

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD and Church GM (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV and Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212.

Šimo L, Koči J, Žitňan D and Park Y (2011) Evidence for D1 dopamine receptor activation by a paracrine signal of dopamine in tick salivary glands. *PLoS ONE* **6**, e16158.

Šimo L, Žitňan D and Park Y (2012) Neural control of salivary glands in ixodid ticks. *Journal of Insect Physiology* **58**, 459–466.

Singh KS and Girschick HJ (2003) Tick-host interactions and their immunological implications in tick-borne diseases. *Current Science* **85**, 1284–1298.

Smith AM, Heisler LE, St Onge RP, Farias-Hesson E, Wallace IM, Bodeau J, Harris AN, Perry KM, Giaever G, Pourmand N and Nislow C (2010) Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. *Nucleic Acids Research* **38**, e142.

Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SB and Hood LE (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* **321,** 674–679.

Soares TS, Oliveira F, Torquato RJS, Sasaki SD, Araujo MS, Paschoalin T and Tanaka AS (2016) BmTI-A, a Kunitz-type inhibitor from *Rhipicephalus microplus* able to interfere in vessel formation. *Veterinary Parasitology* **219**, 44–52.

Sonenshine DE, Bissinger BW, Egekwu N, Donohue KV, Khalil SM and Roe RM. (2011) First transcriptome of the testis-vas deferens-male accessory gland and proteome of the spermatophore from *Dermacentor variabilis* (Acari: Ixodidae). *PLoS ONE* **6**, e24711.

Sonenshine DE: Biology of ticks. 1993, Oxford: Oxford University Press.

Sossai S, Peconick AP, Sales-Junior PA, Marcelino FC, Vargas MI, Neves ES and Patarroyo JH (2005) Polymorphism of the Bm86 gene in South American strains of the cattle tick *Boophilus microplus. Experimental and Applied Acarology* **37**, 199-214.

Strickler SR, Bombarely A and Mueller LA (2012) Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *American Journal of Botany* **99**, 257-66.

Strudwick XL and Cowin AJ (2012) Cytoskeletal regulation of dermal regeneration. *Cells* **1**, 1313–1327.

Sugino M, Imamura S, Mulenga A, Nakajima M, Tsuda A, Ohashi K and Onuma M (2003) A serine proteinase inhibitor (serpin) from ixodid tick *Haemaphysalis longicornis*; cloning and preliminary assessment of its suitability as a candidate for a tick vaccine. V*accine* **21**, 2844–2851.

Sun Y, Yi X, Peng M, Zeng H, Wang D, Li B, Tong Z, Chang L, Jin X and Wang X (2014) Proteomics of *Fusarium oxysporum* Race 1 and Race 4 Reveals Enzymes Involved in Carbohydrate Metabolism and Ion Transport That Might Play Important Roles in Banana Fusarium Wilt. *PLoS One* **9**, e113818.

Tan AWL, Francischetti IMB, Slovak M, Manjunatha KR and Ribeiro JMC (2015) Sexual differences in the sialomes of the zebra tick, *Rhipicephalus pulchellus*. *Journal of Proteomics* **117**, 120–144.

Tanaka AS, Andreotti R, Gomes A, Torquato RJ, Sampaio MU and Sampaio CA (1999) A double headed serine proteinase inhibitor–human plasma kallikrein and elastase inhibitor–from *Boophilus microplus* larvae. *Immunopharmacology* **45**, 171–177.

Tatchell RJ (1967) Salivary secretion in the cattle tick as a means of water elimination. *Nature* **213**, 940-941.

Tawfik DS and Griffiths AD (1998) Man-made cell-like compartments for molecular evolution. *Nature Biotechnology* **16,** 652-656.

Tellam RL, Kemp D, Riding G, Briscoe S, Smith D, Sharp P, Irving D and Willadsen P (2002) Reduced oviposition of *Boophilus microplus* feeding on sheep vaccinated with vitellin. *Veterinary Parasitology* **103**, 141-156.

