In silico and two-hybrid functional characterization of the *Rhipicephalus microplus* (Canestrini, 1888) vaccine antigen Bm86

by

Alanna Margaret James

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

In the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

January 2020

I, **Alanna Margaret James** declare that the dissertation, which I hereby submit for the degree ***Magister Scientiae* with specialisation in Genetics** at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:  DATE: January 2020

ACKNOWLEDGEMENTS

My Supervisor, Prof C. Maritz-Olivier, for her infectious passion for the subject that inspired my love for genetics in her third-year lectures. For sharing her vast knowledge and for providing lasting support, advice and endless encouragement

My Co-Supervisor Dr C. Stutzer for his valuable advice, humour and support

Dr N. Olivier and Dr A. van der Merwe for their technical support with Bioinformatic analyses

Dr D. Coertzen, Dr E. Venter and Mr R. Gilfillan for the use of their laboratory equipment as well as invaluable laboratory support and advice

Dr H. Stoltz and the Hans Hoheisen Wildlife Research Station and staff for the use of their facilities

The entire Ticks and Tick-borne Diseases Research Group

The various funding bodies, including the National Research Foundation (NRF), Technology Innovation Agency (TIA) and the University of Pretoria (UP) without whom this dissertation would not be possible

My family and partner for their unconditional love and endless faith, patience and support throughout my studies

SUMMARY

Ticks are considered economically significant parasitic vectors in the livestock industry globally. Vector directed vaccines are a promising method of tick control. To date only one protective antigen, Bm86, has been commercialised; however, it has variable efficacy across regions and, despite a previous study which proposed a possible role for Bm86 in Phospholipase C signalling (PLC), its biological role is unclear. Antigen 1, a new vaccine candidate, has been identified to interact with Bm86, but the region of protein interaction with Bm86 is unknown.

This study assembled transcriptomes of the larval, nymph, and adult gut, ovary and salivary gland via *de novo* RNA sequence assembly to identify PLC pathway components, Bm86 and Antigen 1 in South African *Rhipicephalus microplus*. G proteins, namely, $G_{\alpha i/o/s/q}$ as well as a G_{γ} and G_{β} subunits and $PLC_{\beta, \gamma, \delta, \eta}$ and $_{\varepsilon,}$ were identified. This is the first study to identify these proteins expressed in *R. microplus.*

Furthermore, regions of sequence variation in the South African Bm86 and Antigen 1 in various life stages and tissues of *R. microplus* ticks were analysed. Also, the study used a yeast-two-hybrid model to study the region of Antigen 1 interaction with Bm86 and found it to be the N-terminal region of Antigen 1. These findings are particularly important in the design of tick control strategies such as vaccine and acaricide development; providing insight into the mechanism of two vaccine antigens as well as identifying various potential novel drug targets in the biology of *R. microplus*.

CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendix

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

LIST OF ABBREVIATIONS

1º	Primary
2º	Secondary
3-AT	3-Amino-1,2,4-triazole
AA	Amino acid
ACE	Angiotensin-converting enzyme
AD	Activation domain
ADE2	Adenine 2 reporter gene
Amp	Ampicillin
Amp^r	Ampicillin resistance gene
AP	Alkaline phosphatase
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BD	Binding domain
BED	Browser extensible data
BLAST	Basic local alignment tool
BLOSUM	Blocks of amino acid substitution matrix
bp	Base pairs
BPTI	Bovine pancreatic trypsin inhibitor
BUSCO	Benchmarking universal single-copy orthologs
cAMP	Cyclic adenosine monophosphate
cDNA	Complementary deoxyribonucleic acid
CEGMA	Core eukaryotic genes mapping approach
CFU	Colony forming units
CWC	Column wash solution
DAF	Decay accelerating factor
DAG	Diacylglycerol
dddH_2O	Double distilled deionised water
DDO	Double drop out
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic Acid
DO	Drop out
DRhoGEF2	Rho guanine nucleotide exchange factor
ds	Double stranded
DTT	Dithiothreitol

EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
ELISA	Enzyme-linked immunosorbent assay
ERB	Endotoxin removal wash
G protein	Guanosine triphosphate binding protein
GABA	Gamma-aminobutyric acid
GDP	Guanosine diphosphate
GPCR	G protein-coupled receptor
GPI	Glycosyl phosphatidylinositol
GPI-PLC	GPI-anchor specific PLC
GPI-PLD	GPI-anchor specific PLD
GTP	Guanosine 5'-triphosphate
HIS3	Histidine 3 reporter gene
HMM	Hidden Markov model
HRP	Horse-radish peroxidase
IP3/4/5/6	Inositol-1,4,5-triphosphate 3/4/5/6
Kan	Kanamycin
Kan^r	Kanamycin resistance gene
KOG	Eukaryotic orthologous groups
KUBP	Antigen 1
LB	Luria-Bertani
LEU2	Leucine 2 reporter gene
MCS	Multiple cloning site
MEL1 (lacZ)	β-galactosidase reporter gene
ML	Maximum Likelihood
NCBI	National centre for Biotechnology information
NEB	New England Biolabs® Inc. (UK)
ng	Nanogram
norpA	No receptor potential A
OD	Optical density
Olf	Olfactory
OPD	Ortho-phenylenediamine
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

PEG	Polyethylene glycol
PI	Phosphatidylinositol
PI3K	Phosphoinositide-3-kinase
PIP_2	Phosphadiylinositol-4,5-bisphosphate
PKC	Protein kinase C
PLA2	Phospholipase A2
PLC	Phospholipase C
PLD	Phospholipase D
PSSM	Position-specific scoring matrix
QC	Quality control
QDO	Quadruple drop out
RDH	Retinol dehydrogenase
RIN	RNA integrity value
RNA	Ribonucleic acid
RNAseq	RNA sequencing
RT	Reverse transcriptase
SAP	Shrimp alkaline phosphatase
SD	Selective dropout
SDO	Single drop out
SP	Sequencing primer
T1	Antigen 1 truncation 1
T2	Antigen 1 truncation 2
T3	Antigen 1 truncation 3 (full length)
TAE	Tris base acetic acid and EDTA buffer
TBS	Tris-buffered saline
TDO	Triple drop out
TF	Transcription factor
tRNA	Transfer ribonucleic acid
TRP1	Tryptophan 1 reporter gene
UDP	Uridine diphosphate
URA3	Uracil 3 reporter gene
Y2H	Yeast-two-hybrid
YNB	Yeast nitrogen base
YPD	Yeast extract, Peptone, Dextrose
YPDA	Yeast extract, Peptone, Dextrose, Adenine

CHAPTER 1

INTRODUCTION

Ticks and tick control.

Tick classification and feeding

Ticks are hematophagous ectoparasitic arthropods belonging to the class Arachnida, subclass Acari, order Parasitiformes, suborder *Ixodida* (Walker *et al*., 2003) (Figure 1.1) and are considered one of the most common vectors for debilitating and lethal zoonotic diseases, second to mosquitos (de la Fuente *et al*., 2008; Gu *et al*., 2014; Sonenshine and Roe, 2014). Ticks and their associated direct and indirect effects on their hosts are considered a significant contributor to livestock morbidity and mortality globally, resulting in substantial economic losses for livestock farmers. Currently, it is estimated that some 80% of the world's cattle population is exposed to ticks and tick-borne pathogens; stressing the need for improved control of tick loads and tick-borne diseases (Castro, 1997; Nyangiwe *et al*., 2018).

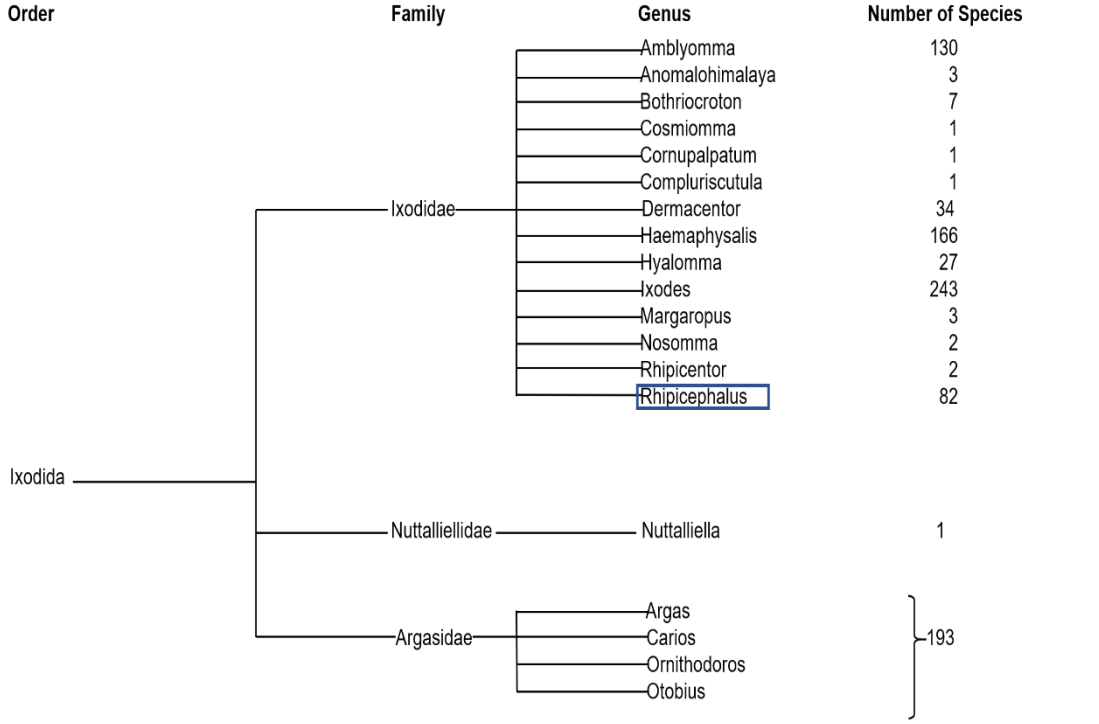


Figure 1.1: Cladogram showing the classification of Acari: Ixodida. *Rhipicephalus* is indicated in the blue box. Ticks are classified based on their morphology into two families, hard (Ixodidae) or soft (Argasidae) ticks, with a third family; the Nutalliellidae having only one species (Mans *et al*. 2012). The figure was adapted from Richards 2015.

As hematophagous ectoparasites, ticks require a blood meal for survival (Reviewed by Anderson and Magnarelli, 2008). Tick feeding takes place in nine sequential steps, beginning with the quest for a suitable host and a suitable feeding site on the host (Appetence, Engagement and Exploration) which is followed by penetrating the host's epidermis to access the blood meal. During attachment, Ixodid ticks secrete a cement-like proteinaceous substance that hardens around the inserted mouthparts to secure the parasite to the host (Reviewed by Anderson and Magnarelli, 2008).

Upon initiation of ingestion, ticks secrete analgesic saliva, containing pharmacologically active molecules (Reviewed by Francischetti *et al*., 2010), into the feeding pool to suppress or counteract host immunological and haemostatic defences (Maritz-Olivier *et al*., 2007). The latter is essential to maintain the fluidity of the blood meal by preventing blood coagulation. Females engorge themselves on a blood meal before detaching from the host to drop to the ground where they lay their eggs in the soil or under leaf litter. (Reviewed by Anderson and Magnarelli, 2008).

The importance of ticks and tick control

Due to the search for better livelihoods, migration from traditional farming areas into urban areas has been increasing in frequency worldwide since the industrial revolution. This results in the loss of arable land and fewer people relying on their property as a source of income and nourishment; depending instead on state food production, and accessing mass-produced sustenance from the commercial sector (FAO, 2018, 2017; United Nations, 2018).

While the global population growth rate is declining, that of low-income countries such as those in Asia and Africa are increasing. It is projected that the combined populations of Asia and Africa, already rife with hunger and malnutrition, will constitute more than 80% of the global population by 2100 (FAO, 2018, 2017). Increasing urban communities means that the demand for animal-derived products is out-pacing supply capabilities. Consequently, food production needs to shift from low intensity to very high-intensity agriculture (FAO, 2018, 2017; United Nations, 2018). For high-intensity agriculture to be feasible, livestock must be healthy (Jeanmonod *et al*., 2018). Livestock well-being is directly influenced by ticks and tick-borne diseases that result in significant losses in animal-derived products such as meat, milk and other downstream by-products like leather and fertilisers. The sectors of livestock farming in

South Africa that are profoundly affected by high tick burdens are the cattle and cattle-related industries (Bigalke 1980; Directorate Agricultural Statistics 2010).

Ixodid ticks are considered of great economic importance to the cattle industry across the globe as they are the most common vectors for debilitating tick-borne cattle diseases like Theileriosis, Anaplasmosis and Babesiosis (Derso and Demessie, 2015; Guerrero *et al*., 2006; Hurtado and Giraldo-Ríos, 2018; McLeod and Kristjanson, 1999). Global economic losses are estimated to be in the range of billions of US dollars annually (de Castro, 1997; Hurtado and Giraldo-Ríos, 2018). While cattle losses are devastating for the commercial sector, the loss is particularly damaging to subsistence farmers who may lose their entire herd rapidly due to Theileriosis, Anaplasmosis or Asiatic Babesiosis (Derso and Demessie, 2015; Ndhlovu *et al*., 2009; Sungirai *et al*., 2016).

Aside from the transmission of disease(s), ticks also impact cattle in a direct, mechanical manner. Direct damage occurs through the ticks' attachment and feeding habits that result in damage to hides, udders and the genitalia of infested animals (de la Fuente *et al*., 2015; Jeanmonod *et al*., 2018; Ndhlovu *et al*., 2009). Furthermore, tick infestation can result in myiasis due to secondary infection of the wounds left by ticks. These secondary infections can include the growth of maggots and secondary microbial infections in live animals (Cheng, 1986; de la Fuente and Contreras, 2015; Jeanmonod *et al*., 2018; Ndhlovu *et al*., 2009). Treatment of these secondary infections typically require drugs which hinder farmers from entering high paying markets which are moving towards antibiotic/drug-free products (Makary *et al*., 2018). Lastly, damage caused by secondary infections often render hides unusable; impacting the leather industry.

Currently, the South African cattle industry contributes as much as 12.1% of the country's gross agricultural product, with around 14 million head of cattle amounting to a gross income value of over R 30.6 billion in 2015/16 (Calculations based on reports by DAFF, 2016, 2017 and Statistics South Africa, 2016). Also, the beef and dairy sectors contribute significantly to the socio-economic development of the country via job creation. In 2016, 25% of South African agricultural households were reported to be involved in the cattle industry (Statistics South Africa, 2016), with over 2 million people (around 4% of the total South African population) depending solely on the beef sector for their livelihoods (DAFF, 2017). Revenue losses due to tick-borne diseases in cattle were estimated to be R 70-200 million per year in the 1980s (Bigalke, 1980), but this figure has since increased; with annual losses due to tick-borne diseases on

beef production alone predicted to range between R 1.3 and R 3.1 billion (Directorate Agricultural Statistics 2010; Oberholster 2014, unpublished).

Any adverse effect on the agricultural sector (such as ticks) resulting in lowered access to quality cattle foodstuffs, impacts negatively on nutritional status and therefore, the nation's development and socio-economic status. Improved nutrition correlates with improved health and thereby increased life expectancy; indeed an increase of 1% in life expectancy correlates with a 6% increase in total GDP and 5% increase in GDP per capita even in high-income countries (Swift, 2011). It is, therefore, of vital economic importance that effective methods of tick control are implemented.

Rhipicephalus microplus *is an economically significant tick species*

There are several endemic and invasive *Rhipicephalus species* of importance in the South African cattle industry. Endemic species include *Rhipicephalus decoloratus* and *Rhipicephalus appendiculatus* that transmit *Babesia bigemina*, and *Theileria parva,* respectively while the invasive Asian blue ticks, *Rhipicephalus microplus*, transmit the highly pathogenic *Babesia bovis* in addition to *Babesia bigemina*, and *Anaplasma spp*. and forms the focus of this study (Horak *et al*., 2015).

Rhipicephalus microplus occurs mainly in temperate, sub-tropical regions around the globe, but have been reported to be unique in their ability to adapt to new environments (Nyangiwe *et al*., 2018); from the harsher desert climates of Namibia (Nyangiwe *et al*., 2013) and Mali (Adakal *et al*., 2013) to the more tropical climates of Burkina Faso, Togo (Adakal *et al*., 2013), Tanzania (Nyangiwe *et al*., 2018), Benin (de Clercq *et al*., 2012), Madagascar (Pothmann *et al*., 2016), Nigeria (Kamani *et al*., 2017) and most recently, Cameroon (Silatsa *et al*., 2019). One-host ticks such as *R. microplus* are found on a single host: feeding, moulting and maturing on the same species (Anderson and Magnarelli, 2008). The need for only one host enables these ticks to have a shorter generation time than multiple-host ticks, allowing them to out-compete multi-host species in the same area.

In South Africa, *R. microplus* was first recorded in King Williams Town in the Eastern Cape by C. W. Howard in 1908. More recently, Nyangiwe *et al.* 2017 documented the distribution of *R. microplus* across South Africa in 51 locations, specifically, the Eastern, Western and Northern Cape, as well as in the Free State (Figure 1.2). They found that *R. microplus* is sympatric in 50% of the localities with *R. decoloratus* and that *R. microplus* was present in more than 80% of the locations while the endemic *R.*

decoloratus species is present in less than 60% of the localities in South Africa. Moreover, Nyangiwe *et al*. 2017 observed larvae displaying morphological characteristics of both species. They, therefore, postulated that some level of hybridisation between the two species occurs despite the genetic incompatibility seen in the *R. decoloratus* and *R. microplus* breeding studies of Spickett and Malan (1978). However, this hypothesis lacks molecular confirmation and therefore requires corroboration beyond morphological observations.

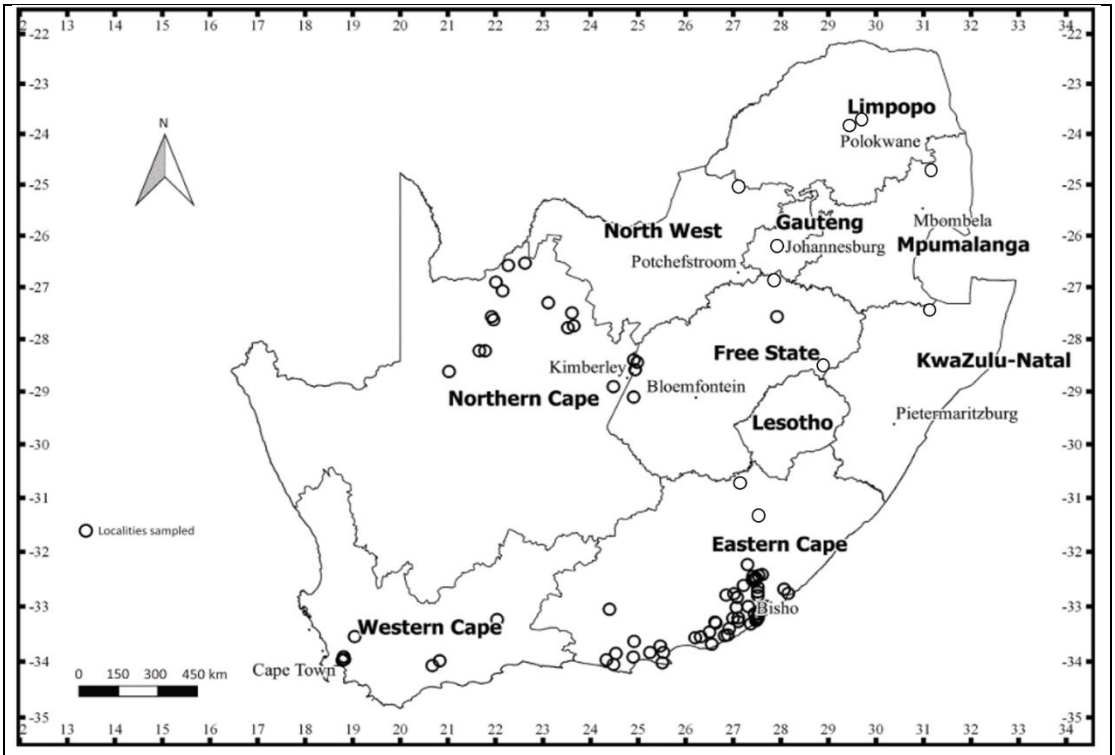


Figure 1.2: Map of areas (Clear circles) sampled in distribution studies of *R. microplus* and *R. decoloratus* adapted from Terkawi *et al*. 2011 and Nyangiwe *et al*. 2017. *R. microplus* and *R. decoloratus* have been identified in areas in each province with more than 80% of the sampling being identified as *R. microplus* in each locality.