Thelen JJ (2007) Quantitative Proteomics in Plants: Choices in Abundance. *Plant Cell* **19**, 3339–3346.

Tirloni L, Islam MS, Kim TK, Diedrich JK, Yates JR III, Pinto AFM, Mulenga A, You MJ and Da Silva Vaz I Jr (2015) Saliva from nymph and adult females of *Haemaphysalis longicornis*: a proteomic study. *Parasites and Vectors* **8,** 338.

Tirloni L, Islam MS, Kim TK, Diedrich JK, YatesIII JR, Pinto AFM, Mulenga A, You MJ and Da Silva VazJ I (2014) Saliva from nymph and adult females of *Haemaphysalis longicornis*: a proteomic study. *Parasites and Vectors* **8**, 1-23.

Tirloni L, Kim TK, Coutinho ML, Ali A, Seixas A, Termignoni C, Mulenga A and da Silva Vaz Jr I (2016) The putative role of *Rhipicephalus microplus* salivary serpins in the tick-host relationship. *Insect Biochemistry and Molecular Biology* **71**, 12–28.

Tirloni L, Reck J, Terra RMS, Martins JR, Mulenga A, Sherman NE, Fox JW, Yates JR, Termignoni C, Pinto AFM and da Silva Vaz I Jr (2014) Proteomic Analysis of Cattle Tick *Rhipicephalus (Boophilus) microplus* Saliva: A Comparison between Partially and Fully Engorged Females. *PLoS ONE* **9**, e94831.

Titus RG and Ribeiro JC (1988) Salivary gland lysates from the sandfly *Lutzomyia longipalpis* enhance *Leishmania* infectivity. *Science* **239,** 1306–1308.

Tønnesen MH, Penzhorn BL, Bryson NR, Stoltsz WH and Masibigiri T (2004) Displacement of *Boophilus decoloratus* by *Boophilus microplus* in the Soutpansberg region, Limpopo Province, South Africa. *Experimental and Applied Acarology* **32,** 199-208.

Trimnell AR, Davies GM, Lissina O, Hails RS and Nuttall PA (2005) A cross-reactive tick cement antigen is a candidate broad-spectrum tick vaccine. *Vaccine* **23**, 4329–4341.

Trimnell AR, Hails RS and Nuttall PA (2002) Dual action ectoparasite vaccine targeting 'exposed' and 'concealed' antigens. *Vaccine* **20**, 3560-3568.

Tsuda A , Mulenga A, Sugimoto C, Nakajima M, Ohashi K and Onuma M (2001) cDNA cloning, characterization and vaccine effect analysis of *Haemaphysalis longicornis* tick saliva proteins. *Vaccine* **19**, 4287-4296.

Ullmann-Moore, A.J. (2007) Tick salivary gland proteins (Salps) as antigens for vaccine development to decrease spirochete load in a murine model of Lyme Borreliosis, Degree of Doctor of Philosophy, Colorado State University, Fort Collins, Colorado.

Untalan PM, Guerrero FD, Haines LR and Pearson TW (2005) Proteome analysis of abundantly expressed proteins from unfed larvae of the cattle tick, *Boophilus microplus*. *Insect Biochemistry and Molecular Biology* **35**, 141–151.

Valenzuela JG (2004) Exploring tick saliva: from biochemistry to“sialomes”and functional genomics. *Parasitology* **129,** 83–94.

Valenzuela JG, Francischetti IM, Pham VM, Garfield MK and Ribeiro JM (2003) Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. *Insect Biochemistry and Molecular Biology* **33**, 717– 732.

Valenzuela JG, Francischetti IMB, Pham VM, Garfield MK, Mather TN and Ribeiro JMC (2002a) Exploring the sialome of the tick, *Ixodes scapularis*. *Journal of Experimental Biology* **205**, 2843–2864.