Current strategies for the control of ticks and tick-borne disease

There is no cure for many tick-borne cattle diseases, but control measures have been developed which have not changed much since their inception in the 1900s (Kocan *et al*., 2003). Antibiotics are the most generic form of treatment for pathogenic infections (such as Anaplasmosis), however antibiotic resistance is rapidly increasing in many pathogens (de la Fuente *et al*., 2007; Dumler, 2005; He *et al*., 1999; Potgieter, 1979).

A simplified way to prevent disease transmission may be to avoid the initial exposure of cattle to the parasites that transmit the pathogens.

The most prevalent means of tick-prevention rely on chemical control measures using acaricides, which are applied via full-body dipping, spraying or pour-on technologies (Sungirai *et al*., 2016), and some are even available as oral medications (Davey *et al*., 2001). Acaricides are the first choice in tick control but ticks rapidly acquire resistance to these chemicals (Schetters *et al*. 2016). Recent surveys on South African tick populations of *R. microplus* demonstrated an alarming level of up to 60% resistance against pyrethroids nationally and 25% and 55% resistance to amitraz across the coastline of South Africa and Mpumalanga, respectively (Baron *et al*., 2018, 2015; Robbertse *et al*., 2016).

Acaricides are unsustainable as they are a source of contamination in the environment and cattle-derived foodstuffs (de la Fuente *et al*., 2007; Kocan *et al*., 2003; Nijhof *et al*., 2007). In regards to the former, there are a variety of adverse effects; first, acaricides can cause damage to entomopathogens (often a fungus) which could have otherwise worked synergistically with regards to the control of parasites (Ferreira *et al*., 2016; Vieira *et al*., 2016). Second, pyrethroid exposure may result in alterations of gut microbiota, especially for mammals in the postnatal period (Nasuti *et al.*, 2016). Third and finally, the widespread use of pesticides, including acaricides results in the leakage of, for example, pyrethroids, into the soil; this has been seen to alter the microbial composition of the affected soil resulting in related consequences for the ecosystem of the area (Qi and Wei, 2017).

Acaricides can also leave toxic residues in cattle foodstuffs such as beef and dairy (Macedo *et al*., 2015; Picinin *et al*., 2016; Qin *et al*., 2017). These residues are dangerous for human health; one example is macrocyclic lactones which target the Gamma-Aminobutyric Acid (GABA) receptors, may cross the blood-brain barrier and may result in various neuropathic symptoms, coma and even death at high doses (Yang, 2012). Another example is pyrethroids which may disrupt the endocrine, reproductive and immune system and have also been linked to breast cancers (Reviewed by Thatheyus and Gnana Selvam, 2013).

The development of new and less environmentally toxic acaricides is a lengthy and expensive process that to date has not provided promising next-generation acaricides. However, Acari growth regulators such as fluazuron (Acatak) (Kemp *et al*., 1990) have been used with success in areas of resistance (Bull *et al*., 1996), but they are often too expensive for small-scale and rural farmers, and there have been reports of resistance

emerging to even these (Reck *et al*., 2014). Recent advances have been made in the development of active compounds in herbal remedies (Amaral *et al*., 2017; Anholeto *et al*., 2017; Fouche *et al*., 2016; Singh *et al*., 2016) and nanoparticles (Gandhi *et al*., 2017; Marimuthu *et al*., 2013) as acaricides and/or tick-deterrents but none have yet had a lasting impact.

Using vaccines to combat ticks reduces environmental contamination caused by acaricides and antibiotics while also limiting the development of resistance among pathogens and ticks alike (Willadsen, 2008). The inclusion of vaccination in a combinatorial approach (with acaricides, grazing rotation and ethical farming practices) to tick control has been proposed since 1939, and the applicability of this form of control in the field has been corroborated since 1979 (Allen and Humphreys, 1979; Trager, 1939). Additionally, vaccines offer a complementary tool to alleviate the selection pressure on chemical acaricides (de la Fuente *et al*., 2007; Vercruysse *et al.*, 2007). One example of the latter is the use of GAVAC® in South America that reduced the use of acaricides by 83.7%, saving some 81.5% on costs while in Cuba some 60% reduction in acaricide usage was observed (de la Fuente *et al*., 2007; de Miranda Santos *et al*., 2018). Tick vaccines hold the promise of reducing tick fecundity on several levels; such as tick attachment, feeding and engorgement weight, oviposition and viability of offspring. If successful, this will minimise tick load and indirectly reduce the transmission of tick-borne diseases.

In the field of ectoparasites, there are two types of antigens currently considered to be protective (De Vos *et al.* 2001; Trimnell *et al.* 2002; Nuttall *et al.* 2006). The first type, known as exposed antigens, comes into direct contact with the host's immune system. These exposed antigens are mostly the secreted saliva proteins that enter the host during feeding. Vaccines that target only exposed antigens have not yet been shown to be effective, presumably due to the array of salivary antigens from similar protein families, where the biological role of a single antigen could be 'rescued' or compensated for by other homologous proteins secreted into the feeding pool (de la Fuente *et al*., 2007; Schetters *et al*., 2016; Stutzer *et al*., 2018). Also, salivary antigens have co-evolved in the tick-host interface for millions of years and may not be as immunogenic and lethal to the tick upon vaccination (Chmelař *et al*., 2017; Mans *et al*., 2017; Šimo *et al*., 2017).

The second type of antigens is termed 'concealed antigens' that are not directly exposed to the host immune system during regular feeding and are usually found in the tick gut and ovaries (Trimnell *et al*., 2002). To date, it is believed that the most

effective vaccines will be those where several targets are vaccinated for at the same time, resulting in improved host immunity and vaccine efficacy (Rodríguez *et al*., 1994; Schetters *et al*., 2016; Stutzer *et al*., 2018).

Currently, the primarily known tick vaccines that are available to the market are Gavac™ (Heber Biotec S.A., Havana, Cuba), Go-Tick/Tick-Vac®, sold by Limor de Colombia®, SA and MK Tecnoquimicas, and Bovimune Ixovac® sold in Latin America (Lapisa S.A., La Piedad, Mexico). The antigen in Go-Tick is not disclosed while the newer Bovimune Ixovac® is based on a whole larval extract from *R. microplus* and the more well-known and widespread Gavac™ is based on the midgut cell membrane protein called Bm86. This antigen is a glycosylphosphatidylinositol (GPI) linked protein with a molecular mass of ~89 kDa and was identified from the *R. microplus* tick species by Willadsen *et al*. (1989). The 1980's study was conducted in Australia, and the Australian *R. microplus* has since been reassigned to *R. australis* (Estrada-Peña *et al*., 2012); so the species from which the original antigen was isolated is now unclear.

With regards to control of *R. microplus* via vaccination, geographical strains have been collected that are less susceptible to immunological intervention via vaccination with the best-known vaccine antigen currently commercialised, Bm86 (Ali *et al*., 2016; de la Fuente *et al*., 2007; Schetters *et al*., 2016). Combined, the ability of *R. microplus* to adapt to new climates, acaricide resistance and lack of control via vaccines highlights the urgent need for new strategies to combat *R. microplus* infestation. (Fyumagwa *et al*., 2009; Nyangiwe *et al*., 2013; Portillo *et al*., 2007).

Bm86 as a protective antigen against *R. microplus*.

Bm86 background

Willadsen *et al*. (1989) identified an antigenic protein that lies on the plasma membrane of tick midgut digest cells, called Bm86. Bm86 was initially reported as a 650 amino acid protein with 6 repeated cysteine residues and includes a 19 amino acid long signal sequence, in addition, it was indicated to have transmembrane regions during the immature life stages which were thought to be replaced by a GPI-anchor in the adult life stage (Rand *et al*., 1989). Domain analyses of Bm86 indicated multiple epidermal growth factor domains which were confirmed by Nijhof *et al.* (2010) from full-length amplicons containing eight full and one partial Epidermal Growth Factor (EGF) domains, a GPI-anchor and the signal peptide along with four N-linked carbohydrate positions and two O-linked carbohydrate positions (Figure 1.3). Currently, there are close to 200 sequence hits for Bm86 in arachnids of GenBank.

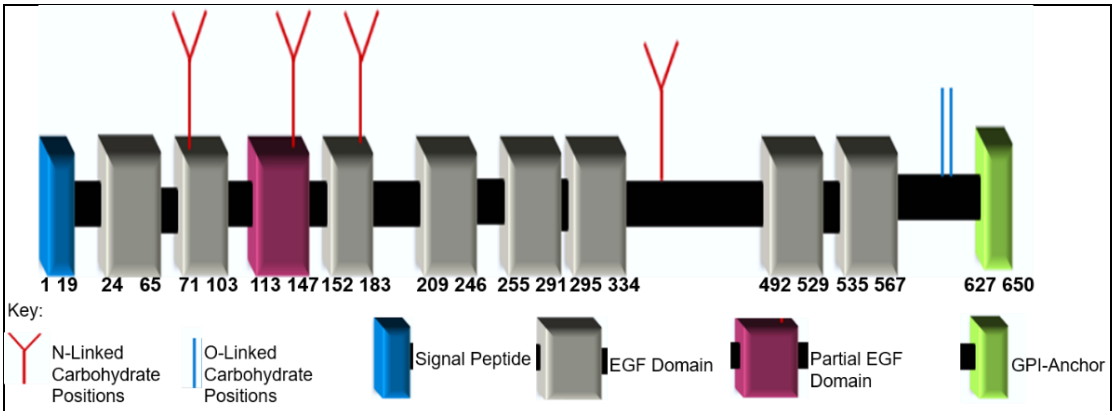


Figure 1.3: The predicted structure of the Bm86 protein. Adapted from a study by Nijhof *et al*. (2010).

The Bm86 gene is differentially expressed in all life stages and tissues. Additionally, each species has a unique expression profile for Bm86 with the highest expression occurring in the adult female gut (Bastos *et al*., 2010; Nijhof *et al*., 2010).

Bm86 is a member of a large group of similar proteins. One homolog of Bm86, Bm95, was sequenced from the Argentinean-A strain of *R. microplus* in Argentina that is less susceptible to vaccination (García-García *et al*., 2000). Bm95 was shown to contain a GPI-anchor and 7 EGF domains, and like Bm86 had N- and O-glycosylation patterns (González et al., 2004). The Bm95 gene has at least two alleles that differ; in that one allele produces a truncated version of the protein, and that this truncated version lacks a transmembrane/GPI-anchor region while the full-length version is identical to the

Bm86 sequence isolated from the same strain. However, when compared to the Cuban Camcord strain Bm86 sequence used in the GAVAC™ vaccine, it showed many amino acid differences in both alleles (García-García *et al*., 2000).

Vaccination trials that included Bm95 showed a similar result to those for the GAVAC™ vaccine on susceptible tick populations. However, on the 100% GAVAC™ resistant *R. microplus* strains, vaccination with the strain-specific Bm95 antigen showed improved efficacy of 58% (García-García *et al*., 2000). Therefore it was considered for inclusion in a multi-factorial vaccine aimed at the specific resistant strain, i.e. GAVAC^{plus} (García-García *et al*., 2000; González *et al*., 2004; Kumar *et al*., 2009; Merino *et al*., 2013).

Additionally, Bm86 shows a broad range of variability across geographical regions, even within a country; exemplified by the numerous Bm86 sequences found in Texas and Thailand alone (Freeman *et al*., 2010; Kaewmongkol *et al*., 2015). A study, conducted by Hüe *et al.* (2017) on *R. australis*, a close cousin of *R. microplus*, found a sequence variation that contains an entire region within Bm86 that is missing, possibly due to alternative splicing. They also found two alternative (Full and alternatively spliced) Bm86 sequences within the same *R. australis* strain.

Nijhof *et al.* (2010) also identified possible Bm86 orthologs in various tick species, as well as a structurally-related protein, BmATAQ, which was found in the Metastriate tick family. This distinction may be indicative of separate Bm86 family gene duplication events in the two families. One ortholog of Bm86 and most BmATAQ orthologs that were identified by Nijhof *et al.* (2010) contain a transmembrane region in place of the characteristic GPI-anchor. The BmATAQ protein is also variably expressed in tick tissues and life stages across species (Nijhof *et al*., 2010).

Despite the extensive research conducted on Bm86 sequence variation and expression, the normal function of the Bm86 protein in the adult tick midgut is currently still unknown, and the precise sequence for Bm86 in each life stage and tissue has not been investigated. Nevertheless, based on analogous structure comparisons, it has been suggested that Bm86 may be involved in cellular communication and signalling (Liao *et al*., 2007; Rand *et al*., 1989).

Bm86 as a protective antigen

Most, if not all, tick vaccines to date have been based on the Bm86 protein which was found to elicit a protective immune response in cattle (Kemp *et al*., 1986; Rand *et al*., 1989; Willadsen *et al*., 1989). The mechanism of action is proposed to entail uptake of

anti-Bm86 antibodies from the vaccinated host upon tick feeding and subsequent binding of the antibodies to Bm86 in the tick gut; resulting in reduced endocytosis (Willadsen *et al*., 1989) and damage to (Kemp *et al*., 1986) digestive gut cells; causing the typical "red phenotype" due to blood leaking into the tick body cavity.

The most significant effect of exposure to vaccinated animals is the reduced fecundity of female ticks, with an overall decrease in their egg-laying abilities. In some studies, it was reduced by as much as 77% (Rand *et al*., 1989; Willadsen *et al*., 1989). However, efficacy varied in field studies conducted in different geographical areas (Richards *et al*., 2015). Additionally, vaccination has little to no effect on the immature larvae, with the most pronounced impact achieved only in adults (Kemp *et al*., 1986; Richards *et al*., 2015).

Vaccines containing the Bm86 protein as an antigen were developed to target *R. microplus* specifically, but also displayed some cross-protection against other tick species (De Vos *et al*., 2001; Derose *et al*., 1999; Riding *et al*., 1994). The GAVAC™ vaccine has been seen to be most effective (100%) against *Rhipicephalus annulatus* (Miller *et al*., 2012) while Tick-GARD was most effective against the *R. annulatus* larvae (Pipano *et al*., 2003), while Bm86-homologs and derivatives have been most effective when used against their target tick strain (reviewed by Richards *et al*., 2015).

The Bm86 vaccine is currently available as a full-length recombinant antigen (Schetters *et al*., 2016). The full-length antigen has been shown to have variable results, with efficacies ranging from 25% to 100% in *R. microplus* (de la Fuente *et al*., 2007; Richards *et al*., 2015; Sonenshine and Roe, 2014) (Table 1.1). This high variability in the protective response has been linked to the way the Bm86 antigen is produced. Trials in which cattle were vaccinated with Bm86 produced in *Escherichia coli* had reduced efficacy and did not exhibit the typical red phenotype, while the effectiveness increased in experiments with Bm86 from a yeast production and these exhibited the red phenotype (Rodríguez *et al.* 1994; Ferreira 2015, unpublished; Sikhosana 2017, unpublished; Stutzer *et al.* 2017, unpublished).

Freeman *et al*. (2010) found significant sequence variations in the Texas outbreak strain Bm86 when compared to the Cuban Camcord strain Bm86 that was used in the GAVAC™ vaccine. They hypothesise that the discrepancy in vaccine efficacy obtained in field trials was influenced by strain-specific diversity in the Bm86 sequence found in a specific area. Evidence for this was seen in a preliminary study by García-García *et al*. (1999) who noted an inverse relationship between Bm86 sequence variation and vaccine efficacy. However, these findings have not been further validated.

Table 1.1: Efficacies of commercialised Bm86 vaccination against *R. microplus* in cattle field trials. Gavac™ is the most widely used with a range in the efficacy of 56% across regions. Coloured according to country.

Vaccine name	Location of the trial	Efficacy (%)
GAVAC™ (Bm86 Based) Adapted from Sonenshine and Roe (2014) and de la Fuente *et al*. (1999)		
Gavac™	"Limonar", Matanzas, Cuba	81
	Cuba	60
	"Los Naranjois", Havana, Cuba	68
	Pinar del Rio, Cuba	53
	Cuba	87
	"Fazenda Restgate", Sao Paulo, Brazil	79
	"Fazenda Restgate", Sao Paulo, Brazil	96
	"Fazenda Restgate", Sao Paulo, Brazil	97
	Barra Mansa, Rio de Janeiro, Brazil	55
	Corrientes, Argentina	55
	Colombia	65
	Colombia	**44** Lowest Efficacy
	Colombia	80
	Doima, Colombia	72
	Santa Cruz, Ilbagué, Colombia	77
	"Peregrino", Tamaulipas, Mexico	**100** Highest Efficacy
	"Kikapu", Chiapas, Mexico	97
	"Tixtla", Tamaulipas, Mexico	67
Other Bm86 Vaccines Adapted from (Andreotti *et al*., 2018; Stutzer *et al*., 2018)		
TickGARD® (**Discontinued**)	Various, unspecified (Australia)	20-30
TickGARD®PLUS (**Discontinued**)	Unspecified (Australia)	56
	"Mato Grosso", do Sul, Brazil	46.4

The difference in efficacy of Bm86 vaccines may complicate vaccination strategies; as antigenic variation in ticks of a specific area may have to be determined during vaccine development stages. Thus, each geographical area could well require tailored vaccines, which are particular to the ticks in the area. *In silico* analyses by Blecha *et al*. (2018) of Bm86 antigenic regions across known sequences for Bm86 and homologs thereof emphasises the relationship of polymorphisms in Bm86 sequence to vaccine efficacy. Furthermore, the study was able to identify that there are conserved antigenic regions among Bm86 sequences and it was postulated that a universal Bm86 sequence containing these common antigenic regions might be synthesised for use as a universal tick vaccine (Blecha *et al*., 2018). However, it is the opinion of this researcher that there is not enough sequence knowledge for Bm86 across tick life stages or tissues since almost all known Bm86 sequences have been identified from

the adult female gut. For the current study using the South African *R. microplus* strain, it is therefore of utmost importance that any Bm86-based vaccine formulation contains all the correct sequences for this antigen identified from each life stage of the parasite.

Insights into the biological function of Bm86

Recently, a yeast-two-hybrid study done by Kiper (2013, invention disclosure) in our research group identified several proteins binding to Bm86. One of the Bm86 interacting proteins is a novel Kunitz protein. Kunitz domains are known to function in the inhibition of an array of proteases in several diverse metabolic pathways. Based on transcriptome data, this Kunitz binding protein (termed Antigen 1) is predicted to be secreted in gut tissues with an unknown biological function (Kiper 2013; invention disclosure).

Small scale vaccination trials with a combinatorial vaccine containing both Bm86 and Antigen 1 showed an increase in the vaccine efficiency compared to using single antigens (Ferreira 2015, unpublished). To date, the tissue distribution and sequence variations of Antigen 1, as well as the regions essential for Antigen 1 interaction with Bm86, remain to be elucidated. Like Bm86, it is expected that that Antigen 1 may also show strain-specific sequence variation and insight into this diversity will be essential to develop a protective vaccine.

A cDNA microarray study was conducted comparing differentially expressed transcripts between ticks fed on Bm86 vaccinated and non-vaccinated cattle to aid in the understanding of the biological function of Bm86 (James 2017, unpublished). Transcripts indicating proteins that may function in the Phospholipase C (PLC) signalling pathway, secreted proteins involved with digestion, as well as proteins involved in endo and exocytosis were found to be significantly down-regulated in ticks fed on Bm86 vaccinated cattle. Using this data, we proposed a hypothetical model for the biological role of Bm86 (Figure 1.4).

Briefly, it was hypothesised that the Bm86-mediated signalling pathway (Figure 1.4 in the Red box) is initiated by binding of a yet unknown ligand which triggers the dimerization of Bm86. Dimerization leads to the downstream activation of a G protein-coupled Phospholipase and the subsequent production of Inositol Triphosphate 3 (IP_3) and release of Ca^{2+} ions.

Since Bm86 is a GPI-anchored protein, it may be necessary for intermediate proteins, such as a GTP-binding (G) protein, to facilitate interaction with a PLC. However, not

all Bm86 orthologs contain GPI-anchors (Nijhof *et al*., 2010) which points towards different functions for Bm86 in the various life stages of *R. microplus*.

The hypothesis is supported by data showing that PLC is a substrate for EGF receptors, and indeed, EGF receptors containing transmembrane regions have been seen to be able to engage in autophosphorylation and directly interact with, and phosphorylate PLC_{γ} (Goldschmidt-Clermont *et al*., 1991; Meisenhelder *et al*., 1989; Vega *et al*., 1992). In this way, EGF domains are implicated in PLC mediated signalling. Several EGF domain-containing proteins and GPI-anchored proteins have also been indicated to form protein dimers upon ligand binding, affording signal transduction (Ullrich and Schlessinger, 1990; Wang *et al*., 2002; Wee and Wang, 2017).

It is therefore hypothesised that Bm86, with multiple EGF domains and a GPI-anchor, is a signal mediator for the activation of pathways mediating secretion of bioactive compounds in gut tissue. Maritz-Olivier *et al.* (2005) previously investigated the signalling pathways regulating exocytosis in the salivary glands of the argasid tick, *Ornithodoros savignyi*, and found them to be similar to those described for the Ixodidae by Sauer *et al.* (2000). In both studies, dopamine was shown to interact with its membrane receptor, coupled to a stimulatory G protein, which activates PLC and leads to a cascade of reactions underlying exocytosis of tick saliva. The ligand stimulating secretion of digestive enzymes from tick midgut tissue remains to be discovered, but it is proposed that a similar calcium-dependent exocytosis pathway may occur in tick midgut tissue.

Apart from identifying components of the PLC pathway, the microarray study by James *et al*. (2017, unpublished) also identified several upregulated genes that may play a role in a Jun/Jun or Jun/Fos compensatory signalling pathway. This upregulation led to the increased production and exocytosis of Serine carboxypeptidases and other uncharacterized tick-specific secreted peptides. These are proposed to allow the tick gut to compensate for the loss of Bm86 function during feeding on vaccinated hosts (Figure 1.4, indicated in blue).

The sections which follow will discuss the mentioned pathways and proteins involved therein as they form some base of the hypotheses under investigation in this study.

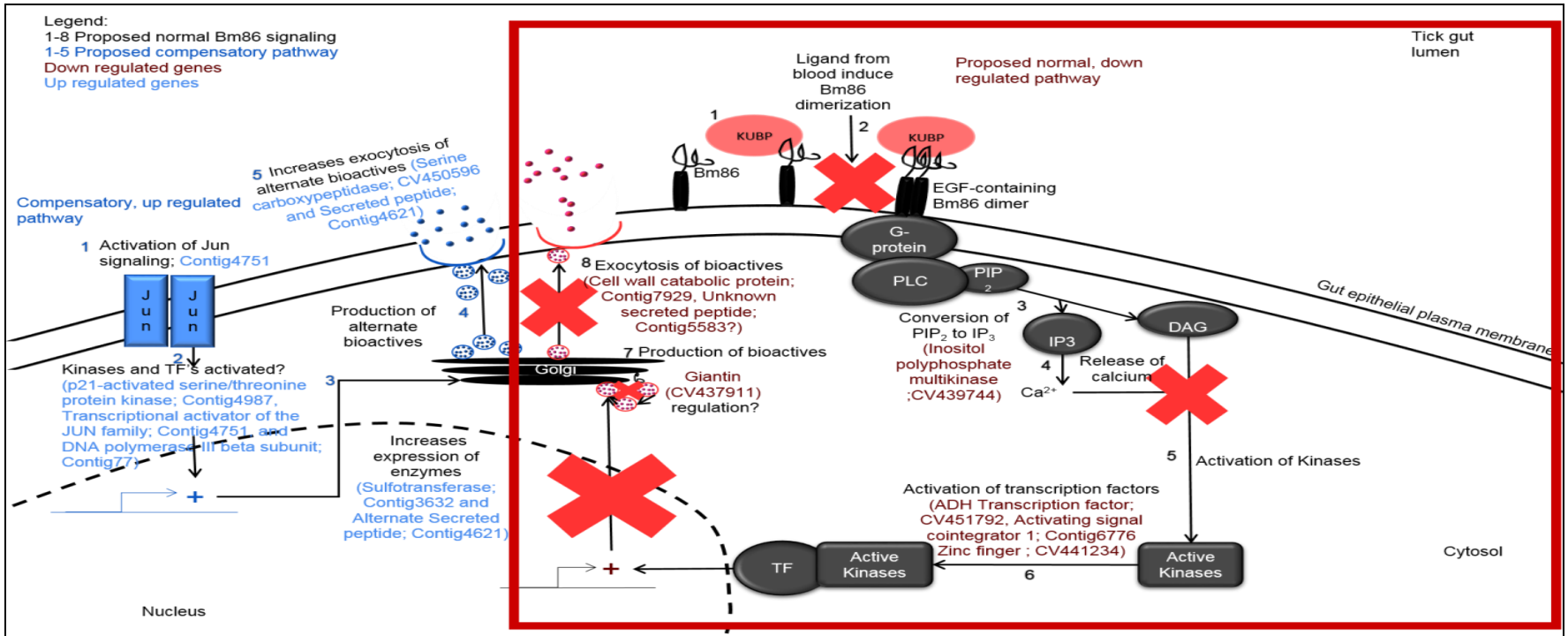


Figure 1.4: Proposed mechanism of Bm86 signalling (Black), down-regulation effects of the vaccine (red) and the proposed up-regulated compensatory mechanism (Blue); numbers 1-8 in black indicate the proposed normal Bm86 signalled pathway, in its native state Bm86 interacts with Antigen 1 (1). Signalling is initiated by ligand binding (2) which triggers dimerization of Bm86, leading to the production of IP_3 and release of Ca^{2+} ions (3-4) which activate kinases, which in turn activate transcription factors to initiate transcription of bioactive molecules (5-7). It is also further proposed that biomolecule production is regulated via a giantin-like modulated mechanism in the Golgi (7) before exocytosis (8). This normal pathway is hypothesised to be inhibited by antibodies imbibed by the tick when feeding on a Bm86 vaccinated animal (red). To survive it is proposed that ticks may employ a Jun/Jun or Jun/Fos compensatory signalling pathway, leading to the increased production and exocytosis of Serine carboxypeptidases and other secreted peptides (1-5 in blue). The red box indicates the focus of this study based on the results that led to the production of Figure 1.4 as one possible mechanism of action of Bm86.

Role of Glycophosphatidylinositol-anchored proteins

Background to GPI-anchors: structure and function

The GPI-anchor is a posttranslational modification of many cell surface proteins Figure 1.5) that attaches proteins to the outer lipid layer of the plasma membrane (Kinoshita and Fujita, 2009). Generally, the GPI-anchor consists of a lipid part that is either phosphatidylinositol (PI) or inositol phosphoceramide (Madore, 1999; Pike, 2004), which can be recognised by enzymes mediating cleavage of the extracellular protein from the membrane to yield a soluble protein (Figure 1.5).

GPI-anchored proteins are not transmembrane proteins, but due to alternative splicing, can exhibit transmembrane, soluble or the GPI-anchored forms depending on the tissue and context in which they are expressed (Saha *et al*., 2016). Many GPI-anchored proteins reside in lipid rafts on the plasma membrane and exhibit lateral motility along the membrane bilayer allowing for protein dimerization and ligand binding (Ishihara *et al*., 1987).

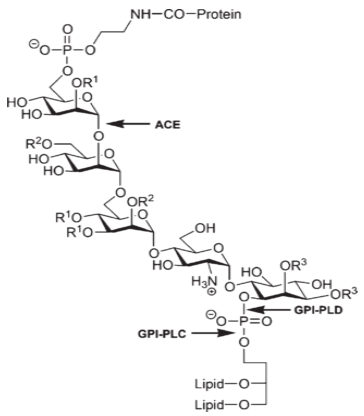


Figure 1.5: The general chemical structure of a GPI-Anchor. Arrows indicate extracellular cleavage sites for angiotensin-converting enzyme (ACE), GPI-Phospholipase C (PLC) and GPI-Phospholipase D (PLD). The lower two lipid chains (shown by Lipid) are embedded in the outer layer of the plasma membrane. Taken from Lauc and Heffer-Lauc (2006).

GPI-anchored proteins have various roles in cellular processes; most notably in cell signalling and adhesion (Saha *et al*., 2016). Furthermore, GPI-anchored proteins function as surface hydrolases, coat proteins, protozoan antigens, toxin binders, receptors, and are vital for embryonic viability (Reviewed by Paulick and Bertozzi (2008) and Saha *et al*. (2016)).

GPI-anchored proteins can be released from the surface of a cell in exosomes or small aggregates with the GPI-Anchor intact through either cleavage by GPI-anchor specific Phospholipase C (GPI-PLC), GPI-anchor specific Phospholipase D (GPI-PLD) or

angiotensin-converting enzyme (ACE). The released protein can then be incorporated into the surface of another cell by direct insertion of the lipid anchor (when released as small aggregates) through a process dubbed “cell surface painting” and the newly incorporated protein is still fully functional (Reviewed by Medof *et al*. (1996) and Lauc and Heffer-Lauc 2006).

This ability of GPI-anchored proteins to relocate and be reintegrated into cell membranes has been proposed for the experimental modification of the extracellular cell surface, in place of traditional gene transfer and expression techniques (Medof *et al*., 1996). Therefore, shedding and uptake of GPI-anchored proteins between cells may confer new abilities on the host cell. An example of this phenomena, attributed to the GPI-anchoring, was Decay-Accelerating Factor (DAF) that prevented haemolysis of rabbit erythrocytes by the human DAF complement and was therefore proposed as a likely mechanism involved in the inhibition of the complement system (Medof *et al*., 1984).

Tumour cells have been seen to suppress the immune response through the shedding and uptake of GPI-anchored proteins, and retroviruses have been seen to exploit this process for replication (Reviewed by Lauc and Heffer-Lauc, 2006). It is, however, not clear what role this process fulfils in healthy cells. It has been proposed that it may be employed by spermatozoa to populate their membranes with proteins and that in other cells it may serve a role in lipid raft modulation, signal transduction modulation and regulation of immune response systems (Reviewed by Lauc and Heffer-Lauc, 2006).

Many GPI-anchored proteins act as receptors for extracellular ligands, and thus function in cell communication and signalling in response to ligand binding (Robinson, 1991; Solomon *et al*., 1996). In mammalian lymphocytes, GPI-anchored proteins are well known as signal transduction proteins (Usually the GPI-linked Thy-1 antigen), where they modulate the immune response (Robinson, 1991; Solomon *et al*., 1996). In neuronal cells, for example, Contactin acts as a recognition molecule involved in cell-cell communication during neurite outgrowth and development (Peles *et al*., 1997, 1995).

In arthropods, GPI-anchored proteins have been seen to function in various signalling pathways, as established from studies conducted on *Drosophila.* Table 1.2. provides examples such as the highly conserved Multiple Inositol Polyphosphate Phosphatase (MIPP-1) isoform in *Drosophila* identified by Cheng and Andrew (2015) which may signal for tracheal cell migration and branch elongation by converting extracellular inositol polyphosphates (IP6, IP5, and IP4) to IP3.

Table 1.2: Arthropod GPI anchored proteins for which reviewed sequences are available. NI indicates when there is little/no definitive information found. (Continues on next page).

GPI-Anchored Proteins in Arthropods					
Protein	**Species**	**Biological Function(s) ("!" Indicates predicted functions)**	**Possible Dimerization**	**Protein-protein interactions ("#" indicates proposed interactions)**	**Reference**
Bm86	*R. microplus*	NI	Homodimerization	Antigen 1[#]	(Kiper, 2013; Willadsen *et al*., 1989)
5'-nucleotidase	*R. microplus*	Degradation of UDP-glucose to uridine monophosphate & glucose-1-phosphate	Homodimerization	NI	(Field *et al*., 1999)
Multiple inositol polyphosphate phosphatase 1	*D. melanogaster*	Dephosphorylation of IP4,5 and 6 to form IP3; facilitate filopodia formation during embryonic tracheal tube elongation.	NI	NI	(Cheng and Andrew, 2015; Chi *et al*., 1999; King *et al*., 2010)
Alkaline phosphatase 4	*D. melanogaster*	Dephosphorylation of pyridoxal-5'-phosphate to pyridoxal & metabolism of calcium phosphates. Transmembrane transport of calcium[!]; Cell differentiation and proliferation in neuronal tissue[!]	Homodimerization	GTP binding proteins; cAMP response element-binding protein	(Torriani, 1968; Yang *et al*., 2000)
Membrane-bound alkaline phosphatase	*Bombyx mori; Anopheles aegypti*	Phosphatase; CRY-toxin receptor	NI	Ribosomal proteins L23A, P0; 60S ribosomal protein L5 and 6-pyruvoyltetrahydropterin synthase	(Itoh *et al*., 1991; Thammasittirong *et al*., 2011)
Acetylcholinesterase	*D. melanogaster; Leptinotarsa decemlineata; Anopheles stephensi*	Hydrolyses choline in synapses	Homodimerization	Lipophorin; cAMP-dependent protein kinase R1; Translational regulator (pumilio)	(Baines, 2003; Chen *et al*., 2008; Fournier *et al*., 1988; Harel *et al*., 2000; Panáková *et al*., 2005); (Zhu and Clark, 1995); (Hall and Malcolm, 1991)
Lazarillo	*Schistocerca americana*	Signalling role in neuronal development[!]	NI	NI	(Ganfornina *et al*., 1995)
Dalley Protein	*D. melanogaster*	Morphogen signalling, JAK/STAT signalling pathway	NI	Lipophorin; Hedgehog; Wingless;	(Eugster *et al*., 2007; Hayashi *et al*., 2012; Nakato *et al*., 1995)
Gram-negative bacteria-binding protein 1	*D. melanogaster*	Toll signalling activation	NI	Peptidoglycan recognition protein SA; Gram-negative bacteria binding protein 1	(Jang *et al*., 2006; Kim *et al*., 2000; Wang *et al*., 2006)

GPI-Anchored Proteins in Arthropods					
Protein	**Species**	**Biological Function(s) ("!" Indicates predicted functions)**	**Possible Dimerization**	**Protein-protein interactions ("#" indicates proposed interactions)**	**Reference**
Sleepless/Quiver	Numerous *Drosophila* species	Signalling protein in regulating sleep via potassium channels	NI	Nicotinic acetylcholine receptors[#]; Shaker; Erect wing	(Giot *et al*., 2003; Koh *et al*., 2008; Wu *et al*., 2016, 2014, 2010); (Clark *et al*., 2007); (Richards, 2005); (Ranz *et al*., 2007)
Lachesin	*S. americana; D. melanogaster*	Homophilic cell adhesin, Signalling role in neuronal development[!]; Tracheal development, cell adhesin, neuronal development[!]	Homodimerization	Ribosomal protein L4; Patched; Septate junction proteins[#]; cell adhesin molecules[#]	(Formstecher *et al*., 2005; Giot *et al*., 2003; Karlstrom *et al*., 1993; Llimargas, 2004; Özkan *et al*., 2013; Strigini *et al*., 2006)
Fasciclin-1	*S. americana; D. melanogaster*	Neuronal cell adhesion	NI	Sarcoplasmic calcium-binding protein 2; Cuticular protein 64Aa; Lipophorin; Cell adhesin molecules[#] and atlastin	(Giot *et al*., 2003; Mazor *et al*., 2016; O'Sullivan *et al*., 2013; Zinn *et al*., 1988)
Fasciclin-2	*D. melanogaster*	Neuronal recognition	NI	Gαs; Lipophorin; Ribosomal proteins[#]; calcium^{2+} ATPase; Various protein kinases[#].	(Grenningloh *et al*., 1991; Rees *et al*., 2011; Wolfgang *et al*., 2004)
Contactin	*D. melanogaster*	Cell-adhesion receptor in axo-glial Septate junctions.	NI	Neurexin & Neuroglian (glycoproteins)	(Faivre-Sarrailh, 2004)
Connectin	*D. melanogaster*	Homophilic cell adhesin	NI	DNA fragmentation factor-related protein 2; Lipophorin	(Panáková *et al*., 2005)
Amalgam	*D. melanogaster*	Cell adhesin ligand	Homodimerization	Neurotactin	(Frémion *et al*., 2000; Özkan *et al*., 2013; Seeger *et al*., 1988)
Membrane alanyl aminopeptidase	*Heliothis virescens; Manduca sexta*	*Bacillus thuringiensis* toxin, CryIA(C) peptidase receptor	NI	NI	(Gill *et al*., 1995); (Knight *et al*., 1995)
Aminopeptidase N	*M. sexta; Plutella xylostella*	The release of an N-terminal amino acid from a peptide; *B. thuringiensis* CRY1AB5 peptidase receptor	NI	NI	(Agrawal *et al*., 2002; Denolf *et al*., 1997)

Richardson *et al.* (1993) also showed that native Bm86 from *R. microplus*, as well as full-length recombinant Bm86 produced in baculoviral-infected insect Sf9 cells, is anchored to the extracellular surface of cell membranes via a GPI-anchor. Initial yeast two-hybrid studies in our laboratory further provide evidence for Bm86 forming homodimers (Ferreira, 2017 unpublished), which is a common characteristic of GPI-anchored proteins that function in signal transduction in arthropods (Table 1.2).

As the GPI-anchored protein only partly spans the outer layer of the plasma membrane bilayer, any intracellular signalling functions must be mediated via other proteins, such as membrane-spanning GTP-binding proteins (better known as G proteins). In a study by Solomon *et al.* (1996) that employed GTP binding assays and immunoprecipitation, G proteins (specifically $G_{\alpha i}$) were found to be physically associated with GPI-anchored proteins in lymphocytes. Suzuki *et al.* (2007) further showed that stimulation of a GPI-anchored protein, CD59, recruited $G_{\alpha i}$ and resulted in the stimulation-induced temporary arrest of lateral diffusion (STALL) of CD59 clusters and activation of Phospholipase C_{γ} via the production of IP_3 and Ca^{2+} signals (Suzuki *et al*., 2007a, 2007b).

Lastly, as stated previously, GPI-anchors contain ACE, PLC and PLD cleavage sites. Cleavage of a GPI-protein from the extracellular membrane is commonly associated with additional extracellular function(s) of such a GPI-protein. The presence of such sites on the Bm86 antigen may, therefore, point towards multiple roles of the Bm86 protein, but this remains to be studied.

G proteins as signalling mediators for GPI-anchored proteins

G proteins are involved in many signal transduction pathways, including stimulation of adenylate cyclase, Phospholipase A2 (PLA2), phosphoinositide-3-kinase (PI3K), PLC, and the regulation of Ca^{2+} channels, among others (Solomon *et al*., 1996; Suzuki *et al*., 2007b, 2007a). G proteins can be grouped into two classes, namely; the large heterotrimeric G proteins and the small G proteins (Neves *et al*., 2002). GPI-anchored proteins are physically associated with large heterotrimeric G proteins in lymphocytes; which emphasises the role of GPI-anchored proteins in signal transduction processes in the cell (Solomon *et al*., 1996; Suzuki *et al*., 2007b).

To date, heterotrimeric G proteins remain to be fully described in ticks but have been described for other arthropods (Table 1.3). As we propose that the GPI-linked Bm86 homodimerizes and then mediates signalling via a heterotrimeric G protein, this section will describe the structure-function relationship of heterotrimeric G proteins and their associated signalling pathways.

Table 1.3: Heterotrimeric G proteins in arthropods for which reviewed sequences are available. NI indicates when there is little/no information found. (Continues on next page).