Valouev A, J. Ichikawa I, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A and Johnson SM (2008) A high resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence dictated positioning. *Genome Research* **18**, 1051–1063.

van de Locht A, Stubbs MT, Bode W, Friedrich T, Bollshweiler C, Hoffken HW and Huber R (1996) The ornithodorin, thrombin crystal structure, a key to the TAP enigma? *EMBO Journal* **15**, 6011–6017.

Van Dijk EL, Auger H, Jaszczyszyn Y and Thermes C (2014) Ten years of next-generation sequencing technology. *Trends in Genetics* **30**, 418-426.

Van Nueten JM, Janssens WJ, Vanhoutte PM (1985) 'Serotonin and vascular smooth muscle', in P. M. Vanhoutte (ed.), Serotonin and the Cardiovascular System, Raven Press, New York, pp. 95–104.

Velculescu VE, Zhang L, Vogelstein B and Kinzler KW (1995) Serial analysis of gene expression. *Science* **270,** 484-487.

Vennestrøm J and Jensen PM (2007) *Ixodes ricinus*: the potential of two-dimensional gel electrophoresis as a tool for studying host-vector-pathogen interactions. *Experimental Parasitology* **115**, 53-58.

Villar M, Popara M, Mangold AM and de la Fuente J (2014b) Comparative proteomics for the characterization of the most relevant *Amblyomma* tick species as vectors of zoonotic pathogens worldwide. *Journal of Proteomics* **105**, 204-216.

Villar M, Popara M, Bonzón-Kulichenko E, Ayllón N, Vázquez J and de la Fuente J (2012) Characterization of the tick-pathogen interface by quantitative proteomics. *Ticks and Tick Borne Diseases* **3**, 154-158.

Villar M, Popara M, Ayllón N, Fernández de Mera IG, Mateos-Hernández L, Galindo RC, Manrique M, Tobes R and de la Fuente J (2014a) A systems biology approach to the characterization of stress response in *Dermacentor reticulatus* tick unfed larvae. *PLoS One* **9**, e89564.

Villar M, Torina A, Nuñez Y, Zivkovic Z, Marina A, Alongi A, Salvatore S, La Barbera G, Caracappa S, Vázquez J, de la Fuente J (2010) Application of highly sensitive saturation labeling to the analysis of differential protein expression in infected ticks from limited samples. *Proteome Science* **8**, 43.

Voelkerding KV, Dames SA and Durtschi JD (2009) Next-generation sequencing: from basic research to diagnostics. *Clinic Chemistry* **55**, 641-658.

Vora A, Taank V, Dutta SM, Anderson JF, Fish D, Sonenshine DE, Catravas JD, Sultana H and Neelakanta G (2017) Ticks elicit variable fibrinogenolytic activities upon feeding on hosts with different immune backgrounds. *Scientific reports* **7**, 44593.

Walker AR, Bouattour A, Camicas JL, Estrada-Pena A, Horak IG, Latif AA, Pegram R and Preston PM (2003) *Ticks of Domestic Animals in Africa: A Guide to Identification of Species***.** *Bioscience Reports*, Edinburgh, United Kingdom, pp 74-221.

Wall PK, Leebens-Mack J, Chanderbali AS, Barakat A, Wolcott E, Liang H, Landherr L, Tomsho LP, Hu Y, Carlson JE, Ma H, Schuster SC, Solti DE, Soltis PS, Altman N and de Pamphilis CW (2009) Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genomics* **10**, 1-19.

Wang H, Paesen GC, Nuttall PA, Barbour AG (1998) Male ticks help their mates to feed. *Nature* **391**, 753–754.

Wang M, Guerrero FD, Pertea G and Nene VM (2007) Global comparative analysis of ESTs from the southern cattle tick, *Rhipicephalus (Boophilus) microplus*. *BMC Genomics* **8,** 1-14.

Wang Z, Gerstein M and Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10,** 57–63.

Waxman L, Smith DE, Arcuri KE and Vlasuk GP (1990) Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. *Science* **248**, 593-596.

Weiss BL and Kaufman RW (2004) Two feeding-induced proteins from the male gonad trigger engorgement of the female tick *Amblyomma hebraeum*. *Proceedings of the National Academy of Sciences* **101**, 5874–5879.