Heterotrimeric G Proteins in Arthropods					
G protein Subunit	**Species**	**Subfamily**	**Possible Process**	**Protein-protein interactions**	**Reference**
α subunits			**Binds and hydrolyses GTP; provides the specificity for receptor and effector combination**		
αi	*D. melanogaster*	Gαi/o/t/z	Adenylate cyclase modulation in neuronal cell division & differentiation	Gβ1; Loco, Rapsynoid; GPCR moody	(Granderath *et al*., 1999; Yu *et al*., 2005)
αo	*D. melanogaster*		Wnt/frizzled & planar/frizzled pathway transduction, adenylate cyclase modulation in neuronal cell division & differentiation	Gβ1; Gγ1; Frizzled; Axin; Loco; Rapsynoid; GPCR moody; PLC 21C	(Dahdal *et al*., 2010; Egger-Adam and Katanaev, 2010; Katanaev *et al*., 2005; Katanayeva *et al*., 2010)
	M. sexta		Adenylate cyclase modulation in neuronal cell differentiation in antennae	B-amyloid protein-like protein	(Horgan *et al*., 1994)
	A. gambiae		Olfactory transduction in female antennae		(Rützler *et al*., 2006)
αs	*D. melanogaster*	Gs	Activates adenylate cyclase in the nervous system & eyes. Mediation of CHIP/LBD complexes in the wing; sensory brittle development.	Gβ1; Gγ1; Fasciclin-2; Dunc; CHIP; Ribosomal proteins S6,13, L26	(Bronstein *et al*., 2010; Dahdal *et al*., 2010; Giot *et al*., 2003; Katanayeva *et al*., 2010; Wolfgang *et al*., 2004)
	D. pseudoobscura pseudoobscura		Activates adenylate cyclase	NI	(Richards, 2005)
	A. gambiae		Adenylate cyclase-activating dopamine receptor signalling pathway	NI	(Rützler *et al*., 2006)
αq	*D. melanogaster*	Gq	Activation of PLC, Visual transduction in the eye. Metabolism of DAG to phosphatidic acid	Gαq; Gβ2; PLC 21C; Diacylglycerol Kinase; Frazzled	(Elia *et al*., 2005; Hardie *et al*., 2002; Hiramoto and Hiromi, 2006; Kain *et al*., 2008)
	A. gambiae		Phototransduction & olfactory transduction	NI	(Rützler *et al*., 2006)
αf	*D. melanogaster*	Gf	Activation of Rho1 for JAK/STAT pathway activation	Hopscotch	(Bausek and Zeidler, 2014; Quan *et al*., 1993)
Concertina	*D. melanogaster*	G12/13	Fog signalling	Cyclin K, GPCR kinase 2; DRhoGEF2	(Fuse *et al*., 2013; Giot *et al*., 2003; Nikolaidou and Barrett, 2004)

Heterotrimeric G-Proteins in Arthropods					
G protein Subunit	**Species**	**Subfamily**	**Possible Process**	**Protein-protein interactions**	**Reference**
γ Subunits			**βγ heterodimer GTPase activation of effectors**		
γ1	*D. melanogaster*	Gγ	Wnt/frizzled & planar/frizzled pathway transduction, adenylate cyclase modulation in neuronal cell division & differentiation	Gαo; Gβ1	(Izumi *et al*., 2004; Katanaev *et al*., 2005)
γ-e	*D. melanogaster; C. vicina*		Visual transduction in the eye	Gαq; Gβ2	(Schulz *et al*., 1999)
β Subunits				NI	
β1	*D. melanogaster*	Gβ	Adenylate cyclase modulation in neuronal cell division & differentiation	Gαi/o; Gγ1	(Izumi *et al*., 2004; Katanayeva *et al*., 2010; Schaefer *et al*., 2001)
β2	*D. melanogaster*		Activation of PLC, Visual transduction in the eye	Gγ-e; Gαq	(Elia *et al*., 2005; Schulz *et al*., 1999)

Heterotrimeric G proteins typically contain three subunits, namely the α, β and γ subunits. The α subunit binds and hydrolyses GTP while the β and γ subunits form a dimer known as βγ (Reviewed by Neer 1995). Heterotrimeric G proteins often function as signal transduction molecules, communicating signals from membrane receptors to intracellular effectors (Neves *et al*., 2002). Subtypes of heterotrimeric G proteins are classified based on the homology of their α subunit, which provides the specificity for receptor and effector combination and is usually implicated in the activation of second messengers involved in the signalling cascade.

When the G protein is inactive, and the α subunit is bound to GDP, the α and βγ subunits associate. However, when a ligand binds to the receptor, the α subunit of the G protein releases GDP, takes up GTP for hydrolysation (Figure 1.6) and the subunit dissociates. GTP cleavage results in the release of the α and βγ subunits to separately activate downstream ligand-specific effectors, such as PLC, PLA2, PI3K to name a few (Table 1.3) (Carty *et al.* 1990; Linder *et al.* 1990; Reviewed by Kaziro *et al.* 1991 and Neer 1995). Each family and subunit will be discussed in turn.

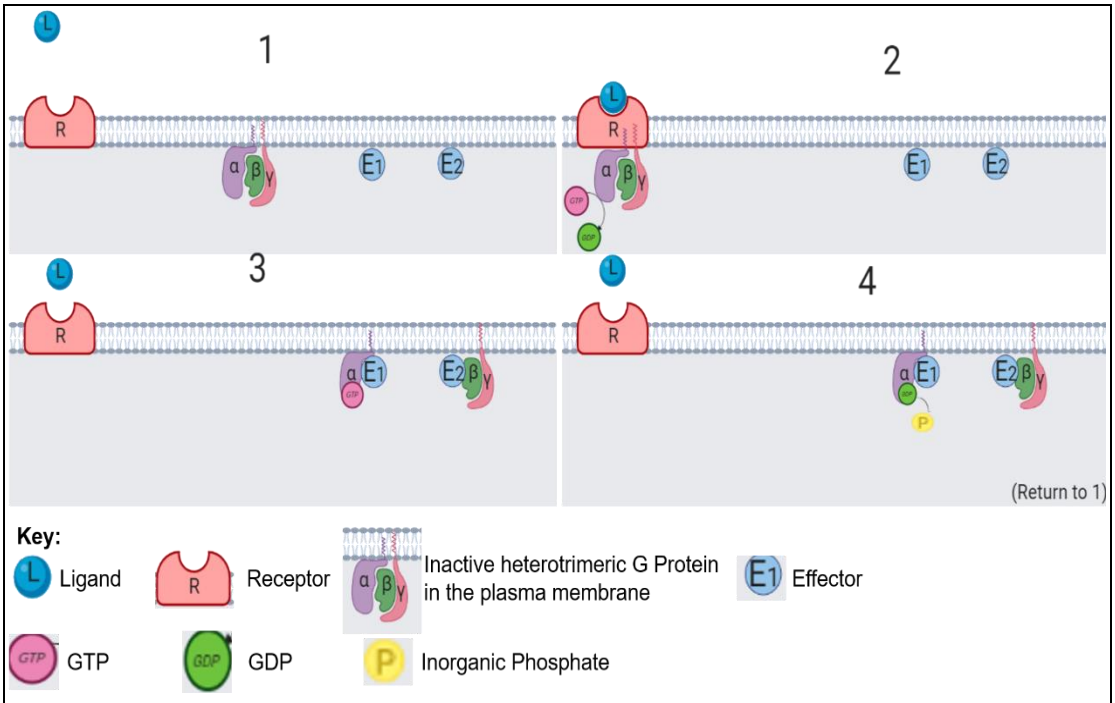


Figure 1.6: Mechanism of G protein signalling. 1: The G protein is inactive, the α and βγ subunits associate. 2-3: When a ligand binds to the receptor, the α subunit of the G-protein releases GDP, takes up GTP for hydrolysation and the subunit dissociates 3-4: The α and βγ subunits separately and ligand-specifically activate effectors, such as PLC, PLA2, PI3K and more. Adapted from Freissmuth *et al.* (1999) using BioRender (https://biorender.com).

The $G_s/_{olf}$ proteins form the first subtype, and their α-subunits are known to activate adenylyl cyclase and SRC tyrosine kinases (Reviewed by Landry *et al*., 2006). The G_s α-subunit is also involved in the activation of calcium channels and are ubiquitously expressed (Reviewed by Landry *et al*., 2006). In contrast, the G_{olf} proteins are only expressed in specific central nervous system ganglia and olfactory cells (Reviewed by Landry *et al*., 2006). In *D. melanogaster*, it has been shown that G_s subunits function in a variety of pathways including regulation of cyclic adenosine monophosphate (cAMP) production (Kimura, 2004), circadian rhythms (Dahdal *et al*., 2010), synaptic development (Wolfgang *et al*., 2004), as well as a putative role in wing development (Bronstein *et al*., 2010; Katanayeva *et al*., 2010).

The second subtype is composed of $G_{i/0/t}$ with the most common effectors being adenylyl cyclase and some calcium and potassium channels (Linder *et al*., 1990). Specifically, $G_{\alpha i}$ protein subunits have been shown to interact with GPI-linked proteins and mediate the subsequent activation of PLC_{γ} in mammalian and murine lymphocytes (Suzuki *et al*., 2007a, 2007b). In *D. melanogaster* neuronal cells it has been shown that $G_{\alpha o}$ leads to the activation of PLC 21C and contributes to the regulation of circadian rhythms (Dahdal *et al*., 2010; Kain *et al*., 2008).

The third subtype consists of the ubiquitously expressed G_q proteins that mainly function to activate Phospholipase C_{β} (Kühn *et al*., 1996). The G protein βγ dimers have a variety of effectors, but of interest is the βγ dimer that occurs in the G_q subtype that also acts via PLC_{β} to affect cellular signalling (Watling, 2001). In both *D. melanogaster* and *Anopheles Gambiae,* G_q subunits have been implicated in photo and olfactory transduction (Elia *et al*., 2005; Hiramoto and Hiromi, 2006; Kain *et al*., 2008; Rützler *et al*., 2006). In *D. melanogaster*, it has been shown that G_q subunit interacts explicitly with the PLC_{β} 21C to modulate olfactory signals in response to odorants (Kain *et al*., 2008).

The fourth subtype is the $G_{12/13}$ proteins which are ubiquitous and activate Rho GTPases (Reviewed by Landry *et al*., 2006). To date, the only documented $G_{12/13}$ subunit in arthropods is the protein Concertina in *D. melanogaster.* Concertina is implicated in folded gastrulation (Fog) signalling and activates Sharpei/DRhoGEF2 (a guanine nucleotide exchange factor) to activate myosin for epithelial invagination and folding (Nikolaidou and Barrett, 2004).

In arthropods, two G_{γ} and two G_{β} subunits have been described to date. The $G_{\gamma 1}$ subunit has been documented in *D. melanogaster* and interacts with the $G_{\alpha o}$ and $G_{\beta 1}$ subunits in signal transduction (Izumi *et al*., 2004; Katanaev *et al*., 2005). The $G_{\gamma\text{-}e}$

subunit interacts with $G_{\alpha q}$ and $G_{\beta 2}$ for visual signal transduction in *D. melanogaster* and has also been found in *Calliphora vicina* (Schulz *et al*., 1999).

A novel G subtype, namely G_f, has been described in *D. melanogaster*. The $G_{\alpha f}$ subunit is expressed in the developing midgut and aminoserosa of these arthropods. It is hypothesised to function downstream of the Janus Kinase/Signal Transducer and Activator of Transcription (JAK/STAT) pathway via interactions with the Ras-like GTP-binding protein, Rho1 (Bausek and Zeidler, 2014; Quan *et al*., 1993).

Currently, there are few papers on heterotrimeric G proteins and their downstream effects in Acari, despite the numerous isoforms that have been identified from genomic and transcriptomic data (i.e. only unreviewed sequences are available) from *Ixodes scapularis*, *Amblyomma variegatum*, *Amblyomma americanum*, *Amblyomma maculatum*, *Ornithodoros turicata*, *Rhipicephalus pulchellus*, and *Hyalomma excavatum* (www.uniprot.org). With regards to G proteins and downstream processes mediated by PLC (via IP_3) and calcium, several studies have been published. Firstly, in both ixodid (Reviewed by Sauer *et al*., 2000) and argasid ticks (Maritz-Olivier *et al*., 2005) it has been shown that dopamine and prostaglandin-dependent signalling pathways mediate the exocytosis of salivary gland proteins in *A. americanum* (Sauer *et al*., 2000) and *O. savignyi* (Maritz-Olivier *et al*., 2005). Secondly, a G protein-coupled leukokinin-like receptor responding to various myokinins, including Lymnokinin and Muscakinin, in *R. microplus* was found to elicit intracellular calcium signals which are postulated to stimulate secretion in Malpighian tubules, but their precise functioning remains to be confirmed (Holmes *et al*., 2003, 2000).

Apart from signalling, G protein-dependent pathways are targets of acaricides like amitraz. In this regard, the G protein-coupled receptor for octopamine is agonised by the amitraz compound, but the exact mechanism of this interaction remains unclear (Baxter and Barker, 1999; Kita *et al*., 2017).

In summary, the link between G proteins, GPI-linked proteins and PLC has been described for numerous organisms, from mammals to arthropods. As several downstream components of the PLC pathway were identified previously in our group using DNA microarrays in ticks fed on cattle vaccinated with Bm86 (Figure 1.4), the next section will describe PLCs, which are known to be activated by $G_{\alpha q}$ and $_{\beta\gamma}$-protein subunits, as well as the $G_{\alpha i}$-subunit.

Phospholipase C (PLC)

Phospholipase C refers to a class of multidomain, soluble protein isozymes under cell surface receptor control (Reviewed by Rhee and Bae, 1997). This group of proteins characteristically consists of X- and Y-box regions that form a catalytic α/β-barrel, around which the other domains are organised (Williams and Katan, 1996). Phospholipase C proteins are divided into six types namely β, γ, δ, ε, ζ, and η; each of which also has more than one alternative splicing variant (Bamji-Mirza and Yoa, 2011). The domain architecture of the murine PLC's is shown in Figure 1.7. Of interest is that the simplest PLC, found in prokaryotes, consists only of the catalytic α/β-barrel (Heinz *et al*., 1996) while the largest, PLC_{ε}, contains two novel additional protein domains (Shibatohge *et al*., 1998).

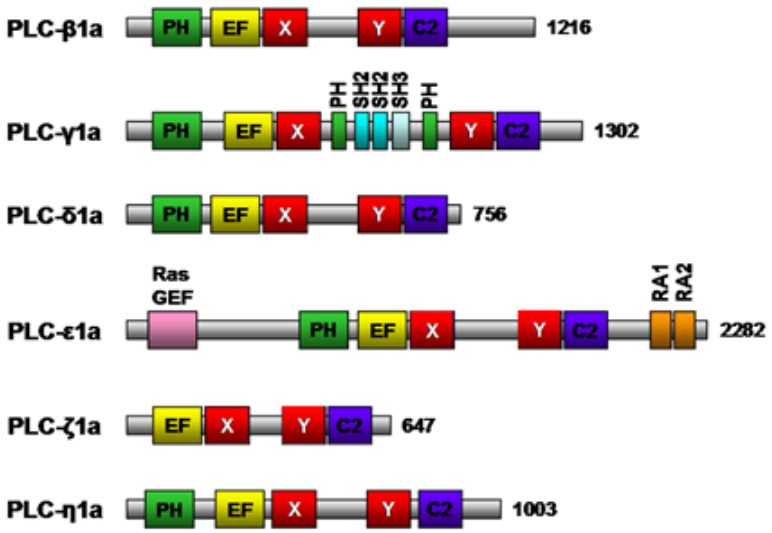


Figure 1.7: The typical domain architecture as derived from the well characterised murine PLC families. Each class of mouse phospholipase C with specific domain order are shown. The number of amino acid residues for each is indicated to the right. Taken from Bamji-Mirza and Yoa (2011)

Phospholipase C proteins cleave the polar head group from inositol phospholipids to hydrolyse phosphatidylinositol 4,5-bisphosphate (PIP_2); generating inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG) (Bamji-Mirza and Yoa, 2011). IP_3 is a second messenger for the release of intracellular calcium and DAG is a known activator of protein kinase C (PKC) (Reviewed by Rebecchi and Pentyala 2000). Both IP_3 and DAG lead to the activation of kinases that activate transcription factors to initiate transcription (Figure 1.8).

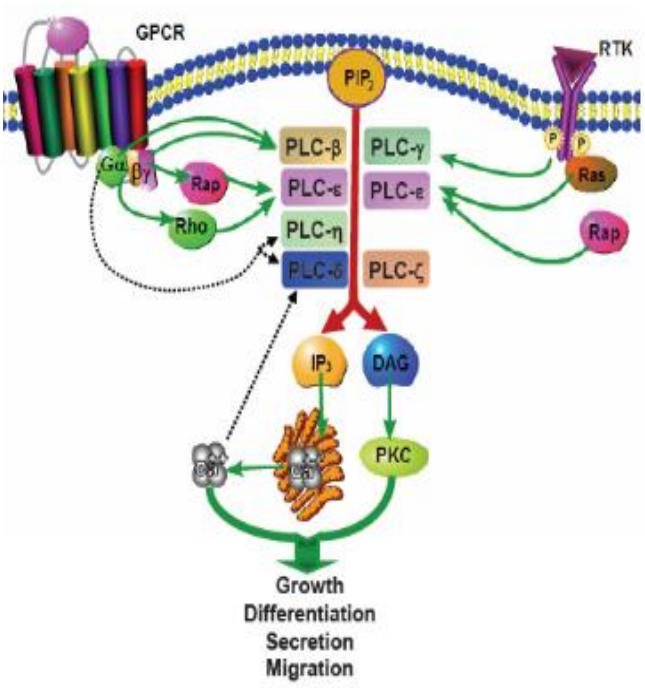


Figure 1.8: The general effect of activation of different classes of PLC. Upper green arrows are showing the activation of PLC either by a G protein or a specific receptor. The red arrow is indicating that PLC hydrolyses PIP_2 to PIP_3 and DAG where the lower green arrows show the release of Ca^{2+} which acts as a second messenger with DAG activating PKC. Taken from Suh *et al*. (2008). Abbreviations correspond to GPCR: G protein-coupled receptor, RTK: Receptor, Rap: Ras-related proteins, PKC: Protein Kinase C, DAG: Diacylglycerol, IP_3 : Inositol 1 ,4, 5- triphosphate

Three classes of PLC families have been identified in arthropods to date (Table 1.4), including PLC_{β} (two subclasses, $PLC_{\beta\text{-}1}$ and PLC 21C($_{\beta\text{-}2}$)), PLC_{γ} and PLC_{ε}. While there is evidence for PLC_{γ} and PLC_{ε} in several arthropod species, the pathways in which they function remain to be elucidated. In *D. melanogaster*, $PLC_{\beta\text{-}1}$ has been implicated in phototransduction and is encoded by the no receptor potential A (norpA) gene (Bloomquist *et al*., 1988). The $PLC_{\beta\text{-}1}$ forms a complex with other proteins and directly interacts with at least one adapter protein, the inactivation no afterpotential D (INAD) protein, and the $G_{\alpha q}$ subunit of the G protein (Bähner *et al*., 2000; Shieh *et al*., 1997). PLC 21C, also a PLC_{β} protein, has been implicated in olfactory transduction in *D. melanogaster* and interacts with $G_{\alpha q}$ and $_{\alpha i}$ (Dahdal *et al*., 2010; Kain *et al*., 2008).

Table 1.4: Phospholipase proteins in arthropods. *Indicates proteins that have sequences which have been reviewed, and which correlate with given functions and interactions. NI indicates when there is little/no information found.

Phospholipase in Arthropods				
Protein	**Species**	**Process**	**Interactions**	**References**
Phospholipase C		Production of IP3 and DAG upon stimulation		
PLCβ				
PLC-β (1)	*D. melanogaster*; R. appendiculatus; R. zambeziensis; Acromyrmex echinatior; Acyrthosiphon pisum; Tropilaelaps mercedesae; Sarcoptes scabiei*	Phototransduction*	INAD*; Gαq*	(Bähner *et al*., 2000; Bloomquist *et al*., 1988; Shieh *et al*., 1997); (de Castro *et al*., 2016); (Nygaard *et al*., 2011); (Richards *et al*., 2010); (Dong *et al*., 2017); (Rider *et al.* 2015)
PLC 21C (β2)	*D. melanogaster*; Papilio machaon*	Olfactory transduction*	Gαq*; Gαo*	(Dahdal *et al*., 2010; Kain *et al*., 2008); (Li *et al*., 2015)
PLCγ				
PLCγ	*A. echinatior I. scapularis A. pisum*	NI	NI	(Nygaard *et al*., 2011); (Ayllón *et al*., 2015); (Richards *et al*., 2010)
PLCε				
PLCε	*A. echinatior A. pisum I. scapularis*	NI	NI	(Nygaard *et al*., 2011); (Richards *et al*., 2010); (Ayllón *et al*., 2015)

Research Question

As several downstream components of the PLC pathway were identified previously in our group using DNA microarrays in *R. microplus* fed on cattle vaccinated with Bm86, this study will now focus on *in silico* mining of transcriptomic data of *R. microplus* to identify whether the components of the proposed signalling pathway are present. Since it is known that Bm86 is a membrane-bound GPI-linked protein containing multiple EGF domains (Nijhof *et al*. 2010) and since GPI-linked proteins do not span the entire membrane, they possibly require a G protein to elicit the cellular responses to ligand binding, it is proposed that Bm86 functions in conjunction with a G protein. It is further suggested that this G protein interaction stimulates a PLC exocytosis pathway in tick midgut digest cells. The focus will be placed on the identification and classification of PLCs, which are known to be activated by G_q α / βγ-protein or $G_{αi}$-subunit as discussed. Whether this pathway is active in the tick midgut remains to be experimentally proven in subsequent studies.

It is not only necessary to corroborate the findings of the microarray gene expression study done by this research group but also to elucidate the precise interaction of Bm86 with the Antigen 1, to understand further the normal biological function of Bm86 in the tick midgut cell wall and consequently the mechanism behind the Bm86 vaccine. The results would allow for improvements to be made to the vaccine and lead to the identification of possible survival mechanisms employed by those ticks which survive feeding on Bm86 vaccinated cattle.

AIMS, HYPOTHESIS, AND OBJECTIVES

Research Questions:

1) Do the Bm86 and Antigen 1 proteins show sequence variation in the South African *R. microplus* laboratory strain from Clinvet?
2) Are G proteins and Phospholipase C proteins present in the *R. microplus* gut?
3) What region of Antigen 1 is implicated in interaction with Bm86?

Hypothesis:

1) The sequence for the *R. microplus* South African laboratory strain Bm86 varies in different life stages and tissues, as does that of the interacting region of Antigen 1 with Bm86.
2) G proteins capable of interacting with Phospholipase C are present in *R. microplus* gut cells.
3) A single region of Antigen 1 is responsible for the interaction of Antigen 1 with Bm86.

Aims and objectives:

1) Identification of sequences for Bm86, Antigen 1 and members of the PLC pathway from various life stages and tissues of a South African strain of *R. microplus*.

    Objectives for aim 1:

    i. *De novo* assembly of transcriptomic data for the Clinvet strain of *R. microplus*.
    ii. Similarity searches for identification of sequences from the data set.
    iii. Construct nucleotide and amino acid alignments for the identification of sequence variation
    iv. Validation of significant sequence variation via Sanger Sequencing
    v. Validation of the presence of identified G proteins and PLC sequences as proteins from tick gut extracts via ELISA assay

2) Map the protein interaction domain of Antigen 1 between Bm86 and the Antigen 1.

    Objectives for aim 2:

    i. Construct plasmids for a yeast two-hybrid study for Bm86 and various lengths of Antigen 1. Binding domain (BD) attached to Bm86 and the activation domains (AD) being the library of different sections of Antigen 1.
    ii. Titre and amplify correct plasmids in *E. coli*, and extract plasmid DNA.
    iii. Verify Plasmid inserts
    iv. Co-transform yeast with extracted plasmids
    v. Plate transformations on appropriate selective medium and observe for interaction indicators

REFERENCES

Adakal, H., Biguezoton, A., Zoungrana, S., Courtin, F., De Clercq, E.M., Madder, M., 2013. Alarming spread of the Asian cattle tick *Rhipicephalus microplus* in West Africa-another three countries are affected: Burkina Faso, Mali and Togo. Exp. Appl. Acarol. 61, 383–386. https://doi.org/10.1007/s10493-013-9706-6

Agrawal, N., Malhotra, P., Bhatnagar, R.K., 2002. Interaction of gene-cloned and insect cell-expressed aminopeptidase N of *Spodoptera litura* with insecticidal crystal protein Cry1C. Appl. Environ. Microbiol. 68, 4583–4592. https://doi.org/10.1128/AEM.68.9.4583-4592.2002

Ali, A., Parizi, L.F., Ferreira, B.R., da Silva Vaz Junior, I., 2016. A revision of two distinct species of *Rhipicephalus*: *R. microplus* and *R. australis*. Ciência Rural 46, 1240–1248. https://doi.org/10.1590/0103-8478cr20151416

Allen, J.R., Humphreys, S.J., 1979. Immunisation of guinea pigs and cattle against ticks. Nature 280, 491–493.

Almazán, C., Lagunes, R., Villar, M., Canales, M., Rosario-Cruz, R., Jongejan, F., de la Fuente, J., 2010. Identification and characterisation of *Rhipicephalus* (*Boophilus*) *microplus* candidate protective antigens for the control of cattle tick infestations. Parasitol. Res. 106, 471–479. https://doi.org/10.1007/s00436-009-1689-1

Amaral, A.C.F., de S. Ramos, A., Pena, M.R., Ferreira, J.L.P., Menezes, J.M.S., Vasconcelos, G.J.N., da Silva, N.M., de Silva, J., 2017. Acaricidal activity of *Derris floribunda* essential oil and its main constituent. Asian Pac. J. Trop. Biomed. 7, 791–796. https://doi.org/10.1016/j.apjtb.2017.08.006

Anderson, J.F., Magnarelli, L.A., 2008. Biology of Ticks. Infect. Dis. Clin. North Am. 22, 195–215. https://doi.org/10.1016/j.idc.2007.12.006

Andreotti, R., Giachetto, P., Cunha, R., 2018. Advances in tick vaccinology in Brazil from gene expression to immunoprotection. Front. Biosci. 10, 127–142. https://doi.org/10.2741/s504

Anholeto, L.A., de Oliveira, P.R., Rodrigues, R.A.F., dos S. Spindola, C., Labruna, M.B., Pizano, M.A., de Castro, K.N., Camargo-Mathias, M.I., 2017. Potential action of extract of *Acmella oleracea* (L.) R.K. Jansen to control *Amblyomma cajennense* (Fabricius, 1787) (*Acari: Ixodidae*) ticks. Ticks Tick. Borne. Dis. 8, 65–72. https://doi.org/10.1016/j.ttbdis.2016.09.018

Ayllón, N., Villar, M., Galindo, R.C., Kocan, K.M., Šíma, R., López, J.A., Vázquez, J., Alberdi, P., Cabezas-Cruz, A., Kopáček, P., de la Fuente, J., 2015. Systems biology of tissue-specific response to *Anaplasma phagocytophilum* reveals differentiated apoptosis in the tick vector *Ixodes scapularis*. PLoS Genet. 11, 1–29. https://doi.org/10.1371/journal.pgen.1005120

Bähner, M., Sander, P., Paulsen, R., Huber, A., 2000. The visual G protein of fly photoreceptors interacts with the PDZ domain assembled INAD signalling complex via direct binding of activated Gα(q) to phospholipase Cβ J. Biol. Chem. 275, 2901–2904. https://doi.org/10.1074/jbc.275.4.2901

Baines, R.A., 2003. Postsynaptic Protein Kinase A Reduces Neuronal Excitability in Response to Increased Synaptic Excitation in the *Drosophila* CNS. J. Neurosci. 23, 8664–8672.

Bamji-Mirza, M., Yoa, Z., 2011. Phospholipases, AOCS Lipid Library. https://doi.org/10.21748/lipidlibrary.39190

Baron, S., Barrero, R.A., Black, M., Bellgard, M.I., van Dalen, E., Maritz-Olivier, C., 2018. Differentially expressed genes in response to amitraz treatment suggests a proposed model of resistance to amitraz in *R. decoloratus* ticks. Int. J. Parasitol. Drugs Drug Resist. 8, 361–371. https://doi.org/10.1016/j.ijpddr.2018.06.005

Baron, S., van der Merwe, N.A., Madder, M., Maritz-Olivier, C., 2015. SNP analysis infers that recombination is involved in the evolution of amitraz resistance in *Rhipicephalus microplus*. PLoS One 10, e0131341. https://doi.org/10.1371/journal.pone.0131341

Bastos, R.G., Ueti, M.W., Knowles, D.P., Scoles, G.A., 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. https://doi.org/10.1186/1756-3305-3-111

Bausek, N., Zeidler, M.P., 2014. G73B is a downstream effector of JAK/STAT signalling and a regulator of Rho1 in *Drosophila* haematopoiesis. J. Cell Sci. 127, 101–110. https://doi.org/10.1242/jcs.132852

Baxter, G.D., Barker, S.C., 1999. Isolation of a cDNA for an octopamine-like, G-protein coupled receptor

from the cattle tick, *Boophilus microplus*. Insect Biochem. Mol. Biol. 29, 461–467. https://doi.org/10.1016/S0965-1748(99)00023-5

Bigalke, R.D., 1980. The control of ticks and tick-borne diseases of cattle in South Africa. Control ticks tick-borne Dis. cattle South Africa. 11, 20–21.

Blecha, I.M.Z., Csordas, B.G., de Aguirre, A., Cunha, R.C., Garcia, M.V., Andreotti, R., 2018. Analysis of Bm86 conserved epitopes: Is a global vaccine against cattle tick *Rhipicephalus microplus* possible? Rev. Bras. Parasitol. Vet. 27, 267–279. https://doi.org/10.1590/s1984-296120180056

Bloomquist, B.T., Shortridge, R.D., Schneuwly, S., Perdew, M., Montell, C., Steller, H., Rubin, G., Pak, W.L., 1988. Isolation of a putative Phospholipase C gene of *Drosophila*, norpA, and its role in phototransduction. Cell 54, 723–733. https://doi.org/10.1016/S0092-8674(88)80017-5

Bronstein, R., Levkovitz, L., Yosef, N., Yanku, M., Ruppin, E., Sharan, R., Westphal, H., Oliver, B., Segal, D., 2010. Transcriptional regulation by CHIP/LDB complexes. PLoS Genet. 6, 12–15. https://doi.org/10.1371/journal.pgen.1001063

Bull, M.S., Swindale, S., Overend, D., Hess, E.A., 1996. Suppression of *Boophilus microplus* populations with fluazuron - An acarine growth regulator. Aust. Vet. J. 74, 468–470. https://doi.org/10.1111/j.1751-0813.1996.tb07575.x

Carty, D.J., Padrell, E., Codina, J., Birnbaumer, L., Hildebrandt, J.D., Iyengar, R., 1990. Distinct guanine nucleotide binding and release properties of the three Gi proteins. J.Biol.Chem. 265, 6268–6273.

Chen, G., Li, W., Zhang, Q.S., Regulski, M., Sinha, N., Barditch, J., Tully, T., Krainer, A.R., Zhang, M.Q., Dubnau, J., 2008. Identification of synaptic targets of *Drosophila* Pumilio. PLoS Comput. Biol. 4. https://doi.org/10.1371/journal.pcbi.1000026

Cheng, T.C., 1986. General Parasitology. Elsevier Science.

Cheng, Y.L.L., Andrew, D.J.J., 2015. Extracellular Mipp1 activity confers migratory advantage to epithelial cells during collective migration. Cell Rep. 13, 2174–2188. https://doi.org/10.1016/j.celrep.2015.10.071

Chi, H., Tiller, G.E., Dasouki, M.J., Romano, P.R., Wang, J., O'Keefe, R.J., Puzas, J.E., Rosier, R.N., Reynolds, P.R., 1999. Multiple Inositol Polyphosphate Phosphatase: Evolution as a distinct group within the Histidine Phosphatase Family and chromosomal localization of the human and mouse genes to chromosomes 10q23 and 19. Genomics 56, 324–336. https://doi.org/10.1006/geno.1998.5736

Chmelař, J., Kotál, J., Langhansová, H., Kotsyfakis, M., 2017. Protease Inhibitors in tick saliva: The role of Serpins and Cystatins in tick-host-pathogen interaction. Front. Cell. Infect. Microbiol. 7, e00216. https://doi.org/10.3389/fcimb.2017.00216

Clark, A.G., Eisen, M.B., Smith, D.R., *et al*., 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450, 203–218. https://doi.org/10.1038/nature06341

DAFF, 2016. Abstract of agricultural statistics. Abstr. Agric. Stat. https://doi.org/10.1017/CBO9781107415324.004

DAFF, 2017. A profile of the South African beef market value Directorate Marketing 1–53.

Dahdal, D., Reeves, D.C., Ruben, M., Akabas, M.H., Blau, J., 2010. *Drosophila* Pacemaker Neurons require G Protein signaling and GABAergic inputs to generate twenty-four-hour behavioral rhythms. Neuron 68, 964–977. https://doi.org/10.1016/j.neuron.2010.11.017

Davey, R.B., Miller, J.A., George, J.E., 2001. Efficacy of daily oral treatments of ivermectin administered to cattle infested with *Boophilus microplus* (*Acari: Ixodidae*). J. Agric. Urban Entomol. 18, 127–137.

de Castro, J.J., 1997. Sustainable tick and tick-borne disease control in livestock improvement in developing countries. Vet. Parasitol. 71, 77–97. https://doi.org/10.1016/S0304-4017(97)00033-2

de Castro, M.H., de Klerk, D., Pienaar, R., Latif, A.A., Rees, D.J.G., Mans, B.J., 2016. *De novo* assembly and annotation of the salivary gland transcriptome of *Rhipicephalus appendiculatus* male and female ticks during blood feeding. Ticks Tick. Borne. Dis. 7, 536–548. https://doi.org/10.1016/j.ttbdis.2016.01.014

de Clercq, E.M., Vanwambeke, S.O., Sungirai, M., Adehan, S., Lokossou, R., Madder, M., 2012. Geographic distribution of the invasive cattle tick *Rhipicephalus microplus*, a country-wide survey in Benin. Exp. Appl. Acarol. 58, 441–452. https://doi.org/10.1007/s10493-012-9587-0

de la Fuente, J., Almazán, C., Canales, M., Pérez de la Lastra, J.M., Kocan, K.M., Willadsen, P., 2007. A ten-year review of commercial vaccine performance for control of tick infestations on cattle. Anim. Health Res. Rev. 8, 23–28. https://doi.org/10.1017/S1466252307001193

de la Fuente, J., Contreras, M., 2015. Tick vaccines: current status and future directions. Expert Rev. Vaccines 14, 1367–1376. https://doi.org/10.1586/14760584.2015.1076339

de la Fuente, J., Estrada-Penn, A., Venzal, J.M., Kocan, K.M., Sonenshine, D.E., 2008. Overview: Ticks as vectors of pathogens that cause disease in humans and animals. Front. Biosci. 13, 6938–6946. https://doi.org/10.1016/S0002-8223(05)01189-2

de La Fuente, J., Kocan, K.M., Contreras, M., 2015. Prevention and control strategies for ticks and pathogen transmission. OIE Rev. Sci. Tech. 34, 249–264.

de la Fuente, J., Rodríguez, M., Montero, C., Redondo, M., García-García, J.C., Méndez, L., Serrano, E., Valdés, M., Enríquez, A., Canales, M., Ramos, E., Boué, O., Machado, H., Lleonart, R., 1999. Vaccination against ticks (*Boophilus* spp.): the experience with the Bm86-based vaccine Gavac™. Genet. Anal. Biomol. Eng. 15, 143–148. https://doi.org/10.1016/S1050-3862(99)00018-2

de Miranda Santos, I.K.F., Garcia, G.R., Oliveira, P.S., Veríssimo, C.J., Katiki, L.M., Rodrigues, L., Szabó, M.P.J., Maritz-Olivier, C., 2018. 4. Acaricides: current status and sustainable alternatives for controlling the cattle tick, *Rhipicephalus microplus*, based on its ecology 91–134. https://doi.org/10.3920/978-90-8686-863-6_4

de Vos, S., Zeinstra, L., Taoufik, A., Willadsen, P., Jongejan, F., 2001. Evidence for the utility of the Bm86 antigen from *Boophilus microplus* in vaccination against other tick species. Exp. Appl. Acarol. 25, 245–261. https://doi.org/10.1023/A:1010609007009

Denolf, P., Hendrickx, K., Van Damme, J., Jansens, S., Peferoen, M., Degheele, D., Van Rie, J., 1997. Cloning and characterisation of *Manduca sexta* and *Plutella xylostella* midgut aminopeptidase N enzymes related to *Bacillus thuringiensis* toxin-binding proteins. Eur. J. Biochem. 248, 748–761.

Derose, R., Mckenna, R. V, Cobon, G., Tennent, J., Zakrzewski, H., Gale, K., Wood, P.R., Scheerlinck, J.P.Y., Willadsen, P., 1999. Bm86 antigen induces a protective immune-response against *Boophilus microplus* following DNA and protein vaccination in sheep. Vet. Immunol. Immunopathol. 71, 151–160.

Derso, S., Demessie, Y., 2015. Tick-borne hemoparasitic diseases of ruminants: A review. Adv. Biol. Res. (Rennes). 9, 210–224. https://doi.org/10.5829/idosi.abr.2015.9.4.9516

Dong, X., Armstrong, S.D., Xia, D., Makepeace, B.L., Darby, A.C., Kadowaki, T., 2017. Draft genome of the honey bee ectoparasitic mite, *Tropilaelaps mercedesae*, is shaped by the parasitic life history. Gigascience 6, 1–17. https://doi.org/10.1093/gigascience/gix008

Dumler, J.S., 2005. *Anaplasma* and *Ehrlichia* infection. Ann. N. Y. Acad. Sci. 1063, 361–373. https://doi.org/10.1196/annals.1355.069

Egger-Adam, D., Katanaev, V.L., 2010. The trimeric G protein Go inflicts a double impact on axin in the Wnt/Frizzled signaling pathway. Dev. Dyn. 239, 168–183. https://doi.org/10.1002/dvdy.22060

Elia, N., Frechter, S., Gedi, Y., Minke, B., Selinger, Z., 2005. Excess of Gβe over Gqαe *in vivo* prevents dark, spontaneous activity of *Drosophila* photoreceptors. J. Cell Biol. 171, 517–526. https://doi.org/10.1083/jcb.200506082

Estrada-Peña, A., Venzal, J.M., Nava, S., Mangold, A., Guglielmone, A.A., Labruna, M.B., de la Fuente, J., 2012. Reinstatement of *Rhipicephalus* (*Boophilus*) *australis* (*Acari*: *Ixodidae*) with redescription of the adult and larval stages. J. Med. Entomol. 49, 794–802. https://doi.org/10.1603/ME11223

Eugster, C., Panáková, D., Mahmoud, A., Eaton, S., 2007. Lipoprotein-Heparan Sulfate interactions in the Hh pathway. Dev. Cell 13, 57–71. https://doi.org/10.1016/j.devcel.2007.04.019

Faivre-Sarrailh, C., 2004. *Drosophila* Contactin, a homolog of vertebrate contactin, is required for septate junction organization and paracellular barrier function. Development 131, 4931–4942. https://doi.org/10.1242/dev.01372

FAO, 2018. The future of food and agriculture - alternative pathways to 2050, Rome. https://doi.org/10.1074/jbc.M308518200

FAO, 2017. The future of food and agriculture, trends and challenges, Channels. https://doi.org/10.1155/2010/178034

Ferreira, G.F., Freitas, T.M., Gonçalves, C.L., Mendes, J.F., Vieira, J.N., Villareal, J.P., Nascente, P.S., 2016. Antiparasitic drugs: in vitro tests against nematophagous fungi. Brazilian J. Biol. 76, 990–993. https://doi.org/10.1590/1519-6984.05615

Field, L.M., James, A.A., Liyou, N., Hamilton, S., Elvin, C., Willadsen, P., 1999. Cloning and expression of ecto 5′-nucleotidase from the cattle tick *Boophilus microplus*. Insect Mol. Biol.

Formstecher, E., Aresta, S., Collura, V., *et al*., 2005. Protein interaction mapping: A *Drosophila* case study. Genome Res. 15, 376–384. https://doi.org/10.1101/gr.2659105

Fouche, G., Ramafuthula, M., Maselela, V., Mokoena, M., Senabe, J., Leboho, T., Sakong, B.M., Adenubi, O.T., Eloff, J.N., Wellington, K.W., 2016. Acaricidal activity of the organic extracts of thirteen South African plants against *Rhipicephalus* (*Boophilus*) *decoloratus* (*Acari: Ixodidae*). Vet. Parasitol. 224, 39–43. https://doi.org/10.1016/j.vetpar.2016.05.011

Fournier, D., Bride, J.M., Karch, F., Bergé, J.B., 1988. Acetylcholinesterase from *Drosophila melanogaster* identification of two subunits encoded by the same gene. FEBS Lett. 238, 333–337. https://doi.org/10.1016/0014-5793(88)80507-6

Francischetti, I.M.B., Sá-Nunes, A., Mans, B.J., Santos, I.M., Ribeiro, J.M.C., 2010. The role of saliva in tick feeding. Front. Biol. Sci. 14, 2051–2088.