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA and Rothberg JM (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**, 872–876.

White MV and Kaliner MA (1987) Neutrophils and mast cells. I. Human neutrophil-derived histamine-releasing activity. *Journal of Immunology* **139**, 1624–1630.

Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF and Williams KL (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology and Genetic Engineering Reviews* **13**, 19-50.

Willadsen P (2001) The molecular revolution in the development of vaccines against ectoparasites. *Veterinary Parasitology* **101**, 353-368.

Willadsen P (2006) Tick control: Thoughts on a research agenda. *Veterinary Parasitology* **138,** 161 – 168.

Willadsen P and Jongejan F (1999) Immunology of the tick–host interaction and the control of ticks and tick-borne diseases. *Parasitology Today* **15,** 258–262.
Willadsen P and Kemp DH (1988) Vaccination with ‘concealed’ antigens for tick control. *Parasitology Today* **4,** 196–198.

Willadsen P, Bird P, Cobon GS and Hungerford J (1995) Commercialization of a recombinant vaccine against *Boophilus microplus*. *Parasitology* **110**, 43–50.

Willadsen P, Cobon G and McKenna RV (1996) Comparative vaccination of cattle against *Boophilus microplus* with recombinant antigen Bm86 alone or in combination with recombinant Bm91. *Parasite Immunology* **18**, 241-246.

Xu XL, Cheng TY, Yang H, Yan F and Yang Y (2015) De novo sequencing, assembly and analysis of salivary gland transcriptome of *Haemaphysalis flava* and identification of sialoprotein genes. *Infection, Genetics and Evolution* **32**, 135-142.

Yalcin EB, Stangl H, Pichu S, Mather TN and King RS (2010) Monoamine neurotransmitters as substrates for novel tick sulfotransferases, homology modeling, molecular docking, and enzyme kinetics. *ACS Chemical Biology* **6**, 176–184.

Yeung ES (2011) Genome-wide Correlation between mRNA and Protein in a Single Cell. *The journal of the Gesellschaft Deutscher Chemiker* **50**, 583–585.

You M and Fujisaki K (2009) Vaccination effects of recombinant chitinase protein from the hard tick *Haemaphysalis longicornis* (Acari: Ixodidae). *Journal of Veterinary Medical Science* **71**, 709-712.

You M, Xuan X, Tsuji N, Kamio T, Igarashi I, Nagasawa H, Mikami T and Fujisaki K (2001) Molecular characterization of a troponin I-like protein from the hard tick *Haemaphysalis longicornis*. *Insect Biochemistry and Molecular Biology* **32**, 67-73.

You MJ (2005) Immunization of mice with recombinant P27/30 protein confers protection against hard tick *Haemaphysalis longicornis* (Acari: Ixodidae) infestation. *Journal of Veterinary Science* **6**, 47-51.

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.

Żmieńko A, Jackowiak P and Figlerowicz M (2011) Transcriptome sequencing: next generation approach to RNA functional analysis. *Journal of Biotechnology, Computational Biology and Bionanotechnology* **92**, 311-319.

Websites and databases

(*Bio-Rad Laboratories*, Hercules, California, USA)

(Invitrogen, Life Technologies, CA, USA).

(Life Technologies, Carlsbad, CA, USA)

(Qiagen, Valencia, CA, USA).

CLCBio, Aarhus, Denmark

http://www.cfsph.iastate.edu/Factsheets/pdfs/boophilus_microplus.pdf

http://phobius.sbc.su.se/

http://www.cbs.dtu.dk/services/TMHMM/

http://www.uniprot.org/)

https://www.expasy.org/

Mascot (Mascot v2.2 (Matrix Science)

NCBI – BLAST nr

RNAlater (Ambion, Austin, Texas, USA).

Scaffold software version 4.0.5 (Proteome Software Inc., Portland, OR)

X-Tandem (http://www.thegpm.org/tandem/)