Freeman, J.M., Davey, R.B., Kappmeyer, L.S., Kammlah, D.M., Olafson, P.U., 2010. Bm86 midgut protein sequence variation in South Texas cattle fever ticks. Parasit. Vectors 3, 101–109. https://doi.org/10.1186/1756-3305-3-101

Freissmuth, M., Waldhoer, M., Bofill-Cardona, E., Nanoff, C., 1999. G protein antagonists signalling mechanism and potential target sites for drug action. TiPS 20, 237–245.

Frémion, F., Darboux, I., Diano, M., Hipeau-Jacquotte, R., Seeger, M.A., Piovant, M., 2000. Amalgam is a ligand for the transmembrane receptor neurotactin and is required for neurotactin-mediated cell adhesion and axon fasciculation in *Drosophila*. EMBO J. 19, 4463–72. https://doi.org/10.1093/emboj/19.17.4463

Fuse, N., Yu, F., Hirose, S., 2013. Gprk2 adjusts Fog signalling to organize cell movements in *Drosophila* gastrulation. Development 140, 4246–4255. https://doi.org/10.1242/dev.093625

Fyumagwa, R.D., Simmler, P., Meli, M.L., Hoare, R., Hofmann-Lehmann, R., Lutz, H., 2009. Prevalence of *Anaplasma marginale* in different tick species from Ngorongoro Crater, Tanzania. Vet. Parasitol. 161, 154–157. https://doi.org/10.1016/j.vetpar.2008.12.018

Gandhi, P.R., Jayaseelan, C., Mary, R.R., Mathivanan, D., Suseem, S.R., 2017. Acaricidal, pediculicidal and larvicidal activity of synthesised ZnO nanoparticles using *Momordica charantia* leaf extract against blood-feeding parasites. Exp. Parasitol. 181, 47–56. https://doi.org/10.1016/j.exppara.2017.07.007

Ganfornina, M.D., Sánchez, D., Bastiani, M.J., *et al.*, 1995. Lazarillo, a new GPI-linked surface lipocalin, is restricted to a subset of neurons in the grasshopper embryo. Development 121, 123–34. https://doi.org/10.1006/jmbi.1990.9999

García-García, J.C., Gonzalez, I.L., González, D.M., Valdés, M., Méndez, L., Lamberti, J., D'Agostino, B., Citroni, D., Fragoso, H., Ortiz, M., Rodríguez, M., De La Fuente, J., 1999. Sequence variations in the *Boophilus microplus* Bm86 locus and implications for immunoprotection in cattle vaccinated with this antigen. Exp. Appl. Acarol. 23, 883–895. https://doi.org/10.1023/A:1006270615158

García-García, J.C., Montero, C., Redondo, M., Vargas, M., Canales, M., Boue, O., Rodríguez, M., Joglar, M., MacHado, H., González, I.L., Valdés, M., Méndez, L., de la Fuente, J., 2000. Control of ticks resistant to immunisation with Bm86 in cattle vaccinated with the recombinant antigen Bm95 isolated from the cattle tick, *Boophilus microplus*. Vaccine 18, 2275–2287. https://doi.org/10.1016/S0264-410X(99)00548-4

Gill, S.S., Cowles, E.A., Francis, V., 1995. Identification, isolation, and cloning of a *Bacillus thuringiensis* CryIAc toxin-binding protein from the midgut of the lepidopteran insect *Heliothis virescens*. J. Biol. Chem. 270, 27277–27282. https://doi.org/10.1074/jbc.270.45.27277

Giot, L., Bader, J.S., Brouwer, C., *et al*.., 2003. A protein interaction map of *Drosophila melanogaster*. Science (80-.). 302, 1727–1736. https://doi.org/10.1126/science.1090289

Goldschmidt-Clermont, P.J., Kim, J.W., Machesky, L.M., Rhee, S.G., Pollard, T.D., 1991. Regulation of Phospholipase C-gamma1 by profilin and tyrosine phosphorylation. Science (80-.). 251, 1231–1233.

González, L.J., Cremata, J.A., Guanche, Y., Ramos, Y., Triguero, A., Cabrera, G., Montesino, R., Huerta, V., Pons, T., Boué, O., Farnós, O., Rodríguez, M., 2004. The cattle tick antigen, Bm95, expressed in *Pichia pastoris*, contains short chains of N- and O-glycans. Arch. Biochem. Biophys. 432, 205–211. https://doi.org/10.1016/j.abb.2004.09.022

Granderath, S., Stollewerk, A, Greig, S., Goodman, C.S., O'Kane, C.J., Klämbt, C., 1999. Loco encodes an RGS protein required for *Drosophila* glial differentiation. Development 126, 1781–1791. https://doi.org/10.1002/(SICI)1096-9861(19981207)402:1<32::AID-CNE3>3.0.CO;2-V

Grenningloh, G., Jay Rehm, E., Goodman, C.S., 1991. Genetic analysis of growth cone guidance in *Drosophila*: Fasciclin II functions as a neuronal recognition molecule. Cell 67, 45–57. https://doi.org/10.1016/0092-8674(91)90571-F

Gu, X. Bin, Liu, G.H., Song, H.Q., Liu, T.Y., Yang, G.Y., Zhu, X.Q., 2014. The complete mitochondrial genome of the scab mite *Psoroptes cuniculi* (*Arthropoda: Arachnida*) provides insights into *Acari* phylogeny. Parasites and Vectors 7, e49461. https://doi.org/10.1186/1756-3305-7-340

Gubler, D.J., 1998. Resurgent vector-borne diseases as a global health problem. Emerg. Infect. Dis. 4, 442–450.

Guerrero, F.D., Miller, R.J., Pérez de León, A.A., 2012. Cattle tick vaccines: Many candidate antigens, but will a commercially viable product emerge? Int. J. Parasitol. 42, 421–427. https://doi.org/10.1016/j.ijpara.2012.04.003

Guerrero, F.D., Nene, V.M., George, J.E., Barker, S.C., Willadsen, P., 2006. Sequencing a new target genome : The *Boophilus microplus* (*Acari : Ixodidae*) Genome Project. J. Med. Entomol. 43, 9–16.

Hall, L.M., Malcolm, C.A., 1991. The acetylcholinesterase gene of *Anopheles stephensi*. Cell Mol Neurobiol 11, 131–141.

Hardie, R.C., Martin, F., Cochrane, G.W., Juusola, M., Georgiev, P., Raghu, P., 2002. Molecular basis of amplification in *Drosophila* phototransduction: roles for G protein, Phospholipase C, and Diacylglycerol Kinase. Neuron 36, 689–701. https://doi.org/10.1016/S0896-6273(02)01048-6

Harel, M., Kryger, G., Rosenberry, T.L., Mallender, W.D., Lewis, T., Fletcher, R.J., Guss, J.M., Silman, I., Sussman, J.L., 2000. Three-dimensional structures of *Drosophila melanogaster* Acetylcholinesterase and of its complexes with two potent inhibitors. Protein Sci. 9, 1063–1072. https://doi.org/10.1110/ps.9.6.1063

Hayashi, Y., Sexton, T.R., Dejima, K., Perry, D.W., Takemura, M., Kobayashi, S., Nakato, H., Harrison, D.A., 2012. Glypicans regulate JAK/STAT signalling and distribution of the unpaired morphogen. Development 139, 4162–4171. https://doi.org/10.1242/dev.078055

He, H., Chen, A.C., Davey, R.B., Ivie, G.W., George, J.E., 1999. Identification of a point mutation in the para-type sodium channel gene from a pyrethroid-resistant cattle tick. Biochem. Biophys. Res. Commun. 261, 558–61. https://doi.org/10.1006/bbrc.1999.1076

Heinz, D.W., Ryan, M., Smith, M.P., Weaver, L.H., Keana, J.F.W., Hayes Griffith, O., 1996. Crystal structure of Phosphatidylinositol-specific Phospholipase C from *Bacillus cereus* in complex with glucosaminyl(α1→6)-D-myoinositol, an essential fragment of GPI anchors. Biochemistry 35, 9496–9504. https://doi.org/10.1021/bi9606105

Hiramoto, M., Hiromi, Y., 2006. ROBO directs axon crossing of segmental boundaries by suppressing responsiveness to relocalized Netrin. Nat. Neurosci. 9, 58–66. https://doi.org/10.1038/nn1612

Holmes, S.P., Barhoumi, R., Nachman, R.J., Pietrantonio, P. V., 2003. Functional analysis of a G protein-coupled receptor from the Southern cattle tick *Boophilus microplus* (*Acari: Ixodidae*) identifies it as the first arthropod myokinin receptor. Insect Mol. Biol. 12, 27–38. https://doi.org/10.1046/j.1365-2583.2003.00384.x

Holmes, S.P., He, H., Chen, A.C., Lvie, G.W., Pietrantonio, P. V., 2000. Cloning and transcriptional expression of a leucokinin-like peptide receptor from the Southern cattle tick, *Boophilus microplus* (*Acari: Ixodidae*). Insect Mol. Biol. 9, 457–465. https://doi.org/10.1046/j.1365-2583.2000.00208.x

Horak, I.G., Jordaan, A.J., Nel, P.J., Van Heerden, J., Heyne, H., Van Dalen, E.M., 2015. Distribution of endemic and introduced tick species in Free State Province, South Africa. J. S. Afr. Vet. Assoc. 86, 1–9. https://doi.org/10.4102/jsava.v86i1.1255

Horgan, A. M., Lagrange, M.T., Copenhaver, P.F., 1994. Developmental expression of G proteins in a migratory population of embryonic neurons. Development 120, 729–42.

Howard, C.W., 1908. A list of the ticks of South Africa, with descriptions and keys to all the forms known., Annals of the Transvaal Museum. Transvaal Museum, [Pretoria] : https://doi.org/10.5962/bhl.title.84612

Hüe, T., Petermann, J., Bonnefond, R., Mermoud, I., Rantoen, D., Vuocolo, T., 2017. Experimental efficacy of a vaccine against *Rhipicephalus australis*. Exp. Appl. Acarol. 73, 245–256. https://doi.org/10.1007/s10493-017-0184-0

Hurtado, O., Giraldo-Ríos, C., 2018. Economic and health impact of the ticks in production animals, in: Ticks and Tick-Borne Pathogens. IntechOpen. https://doi.org/10.5772/intechopen.81167

Ikonen, E., 2001. Roles of lipid rafts in membrane transport. Curr. Opin. Cell Biol. 13, 470–477.

Ishihara, A., Hou, Y., Jacobson, K., 1987. The Thy-1 antigen exhibits rapid lateral diffusion in the plasma membrane of rodent lymphoid cells and fibroblasts. Proc. Natl. Acad. Sci. 84, 1290–1293. https://doi.org/10.1073/pnas.84.5.1290

Itoh, M., I, S.T., Yamamoto, H., Tomino, S., Eguchi, M., 1991. Cloning and sequence analysis of membrane-bound alkaline phosphatase cDNA of the silkworm, *Bombyx mori*. Biochimica et Biophysica Acta 1129, 135–138.

Izumi, Y., Ohta, N., Itoh-Furuya, A., Fuse, N., Matsuzaki, F., 2004. Differential functions of G protein and Baz-aPKC signaling pathways in *Drosophila* neuroblast asymmetric division. J. Cell Biol. 164, 729–738. https://doi.org/10.1083/jcb.200309162

Jang, I.H., Chosa, N., Kim, S.H., Nam, H.J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., Brey, P.T., Lee, W.J., 2006. A Spätzle-processing enzyme required for Toll signalling activation in *Drosophila* innate immunity. Dev. Cell 10, 45–55. https://doi.org/10.1016/j.devcel.2005.11.013

Jeanmonod, D.J., Rebecca, Suzuki, K. *et al*., 2018. Economic and health impact of the ticks in production animals, in: Ticks and Tick-Borne Pathogens. p. 64. https://doi.org/10.5772/32009

Kaewmongkol, S., Kaewmongkol, G., Inthong, N., Lakkitjaroen, N., Sirinarumitr, T., Berry, C.M., Jonsson, N.N., Stich, R.W., Jittapalapong, S., 2015. Variation among Bm86 sequences in *Rhipicephalus* (*Boophilus*) *microplus* ticks collected from cattle across Thailand. Exp. Appl. Acarol. 66, 247–256. https://doi.org/10.1007/s10493-015-9897-0

Kain, P., Chakraborty, T.S., Sundaram, S., Siddiqi, O., Rodrigues, V., Hasan, G., 2008. Reduced odour responses from antennal neurons of Gq, Phospholipase C, and rdgA mutants in *Drosophila* support a role for a phospholipid intermediate in insect olfactory transduction. J. Neurosci. 28, 4745–4755. https://doi.org/10.1523/JNEUROSCI.5306-07.2008

Kamani, J., Apanaskevich, D.A., Gutiérrez, R., Nachum-Biala, Y., Baneth, G., Harrus, S., 2017. Morphological and molecular identification of *Rhipicephalus* (*Boophilus*) *microplus* in Nigeria, West Africa: a threat to livestock health. Exp. Appl. Acarol. 73, 283–296. https://doi.org/10.1007/s10493-017-0177-z

Karlstrom, R.O., Wilder, L.P., Bastiani, M.J., 1993. Lachesin: an immunoglobulin superfamily protein whose expression correlates with neurogenesis in grasshopper embryos. Development 118, 509–522.

Katanaev, V.L., Ponzielli, R., Sémériva, M., Tomlinson, A., 2005. Trimeric G protein-dependent frizzled signaling in *Drosophila*. Cell 120, 111–122. https://doi.org/10.1016/j.cell.2004.11.014

Katanayeva, N., Kopein, D., Portmann, R., Hess, D., Katanaev, V.L., 2010. Competing activities of Heterotrimeric G proteins in *Drosophila* wing maturation. PLoS One 5. https://doi.org/10.1371/journal.pone.0012331

Kaziro, Y., Itoh, H., Kozasa, T., Nakafuku, M., Satoh, T., 1991. Signal-transducing structure and function of GTP-binding proteins. Annu. Rev. Biochem. 60, 349–400.

Kemp, D.H., Agbede, R.I.S., Johnston, L.A.Y., Gough, J.M., 1986. Immunisation of cattle against *Boophilus microplus* using extracts derived from adult female ticks : Feeding and survival of the parasite on vaccinated cattle. Int. J. Parasitol. 16, 115–120.

Kemp, D.H., Hughes, S., Binnington, K.C., Bird, P., Nolan, J., 1990. Mode of action of CGA 157419 on the cattle tick *Boophilus microplus*.

Kim, Y.S., Ryu, J.H., Han, S.J., Choi, K.H., Nam, K.B., Jang, I.H., Lemaitre, B., Brey, P.T., Lee, W.J., 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide

and β-1,3-glucan that mediates the signalling for the induction of innate immune genes in *Drosophila melanogaster* cells. J. Biol. Chem. 275, 32721–32727. https://doi.org/10.1074/jbc.M003934200

Kimura, K. -i., 2004. Activation of the cAMP/PKA signalling pathway is required for post-ecdysial cell death in wing epidermal cells of *Drosophila melanogaster*. Development 131, 1597–1606. https://doi.org/10.1242/dev.01049

King, J., Keim, M., Teo, R., Weening, K.E., Kapur, M., McQuillan, K., Ryves, J., Rogers, B., Dalton, E., Williams, R.S.B., Harwood, A.J., 2010. Genetic control of lithium sensitivity and regulation of inositol biosynthetic genes. PLoS One 5. https://doi.org/10.1371/journal.pone.0011151

Kinoshita, T., 2016. Glycosylphosphatidylinositol (GPI) Anchors : Biochemistry and Cell Biology : Introduction to a thematic review series 57, 4–5. https://doi.org/10.1194/jlr.E065417

Kinoshita, T., Fujita, M., 2009. Chapter 1 Overview of GPI Biosynthesis. Enzymes 26, 1–30. https://doi.org/10.1016/S1874-6047(09)26001-X

Kiper, I., 2013. Two-hybrid analysis and functional annotation of Bm86 and ATAQ from *Rhipicephalus microplus*. University of Pretoria.

Kita, T., Hayashi, T., Ohtani, T., Takao, H., Takasu, H., Liu, G., Ohta, H., Ozoe, F., Ozoe, Y., 2017. Amitraz and its metabolite differentially activate α - and β -adrenergic-like octopamine receptors. Pest Manag. Sci. 73, 984–990. https://doi.org/10.1002/ps.4412

Knight, P.J.K., Knowles, B.H., Ellar, D.J., 1995. Molecular cloning of an insect aminopeptidase N that serves as a receptor for *Bacillus thuringiensis* CryIA(c) toxin. J. Biol. Chem. https://doi.org/10.1074/jbc.270.30.17765

Kocan, K.M., Fuente, J. De, Alberto, A., Meléndez, R.D., Fuente, D., Guglielmone, A. a, Mele, R.D., 2003. Antigens and alternatives for control of Anaplasma marginale infection in cattle. Clin. Microbiol. Rev. 16, 698–712. https://doi.org/10.1128/CMR.16.4.698

Koh, K., Joiner, W.J., Wu, M.N., Yue, Z., Smith, C.J., Sehgal, A., 2008. Identification of Sleepless, a sleep-promoting factor. Science 321, 372–6. https://doi.org/10.1126/science.1155942

Kühn, B., Kühn, B., Schmid, A., Schmid, A., Harteneck, C., Harteneck, C., Gudermann, T., Gudermann, T., Schultz, G., Schultz, G., 1996. G proteins of the Gq family couple the H2 histamine receptor to Phospholipase C. Mol. Endocrinol. 10, 1697–1707. https://doi.org/10.1210/mend.10.12.8961278

Kumar, A., Garg, R., Yadav, C.L., Vatsya, S., Kumar, R.R., Sugumar, P., Chandran, D., Mangamoorib, L.N., Bedarkar, S.N., 2009. Immune responses against recombinant tick antigen, Bm95, for the control of *Rhipicephalus* (*Boophilus*) *microplus* ticks in cattle. Vet. Parasitol. 165, 119–124. https://doi.org/10.1016/j.vetpar.2009.06.030

Landry, Y., Niederhoffer, N., Sick, E., Gies, J., 2006. Heptahelical and other G-protein-coupled receptors (GPCRs) signalling. Curr. Med. Chem. 13, 51–63. https://doi.org/10.2174/092986706775197953

Lauc, G., Heffer-Lauc, M., 2006. Shedding and uptake of gangliosides and glycosylphosphatidylinositol-anchored proteins. Biochim. Biophys. Acta - Gen. Subj. 1760, 584–602. https://doi.org/10.1016/j.bbagen.2005.11.014

Li, X., Fan, D., Zhang, W., *et al.*, 2015. Outbred genome sequencing and CRISPR/Cas9 gene editing in Butterflies. Nat. Commun. 6, 1–10. https://doi.org/10.1038/ncomms9212

Liao, M., Zhou, J., Hatta, T., Umemiya, R., Miyoshi, T., Tsuji, N., Xuan, X., Fujisaki, K., 2007. Molecular characterisation of *Rhipicephalus* (*Boophilus*) *microplus* Bm86 homologue from *Haemaphysalis longicornis* ticks. Vet. Parasitol. 146, 148–157. https://doi.org/10.1016/j.vetpar.2007.01.015

Linder, M.E., Ewald, D. a, Miller, R.J., Gilman, a G., 1990. Purification and characterization of Go alpha and three types of Gi alpha after expression in *Escherichia coli*. J. Biol. Chem. 265, 8243–8251.

Llimargas, M., 2004. Lachesin is a component of a septate junction-based mechanism that controls tube size and epithelial integrity in the *Drosophila* tracheal system. Development 131, 181–190. https://doi.org/10.1242/dev.00917

Low, M.G., Stiernberg, J., Waneck, G.L., Flavell, R.A., Kincade, P.W., 1988. Cell-specific heterogeneity in sensitivity of phosphatidylinositol-anchored membrane antigens to release by Phospholipase C. J. Immunol. Methods 113, 101–111. https://doi.org/10.1016/0022-1759(88)90386-9

Macedo, F., Marsico, E.T., Conte-Júnior, C.A., Furtado, L. de A., Brasil, T.F., Pereira Netto, A.D., 2015. Short communication: Macrocyclic lactone residues in butter from Brazilian markets. J. Dairy Sci.

98, 3695–3700. https://doi.org/10.3168/jds.2014-9130

Madore, N., 1999. Functionally different GPI proteins are organized in different domains on the neuronal surface. EMBO J. 18, 6917–6926. https://doi.org/10.1093/emboj/18.24.6917

Makary, M.A., Kaczmarski, K., Nachman, K., 2018. A call for doctors to recommend antibiotic-free foods: agricultural antibiotics and the public health crisis of antimicrobial resistance. J. Antibiot. (Tokyo). 71, 685–687. https://doi.org/10.1038/s41429-018-0062-y

Mans, B.J., Featherston, J., de Castro, M.H., Pienaar, R., 2017. Gene duplication and protein evolution in tick-host interactions. Front. Cell. Infect. Microbiol. 7, e00413. https://doi.org/10.3389/fcimb.2017.00413

Marimuthu, S., Rahuman, A.A., Jayaseelan, C., Kirthi, A.V., Santhoshkumar, T., Velayutham, K., Bagavan, A., Kamaraj, C., Elango, G., Iyappan, M., Siva, C., Karthik, L., Rao, K.V.B., 2013. Acaricidal activity of synthesised titanium dioxide nanoparticles using *Calotropis gigantea* against *Rhipicephalus microplus* and *Haemaphysalis bispinosa.* Asian Pac. J. Trop. Med. 6, 682–688. https://doi.org/10.1016/S1995-7645(13)60118-2

Maritz-Olivier, C., Louw, A.I., Neitz, A.W.H., 2005. Similar mechanisms regulate protein exocytosis from the salivary glands of *Ixodid* and *Argasid* ticks. J. Insect Physiol. 51, 1390–1396. https://doi.org/10.1016/j.jinsphys.2005.08.012

Maritz-Olivier, C., Stutzer, C., Jongejan, F., Neitz, A.W.H., Gaspar, A.R.M., 2007. Tick anti-hemostatics: targets for future vaccines and therapeutics. Trends Parasitol. 23, 397–407. https://doi.org/10.1016/j.pt.2007.07.005

Mazor, T., Pankov, A., Song, J.S., Costello, J.F., 2016. Intratumoral heterogeneity of the epigenome. Cancer Cell 29, 440–451. https://doi.org/10.1016/j.ccell.2016.03.009

McLeod, R., Kristjanson, P., 1999. Economic impact of ticks and tick-borne diseases to livestock in Africa, Asia and Australia., Final report of joint esys/ILRI/ACIAR Tick Cost project. Nairobi. https://doi.org/10.1016/S0264-410X(12)01439-9

Medof, B.Y.M.E., Kinoshita, T., Nussenzweig, V., 1984. Inhibition of complement of cells after activation on the of incorporation (Daf) into their membranes. J. Exp. Med. 160, 1558–1578.

Medof, M.E., Nagarajan, S., Tykocinski, M.L., 1996. Cell-surface engineering with GPI-anchored proteins. FASEB J. 10, 574–586. https://doi.org/10.1096/fasebj.10.5.8621057

Meisenhelder, J., Suh, P.-G., Rhee, S.G., Hunter, T., 1989. Phospholipase C-γ is a substrate for the PDGF and EGF receptor protein-tyrosine kinases in vivo and in vitro. Cell 57, 1109–1122. https://doi.org/10.1016/0092-8674(89)90048-2

Merino, O., Alberdi, P., Pérez de la Lastra, J.M., de la Fuente, J., 2013. Tick vaccines and the control of tick-borne pathogens. Front. Cell. Infect. Microbiol. 3, 1–10. https://doi.org/10.3389/fcimb.2013.00030

Miller, R., Estrada-Peña, A., Almazán, C., Allen, A., Jory, L., Yeater, K., Messenger, M., Ellis, D., Pérez de León, A.A., 2012. Exploring the use of an anti-tick vaccine as a tool for the integrated eradication of the cattle fever tick, *Rhipicephalus* (*Boophilus*) *annulatus*. Vaccine 30, 5682–5687. https://doi.org/10.1016/j.vaccine.2012.05.061

Nakato, H., Futch, T.A., Selleck, S.B., *et al*., 1995. The division abnormally delayed (dally) gene: a putative integral membrane proteoglycan required for cell division patterning during postembryonic development of the nervous system in *Drosophila*. Development 121, 3687–702. https://doi.org/10.1073/pnas.58.3.1112

Nasuti, C., Coman, M.M., Olek, R.A., Fiorini, D., Verdenelli, M.C., Cecchini, C., Silvi, S., Fedeli, D., Gabbianelli, R., 2016. Changes in faecal microbiota in rats exposed to permethrin during postnatal development. Environ. Sci. Pollut. Res. 23, 10930–10937. https://doi.org/10.1007/s11356-016-6297-x

Nations, U., 2018. The 2018 Revision of the World Urbanization Prospects.

Ndhlovu, D.N., Makaya, P. V, Penzhorn, B.L., 2009. Tick infestation, and udder and teat damage in selected cattle herds of Matabeleland South, Zimbabwe. Onderstepoort J. Vet. Res. 76, 235–248. https://doi.org/10.4102/ojvr.v76i2.48

Neer, E.J., 1995. Heterotrimeric C proteins: Organizers of transmembrane signals. Cell 80, 249–257. https://doi.org/10.1016/0092-8674(95)90407-7

Neves, S.R., Ram, P.T., Iyengar, R., 2002. G Protein Pathways. Science (80-.). 296, 1636–1639. https://doi.org/10.1126/science.1071550

Nijhof, A.M., Balk, J.A., Postigo, M., Rhebergen, A.M., Taoufik, A., Jongejan, F., 2010. Bm86 homologues and novel ATAQ proteins with multiple epidermal growth factor (EGF)-like domains from hard and soft ticks. Int. J. Parasitol. 40, 1587–1597. https://doi.org/10.1016/j.ijpara.2010.06.003

Nijhof, A.M., Taoufik, A., De La Fuente, J., Kocan, K.M., De Vries, E., Jongejan, F., 2007. Gene silencing of the tick protective antigens, Bm86, Bm91 and subolesin, in the one-host tick *Boophilus microplus* by RNA interference. Int. J. Parasitol. 37, 653–662. https://doi.org/10.1016/j.ijpara.2006.11.005

Nikolaidou, K.K., Barrett, K., 2004. A Rho GTPase signalling pathway is used reiteratively in epithelial folding and potentially selects the outcome of Rho activation. Curr. Biol. 14, 1822–1826. https://doi.org/10.1016/j.cub.2004.09.080

Nuttall, P.A., Trimnell, A.R., Kazimirova, M., Labuda, M., 2006. Exposed and concealed antigens as vaccine targets for controlling ticks and tick-borne diseases. Parasite Immunol. 28, 155–163. https://doi.org/10.1111/j.1365-3024.2006.00806.x

Nyangiwe, N., Harrison, A., Horak, I.G., 2013. Displacement of *Rhipicephalus decoloratus* by *Rhipicephalus microplus* (*Acari: Ixodidae*) in the Eastern Cape Province, South Africa. Exp. Appl. Acarol. 61, 371–382. https://doi.org/10.1007/s10493-013-9705-7

Nyangiwe, N., Horak, I.G., Mescht, L. Van Der, Matthee S., der Mescht, V., 2017. Range expansion of the economically important Asiatic blue tick , *Rhipicephalus microplus*, in South Africa. J. S. Afr. Vet. Assoc. 88, 1–7. https://doi.org/10.4102/jsava.v88i0.1482

Nyangiwe, Nkululeko, Matthee, C., Horak, I., Matthee, S., 2013. First record of the pantropical blue tick *Rhipicephalus microplus* in Namibia. Exp. Appl. Acarol. 61, 503–507. https://doi.org/10.1007/s10493-013-9717-3

Nyangiwe, N., Yawa, M., Muchenje, V., 2018. Driving forces for changes in geographic range of cattle ticks (*Acari: Ixodidae*) in Africa: A review. S. Afr. J. Anim. Sci. 48, 829. https://doi.org/10.4314/sajas.v48i5.4

Nygaard, S., Zhang, G., Schiøtt, M., Li, C., Wurm, Y., Hu, H., Zhou, J., Ji, L., Qiu, F., Rasmussen, M., Pan, H., Hauser, F., Krogh, A., Grimmelikhuijzen, C.J.P., Wang, J., Boomsma, J.J., 2011. The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. Genome Res. 21, 1339–1348. https://doi.org/10.1101/gr.121392.111

O'Sullivan, N.C., Dräger, N., O'Kane, C.J., 2013. Characterisation of the *Drosophila* Atlastin interactome reveals VCP as a functionally related interactor. J. Genet. Genomics 40, 297–306. https://doi.org/10.1016/j.jgg.2013.04.008

Özkan, E., Carrillo, R.A., Eastman, C.L., Weiszmann, R., Waghray, D., Johnson, K.G., Zinn, K., Celniker, S.E., Garcia, K.C., 2013. An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. Cell 154. https://doi.org/10.1016/j.cell.2013.06.006

Panáková, D., Sprong, H., Marois, E., Thiele, C., Eaton, S., 2005. Lipoprotein particles are required for Hedgehog and Wingless signalling. Nature 435, 58–65. https://doi.org/10.1038/nature03504

Paulick, M., Bertozzi, C., 2008. The Glycosylphosphatidylinositol Anchor: A complex membrane-anchoring structure for proteins. Biochemistry 6991–7000. https://doi.org/10.1021/bi8006324

Peles, E., Nativ, M., Campbell, P.L., Sakurai, T., Martinez, R., Levt, S., Clary, D.O., Schilling, J., Barnea, G., Plowman, G.D., Grumet, M., Schlessinger, J., 1995. The carbonic anhydrase domain of receptor tyrosine phosphatase β is a functional ligand for the axonal cell recognition molecule contactin. Cell 82, 251–260. https://doi.org/10.1016/0092-8674(95)90312-7

Peles, E., Nativ, M., Lustig, M., Grumet, M., Schilling, J., Martinez, R., Plowman, G.D., Schlessinger, J., 1997. Identification of a novel contactin-associated transmembrane receptor with multiple domains implicated in protein-protein interactions. EMBO J. 16, 978–988. https://doi.org/10.1093/emboj/16.5.978

Picinin, L.C.A., Toaldo, I.M., Hoff, R.B., Souza, F.N., Leite, M.O., Fonseca, L.M., Diniz, S.A., Silva, M.X., Haddad, J.P.A., Cerqueira, M.M.O.P., Luiz, M.T.B., 2016. Climate conditions associated with the occurrence of pyrethroid residues in bulk milk tank. Arq. Bras. Med. Vet. e Zootec. 68, 1721–1726. https://doi.org/10.1590/1678-4162-8880

Pike, L.J., 2004. Lipid rafts: heterogeneity on the high seas. Biochem. J. 378, 281–292. https://doi.org/10.1042/bj20031672

Pipano, E., Alekceev, E., Galker, F., Fish, L., Samish, M., Shkap, V., 2003. Immunity against *Boophilus annulatus* induced by the Bm86 (Tick-GARD) vaccine. Exp. Appl. Acarol. 29, 141–149. https://doi.org/10.1023/A:1024246903197

Portillo, A., Pérez-Martínez, L., Santibáñez, S., Blanco, J.R., Ibarra, V., Oteo, J.A., 2007. Short report: detection of *Rickettsia africae* in *Rhipicephalus* (*Boophilus*) *decoloratus* ticks from the Republic of Botswana, South Africa. Am. J. Trop. Med. Hyg. 77, 376–377.

Potgieter, F., 1979. Epizootiology and control of Anaplasmosis in South Africa. J. South African Vet. Assoc. 50, 367–372.

Pothmann, D., Poppert, S., Rakotozandrindrainy, R., Hogan, B., Mastropaolo, M., Thiel, C., Silaghi, C., 2016. Prevalence and genetic characterisation of *Anaplasma marginale* in zebu cattle (*Bos indicus*) and their ticks (*Amblyomma variegatum*, *Rhipicephalus microplus*) from Madagascar. Ticks Tick. Borne. Dis. 7, 1116–1123. https://doi.org/10.1016/j.ttbdis.2016.08.013

Qi, Z., Wei, Z., 2017. Microbial flora analysis for the degradation of beta-cypermethrin. Environ. Sci. Pollut. Res. 24, 6554–6562. https://doi.org/10.1007/s11356-017-8370-5

Qin, Y., Jatamunua, F., Zhang, J., Li, Y., Han, Y., Zou, N., Shan, J., Jiang, Y., Pan, C., 2017. Analysis of sulfonamides, tilmicosin and avermectins residues in typical animal matrices with multi-plug filtration cleanup by liquid chromatography-tandem mass spectrometry detection. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 1053, 27–33. https://doi.org/10.1016/j.jchromb.2017.04.006

Quan, F., Wolfgang, W.J., Forte, M., 1993. A *Drosophila* G-protein alpha subunit, Gf alpha, expressed in a spatially and temporally restricted pattern during *Drosophila* development. Proc. Natl. Acad. Sci. U. S. A. 90, 4236–40.

Rand, K.N., Moore, T., Sriskantha, A., Spring, K., Tellam, R., Willadsen, P., Cobon, G.S., 1989. Cloning and expression of a protective antigen from the cattle tick *Boophilus microplus*. Proc. Natl. Acad. Sci. U. S. A. 86, 9657–61. https://doi.org/10.1073/pnas.86.24.9657

Ranz, J.M., Maurin, D., Chan, Y.S., Von Grotthuss, M., Hillier, L.D.W., Roote, J., Ashburner, M., Bergman, C.M., 2007. Principles of genome evolution in the *Drosophila melanogaster* species group. PLoS Biol. 5, 1366–1381. https://doi.org/10.1371/journal.pbio.0050152

Rebecchi, M.J., Pentyala, S.N., 2000. Structure, Function, and Control of Phosphoinositide-Specific Phospholipase C 80, 1291–1335. https://doi.org/0031-9333/00

Reck, J., Klafke, G.M., Webster, A., Dall'Agnol, B., Scheffer, R., Souza, U.A., Corassini, V.B., Vargas, R., dos Santos, J.S., de Souza Martins, J.R., 2014. First report of fluazuron resistance in *Rhipicephalus microplus*: A field tick population resistant to six classes of acaricides. Vet. Parasitol. 201, 128–136. https://doi.org/10.1016/j.vetpar.2014.01.012

Rees, J.S., Lowe, N., Armean, I.M., Roote, J., Johnson, G., Drummond, E., Spriggs, H., Ryder, E., Russell, S., Johnston, D.S., Lilley, K.S., 2011. *In Vivo* Analysis of proteomes and interactomes using parallel affinity capture (iPAC) coupled to mass spectrometry. Mol. Cell. Proteomics 10, M110.002386. https://doi.org/10.1074/mcp.M110.002386

Rhee, S.G., Bae, Y.S., 1997. Regulation of Phosphoinositide-specific Phospholipase C isozymes. J. Biol. Chem. 272, 15045–15048. https://doi.org/10.1074/jbc.272.24.15045

Richards, S., 2005. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. Genome Res. 15, 1–18. https://doi.org/10.1101/gr.3059305

Richards, S., Gibbs, R.A., Gerardo, N.M., *et al*., 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol. 8. https://doi.org/10.1371/journal.pbio.1000313

Richards, S.A., Stutzer, C., Bosman, A.-M., Maritz-Olivier, C., 2015. Transmembrane proteins – Mining the cattle tick transcriptome. Ticks Tick. Borne. Dis. 6, 695–710. https://doi.org/10.1016/j.ttbdis.2015.06.002

Richardson, M.A., Smith, D.R.J., Kemp, D.H., Tellam, R.L., 1993. Native and baculovirus-expressed forms of the immunoprotective protein Bm86 from *Boophilus microplus* are anchored to the cell membrane by a glycosylphosphatidyl inositol linkage. Insect Mol. Biol. 1, 139–147. https://doi.org/10.1111/j.1365-2583.1993.tb00115.x

Rider, S.D., Morgan, M.S., Arlian, L.G., 2015. Draft genome of the scabies mite. Parasites and Vectors 8, 1–14. https://doi.org/10.1186/s13071-015-1198-2

Riding, G. a, Jarmey, J., McKenna, R. V, Pearson, R., Cobon, G.S., Willadsen, P., 1994. A protective

"concealed" antigen from *Boophilus microplus*. Purification, localisation, and possible function. J. Immunol. 153, 5158–5166.

Robbertse, L., Baron, S., van der Merwe, N.A., Madder, M., Stoltsz, W.H., Maritz-Olivier, C., 2016. Genetic diversity, acaricide resistance status and evolutionary potential of a *Rhipicephalus microplus* population from a disease-controlled cattle farming area in South Africa. Ticks Tick. Borne. Dis. 7, 595–603. https://doi.org/10.1016/j.ttbdis.2016.02.018

Robinson, P.J., 1991. Signal transduction by GPI-Anchored membrane proteins. Cell Biol. 15, 761–767.

Rodríguez, M., Rubiera, R., Penichet, M., Montesinos, R., Cremata, J., Falcón, V., Sánchez, G., Bringas, R., Cordovés, C., Valdés, M., Lleonart, R., Herrera, L., 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. J. Biotechnol. 33, 135–146. https://doi.org/10.1016/0168-1656(94)90106-6

Rützler, M., Lu, T., Zwiebel, L.J., 2006. Gα encoding gene family of the malaria vector mosquito *Anopheles gambiae*: Expression analysis and immunolocalisation of Gαq and Gαo in female antennae. J. Comp. Neurol. 499, 533–545. https://doi.org/10.1002/cne.21083

Saha, S., Anilkumar, A.A., Mayor, S., 2016. GPI-anchored protein organisation and dynamics at the cell surface. J. Lipid Res. 57, 159–175. https://doi.org/10.1194/jlr.R062885

Sauer, J.R., Essenberg, R.C., Bowman, A.S., 2000. Salivary glands in *Ixodid* ticks: control and mechanism of secretion. J. Insect Physiol. 46, 1069–1078.

Schaefer, M., Petronczki, M., Dorner, D., Forte, M., Knoblich, J.A., 2001. Heterotrimeric G proteins direct two modes of asymmetric cell division in the *Drosophila* nervous system. Cell 107, 183–194. https://doi.org/10.1016/S0092-8674(01)00521-9

Schetters, T., Bishop, R., Crampton, M., Kopáček, P., Lew-Tabor, A., Maritz-Olivier, C., Miller, R., Mosqueda, J., Patarroyo, J., Rodriguez-Valle, M., Scoles, G.A., de la Fuente, J., 2016. Cattle tick vaccine researchers join forces in CATVAC. Parasit. Vectors 9, 105. https://doi.org/10.1186/s13071-016-1386-8

Schulz, S., Huber, A., Schwab, K., Paulsen, R., 1999. A novel Gα isolated from *Drosophila* constitutes a visual G protein subunit of the fly compound eye. J. Biol. Chem. 274, 37605–37610. https://doi.org/10.1074/jbc.274.53.37605

Seeger, M.A., Haffley, L., Kaufman, T.C., 1988. Characterisation of Amalgam: A member of the immunoglobulin superfamily from *Drosophila*. Cell 55, 589–600. https://doi.org/10.1016/0092-8674(88)90217-6

Shibatohge, M., Kariya, K.I., Liao, Y., Hu, C.D., Watari, Y., Goshima, M., Shima, F., Kataoka, T., 1998. Identification of PLC210, a *Caenorhabditis elegans* Phospholipase C, as a putative effector of Ras. J. Biol. Chem. 273, 6218–6222. https://doi.org/10.1074/jbc.273.11.6218

Shieh, B.H., Zhu, M.Y., Lee, J.K., Kelly, I.M., Bahiraei, F., 1997. Association of INAD with NORPA is essential for controlled activation and deactivation of *Drosophila* phototransduction *in vivo*. Proc. Natl. Acad. Sci. U. S. A. 94, 12682–7. https://doi.org/VL - 94

Silatsa, B.A., Kuiate, J.-R., Njiokou, F., Simo, G., Feussom, J.-M.K., Tunrayo, A., Amzati, G.S., Bett, B., Bishop, R., Githaka, N., Opiyo, S.O., Djikeng, A., Pelle, R., 2019. A countrywide molecular survey leads to a seminal identification of the invasive cattle tick *Rhipicephalus* (*Boophilus*) *microplus* in Cameroon, a decade after it was reported in Cote d'Ivoire. Ticks Tick. Borne. Dis. 10, 585–593. https://doi.org/10.1016/j.ttbdis.2019.02.002

Šimo, L., Kazimirova, M., Richardson, J., Bonnet, S.I., 2017. The essential role of tick salivary glands and saliva in tick feeding and pathogen transmission. Front. Cell. Infect. Microbiol. 7, e00281. https://doi.org/10.3389/fcimb.2017.00281

Singh, N.K., Jyoti, Vemu, B., Prerna, M., Singh, H., Dumka, V.K., Sharma, S.K., 2016. Acaricidal activity of leaf extracts of *Dalbergia sissoo* (*Fabaceae*) against synthetic pyrethroid resistant *Rhipicephalus* (*Boophilus*) *microplus*. Res. Vet. Sci. 106, e01603002. https://doi.org/10.1016/j.rvsc.2016.03.002

Singh, R.K., Sanyal, P.K., Patel, N.K., Sarkar, A.K., Santra, A.K., Pal, S., Mandal, S.C., 2010. Fungus-benzimidazole interactions: A prerequisite to deploying egg-parasitic fungi *Paecilomyces lilacinus* and *Verticillium chlamydosporium* as biocontrol agents against *fascioliasis* and *amphistomiasis* in ruminant livestock. J. Helminthol. 84, 123–131. https://doi.org/10.1017/S0022149X09990344

Solomon, K.R., Rudd, C.E., Finberg, R.W., 1996. The association between glycosylphosphatidylinositol-anchored proteins and Heterotrimeric G protein alpha subunits in lymphocytes. Proc Natl Acad Sci

U S A 93, 6053–8. https://doi.org/10.1073/pnas.93.12.6053

Sonenshine, D.E., Roe, R.M., 2014. Biology of ticks. Oxford University Press.

Spickett, A.M., Malan, R., 1978. Genetic incompatibility between *Boophilus decoloratus* and *Boophilus microplus* and hybrid sterility of Australian and South African *Boophilus microplus*. Onderstepoort J. Vet. Res. 45, 149–153.

Statistics South Africa, 2016. Community Survey 2016 Agricultural households, Report no. 03-01-05 (2016).

Strigini, M., Cantera, R., Morin, X., Bastiani, M.J., Bate, M., Karagogeos, D., 2006. The IgLON protein Lachesin is required for the blood-brain barrier in *Drosophila*. Mol. Cell. Neurosci. 32, 91–101. https://doi.org/10.1016/j.mcn.2006.03.001

Stutzer, C., Richards, S.A., Ferreira, M., Baron, S., Maritz-Olivier, C., 2018. Metazoan parasite vaccines: present status and future prospects. Front. Cell. Infect. Microbiol. 8, Article 67. https://doi.org/10.3389/fcimb.2018.00067

Suh, P.-G., Park, J.-I., Manzoli, L., Cocco, L., Peak, J.C., Katan, M., Fukami, K., Kataoka, T., Ryu, S.Y.& S.H., 2008. Multiple roles of Phosphoinositide-specific Phospholipase C isozymes. BMB Rep. 41, 415–434.

Sungirai, M., Moyo, D.Z., De Clercq, P., Madder, M., 2016. Communal farmers' perceptions of tick-borne diseases affecting cattle and investigation of tick control methods practised in Zimbabwe. Ticks Tick. Borne. Dis. 7, 1–9. https://doi.org/10.1016/j.ttbdis.2015.07.015

Suzuki, K., Fujiwara, T.K., Edidin, M., Kusumi, A., 2007a. Dynamic recruitment of Phospholipase Cγ at transiently immobilised GPI-anchored receptor clusters induces IP 3 –Ca 2+ signalling: single-molecule tracking study 2. J. Cell Biol. 177, 731–742. https://doi.org/10.1083/jcb.200609175

Suzuki, K., Fujiwara, T.K., Sanematsu, F., Iino, R., Edidin, M., Kusumi, A., 2007b. GPI-anchored receptor clusters transiently recruit Lyn and Gα for temporary cluster immobilisation and Lyn activation: single-molecule tracking study 1. J. Cell Biol. 177, 717–730. https://doi.org/10.1083/jcb.200609174

Swift, R., 2011. The relationship between health and GDP in OECD countries in the very long run. Health Econ. 20, 306–322. https://doi.org/10.1002/hec.1590

Thammasittirong, A., Dechklar, M., Leetachewa, S., Pootanakit, K., Angsuthanasombat, C., 2011. *Aedes aegypti* membrane-bound alkaline phosphatase expressed in *Escherichia coli* retains high-affinity binding for *Bacillus thuringiensis* Cry4Ba toxin. Appl. Environ. Microbiol. 77, 6836–6840. https://doi.org/10.1128/AEM.05775-11

Thatheyus, A.J., Gnana Selvam, A.D., 2013. Synthetic Pyrethroids: toxicity and biodegradation. Appl. Ecol. Environ. Sci. 1, 33–36. https://doi.org/10.12691/aees-1-3-2

Torriani, A., 1968. Alkaline phosphatase subunits and their dimerisation *in vivo*. J. Bacteriol. 96, 1200–1207.

Trager, W., 1939. Acquired Immunity to Ticks. Am. Soc. Parasitol. 25, 57–81.

Trimnell, A.R., Hails, R.S., Nuttall, P.A., 2002. Dual-action ectoparasite vaccine targeting 'exposed' and 'concealed' antigens. Vaccine 20, 3560–3568. https://doi.org/10.1016/S0264-410X(02)00334-1

Ullrich, A., Schlessinger, J., 1990. Signal transduction by receptors with Tyrosine Kinase activity. Cell 61, 203–212. https://doi.org/10.1007/s00497-011-0177-9

Vega, Q.C., Cochet, C., Filhol, O., Chang, C.-P., Rhee, S.G., Gill, G.N., 1992. A site of tyrosine phosphorylation in the C terminus of the epidermal growth factor receptor is required to activate Phospholipase C. Mol. Cell Biol. 12, 128–135.

Vercruysse, J., Schetters, T.P.M., Knox, D.P., Willadsen, P., Claerebout, E., 2007. Control of parasitic disease using vaccines: an answer to drug resistance? Rev. Sci. Tech. 26, 105–15.

Vieira, J.N., Maia Filho, F.S., Ferreira, G.F., Mendes, J.F., Gonçalves, C.L., Villela, M.M., Pereira, D.I.B., Nascente, P.S., 2016. In vitro susceptibility of nematophagous fungi to antiparasitic drugs: interactions and implications for biological control. Brazilian J. Biol. 77, 476–479. https://doi.org/10.1590/1519-6984.15715

Walker, A.R., Bouattour, A., Camicas, J.L., Estrada-Peña, A., Horak, I.G., Latif, A.A., Pegram, R.G., Preston, P.M., 2003. Ticks of domestic animals in Africa: a guide to identification of species.

Bioscience Reports, Edinburgh Scotland, U.K.

Wang, L., Weber, A.N.R., Atilano, M.L., Filipe, S.R., Gay, N.J., Ligoxygakis, P., 2006. Sensing of Gram-positive bacteria in *Drosophila*: GNBP1 is needed to process and present peptidoglycan to PGRP-SA. EMBO J. 25, 5005–5014. https://doi.org/10.1038/sj.emboj.7601363

Wang, Y., Pennock, S., Chen, X., Wang, Z., 2002. Endosomal signalling of epidermal growth factor receptor stimulates signal transduction pathways leading to cell survival. Mol. Cell. Biol. 22, 7279–7290. https://doi.org/10.1128/MCB.22.20.7279

Watling, K.J., 2001. The SIGMA-RBI handbook of receptor classification and signal transduction, 4th Editio. ed. Sigma-RBI, Natick.

Wee, P., Wang, Z., 2017. Epidermal growth factor receptor cell proliferation signalling pathways. Cancers (Basel). 9, 52–97. https://doi.org/10.3390/cancers9050052

Willadsen, P., 2008. Anti-tick vaccines, in: Ticks: Biology, Disease and Control. Cambridge University Press, pp. 424–446. https://doi.org/10.1017/CBO9780511551802.020

Willadsen, P., Riding, G.A., Mckenna, R. V, Kemp, D.H., Tellam, R.L., Nielsen, J.N., Lahnstein, J., Cobon, G.S., Gough, J.M., 1989. Immunologic control of a parasitic arthropod. Identification of a protective antigen from *Boophilus microplus*. J. Immunol. 143, 1346–1351.

Williams, R.L., Katan, M., 1996. Structural views of phosphoinositide-specific phospholipase C: Signalling the way ahead. Structure 4, 1387–1394. https://doi.org/10.1016/S0969-2126(96)00146-3

Wolfgang, W.J., Clay, C., Parker, J., Delgado, R., Labarca, P., Kidokoro, Y., Forte, M., 2004. Signalling through Gsα is required for the growth and function of neuromuscular synapses in *Drosophila*. Dev. Biol. 268, 295–311. https://doi.org/10.1016/j.ydbio.2004.01.007

Wu, M., Liu, C.Z., Joiner, W.J., 2016. Structural analysis and deletion mutagenesis define regions of QUIVER/SLEEPLESS that are responsible for interactions with shaker- Type potassium channels and nicotinic acetylcholine receptors. PLoS One 11, 1–14. https://doi.org/10.1371/journal.pone.0148215

Wu, M., Robinson, J.E., Joiner, W.J., 2014. SLEEPLESS is a bifunctional regulator of excitability and cholinergic synaptic transmission. Curr. Biol. 24, 621–629. https://doi.org/10.1016/j.cub.2014.02.026

Wu, M.N., Joiner, W.J., Dean, T., Yue, Z., Smith, C.J., Chen, D., Hoshi, T., Sehgal, A., Koh, K., 2010. SLEEPLESS, a Ly-6/neurotoxin family member, regulates the levels, localisation and activity of Shaker. Nat. Neurosci. 13, 69–75. https://doi.org/10.1038/nn.2454

Yang, C.-C., 2012. Acute human toxicity of Macrocyclic Lactones. Curr. Pharm. Biotechnol. 13, 999–1003. https://doi.org/10.2174/138920112800399059

Yang, M.Y., Wang, Z., MacPherson, M., Dow, J.A.T., Kaiser, K., 2000. A novel *Drosophila* Alkaline Phosphatase specific to the ellipsoid body of the adult brain and the lower Malpighian (renal) tubule. Genetics 154, 285–297.

Yu, F., Wang, H., Qian, H., Kaushik, R., Bownes, M., Yang, X., Chia, W., 2005. Locomotion defects, together with Pins, regulates Heterotrimeric G-protein signalling during *Drosophila* neuroblast asymmetric divisions. Genes Dev. 19, 1341–1353. https://doi.org/10.1101/gad.1295505

Zhu, K.Y., Clark, J.M., 1995. Cloning and sequencing of a cDNA encoding acetylcholinesterase in Colorado potato beetle, *Leptinotarsa decemlineata* (Say). Insect Biochem. Mol. Biol. 25, 1129–1138.

Zinn, K., McAllister, L., Goodman, C.S., 1988. Sequence analysis and neuronal expression of fasciclin I in grasshopper and *Drosophila*. Cell 53, 577–587. https://doi.org/10.1016/0092-8674(88)90574-0

CHAPTER 2: MATERIALS AND METHODS

INTRODUCTION: AN OVERVIEW OF METHODS EMPLOYED IN THIS STUDY

Assembly of the *Rhipicephalus microplus* transcriptome for annotation

Quality control and de novo *assembly of RNAseq data*

RNA sequencing of larvae, nymphs and three tissues from adults (salivary glands, guts and ovaries) was performed by Dr C. Stutzer at the University of Pretoria in our research group. As there is no full genome sequence available for any *Rhipicephalus* tick species, *de novo* assemblies of the RNAseq sequence data were performed in this study.

The quality of the raw sequencing data could have been negatively affected and may exhibit various biases (Such as coverage and error bias (Ross *et al*., 2013)) from several factors. Firstly, anomalies originating in the sequencer itself, for example; could have led to the sequencer misreading or not reading the full length of transcripts (Conesa *et al*., 2016). The latter can be accessed via the sequencer's own automatically generated quality control (QC) report. Secondly, errors in the sequence or starting material which are identified with QC software. For example; a bias may result from the integrity of the RNA after extraction, which should have RNA integrity (RIN) numbers greater than 6, indicating minimal degradation of RNA before sequencing. A RIN < 6 may result in uneven coverage (Haidula, 2016), and/or during later steps in the reaction, some transcripts may be lost (Conesa *et al*., 2016). Quality assessment of the raw RNAseq sequence data is the first step in the *de novo* assembly process, utilizing software such as FastQC (Andrews, 2010), RNA-SeQC (Deluca *et al*., 2012), Qualimap 2 (Okonechnikov *et al*., 2016) and MultiQC (Ewels *et al*., 2016). The QC software will usually assess the quality of each base call in the read, the overall quality of each sequence and the quality of the data set as a whole (Andrews, 2010), and some will even provide automatic comparisons between samples (Ewels *et al*., 2016; Okonechnikov *et al*., 2016). For example, FastQC assesses various quality metrics, such as sample randomness and diversity, base call quality, GC content, the number of ambiguous base calls, read length uniformity and lastly over-represented samples and k-mers, to provide colour coded outputs to indicate whether the relevant data set passes, fails or is borderline for the respective metrics (Andrews, 2010).

Following quality assessment the sequence data is processed to remove aberrant sequence stretches and filter out low-quality reads using Trimmomatic; which is a multithreaded command-line tool that can be used to filter read quality and trim and remove adaptors from Illumina FastQC data (Bolger, A. M., Lohse, M., & Usadel, 2014). The first step in Trimmomatic is called Illuminaclip. This step removes illumina adapters, which are 100-1000 base pair long sequences that are ligated to either end of the library sequences to enable sequencing (www.illumina.com). A common source of such adaptor contaminants occurs when the beginning of the read contains a sequence of interest, but at the end of the fragment, the sequencer continued to read the adaptor's sequence, resulting in a partial or full adapter sequence towards the 3' end of the read (www.illumina.com). Trimmomatic is used to identify the start of adapter sequences and remove them from the read (Bolger, A. M., Lohse, M., & Usadel, 2014).

Once IIlumina adaptors are removed, Trimmomatic can remove leading and trailing low-quality bases (below quality 3 for example) and scan the reads with a sliding window (4 bases for example) while cutting each read when the quality per base is low (below 15 for example). Lastly, Trimmomatic can drop reads entirely which are less than, for example, 36 bases long after all trimming steps have been performed (Bolger, A. M., Lohse, M., & Usadel, 2014).

Only once the sequencing data has been sufficiently prepared can the *de novo* assembly itself be performed. To assemble the reads, software packages such as Oases (Schulz *et al*., 2012) or Trinity (Grabherr *et al*., 2013) are used. The former is considered an assembly-first approach, as opposed to a mapping first approach (Schulz *et al*., 2012). The latter, Trinity, used in this study, makes sequential use of three software modules; Inchworm, Chrysalis, and Butterfly (Grabherr *et al*., 2011). Inchworm assembles the read data set to produce linear contigs. Once contigs have been formed, Trinity uses Chrysalis to pool the contigs and build individual de Bruijn graphs from each pool. Butterfly trims spurious edges and compacts linear paths (middle) based on the de Bruijn graphs from Chrysalis, subsequently reconciling the graphs with the reads and pairs to produce one linear sequence for each transcript in the graph (Grabherr *et al*., 2011). Trinity has shown low base-error rates and can capture multiple isoforms (Grabherr *et al*., 2011).

TransDecoder (https://github.com/TransDecoder/TransDecoder/wiki) is included as a part of the Trinity package and is used to predict protein-coding regions from Trinity reconstructed transcripts (Haas *et al*., 2013). It does this based on nucleotide

composition and predicting open reading frames in transcriptomic data, with likely open reading frames based on specific criteria. These criteria include; that a minimum length open reading frame is found in a transcript sequence, a log-likelihood score greater than 0, that the greatest above coding score is when the open reading frame is scored in the first reading frame compared to subsequent reading frame scores and that the most extended open reading frame is reported. A Position-Specific Scoring Matrix (PSSM) is used to refine the start codon prediction and the optional criteria that the putative peptide matches a Pfam domain above the noise cut-off score (https://github.com/TransDecoder/TransDecoder/wiki)

Following assembly, all transcripts shorter than 200 base pairs are generally removed. A representative transcriptome is created by removing highly-homologous sequences. This removal is done by CD-HIT (95% identity). CD-HIT is a program to cluster and compare sequence data (Fu *et al*., 2012). CD-HIT begins by identifying the longest input sequence as the first cluster representative. It then processes the remaining sequences from longest to shortest and classifies each sequence as either redundant or representative, depending on its similarity to the existing representatives (longer sequences) based on the word counts (Fu *et al*., 2012). In this manner, unnecessary sequence alignments are removed, in this case, those less than 200 base pairs and with high homology to longer reads (Fu *et al*., 2012).

The assembly can then be quantitatively assessed against a reference data set, previously done with the Core Eukaryotic Genes Mapping Approach (CEGMA) (Parra *et al*., 2007), but most recently by the Benchmarking Universal Single-Copy Orthologs (BUSCO) (Simão *et al*., 2015) that has replaced CEGMA. BUSCO is the standard programme to assess the quality and completeness of the RNAseq data. The assessment is based on a comparison of the transcriptome to a reference transcriptome or genome and uses hidden Markov models (HMM) to investigate the data set for specific, expected signature sequences in the reference BUSCO data set (Simão *et al*., 2015). The output reports on the number of signatures found in the data set under investigation (Simão *et al*., 2015). A general eukaryotic reference set of 248 genes are applied for all eukaryotes, while more lineage-specific genes provide more focused analyses of specific classes (Simão *et al*., 2015). An example is the availability of an arthropod gene set containing 1 066 arthropod-specific genes. These are identified in the assembly and are considered the core-genes for the class in which Acari falls.

An assembly is classified as 'complete' when the transcript lengths are within two standard deviations of the BUSCO group (Arthropod transcriptome) mean length (Simão *et al*., 2015). Transcripts with more than one copy of a signature sequence are classified as 'duplicated'. The unexpected occurrence of many duplicates may indicate errors in the assembly of haplotypes (Simão *et al*., 2015). Transcripts identified as being only partially complete are classified as 'fragmented', and transcripts not identified are classified as 'missing' (Simão *et al*., 2015).

BLAST analysis to identify open reading frames of interest

The most widely used tool for the analysis of nucleotide and protein sequences is a basic local alignment tool (BLAST) (Altschul *et al*., 1990). It is a computer algorithm that works using alignment matrices to find regions of similarity between sequences, as well as calculate the statistical significance of matches. BLAST is accessed most often through the National Centre for Biotechnology Innovation (NCBI) (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

BLAST creates alignment "seeds" by fractionating the input sequence into short "words" of a specified starting length, usually 3 residues and, once the alignment is "seeded", BLAST extends the words used and creates a set of acceptable changes in a sequence due to mutation, also dubbed "synonyms" (Altschul *et al*., 1990; Kerfeld and Scott, 2011). These words and synonyms are scored based on a substitution matrix (e.g. BLOcks of Amino Acid Substitution Matrix, BLOSUM) on how well they match the input sequence (Kerfeld and Scott, 2011). The sequences in the BLAST database that have the best-scored match to the input sequence are used to search for homologs. The so-called "best" alignments are recognised as those that exceed a neighbourhood score above the threshold (defined by the user) and above the cut-off value (also defined by the user) (Altschul *et al*., 1990; Lobo, 2008). However, some sequences may be recognised as a match by chance rather than because they are homologs. To circumvent this issue, the NCBI BLAST tool also calculates an 'e-value'. E-values indicate how likely a particular result is to be due to chance alone, the smaller the e-value, the less likely the result is to occur by chance and the more likely the match is to be a homolog or the input query. The usual cut-off value for a BLAST search is an e-value of 10^{-10}; however, the smaller and closer to 0 this value is, the better (Kerfeld and Scott, 2011).

BLAST can either compare nucleotide sequences with a database of nucleotide sequences (BLASTn), protein sequences with a database of protein sequences (BLASTp), nucleotide sequences (translated) with its database of protein sequences (BLASTx) or it can compare a protein sequence with a database of translated nucleotides (tBLASTn). To find specific transcripts in a transcriptome assembly, a local custom nucleotide database is created for each dataset which can then be used against a known reference protein sequence in a tBLASTn search for transcripts of interest.

The output file gives the coordinates in the database of the transcript that best matches the query; these coordinates can then be used to create a Browser Extensible Data (BED) format file. The BED-format file can be used to extract the required nucleic acid (as well as encoded protein) sequence from the transcriptome assembly using the open-source and LINUX-based BEDtools getfasta package that offers a suite of tools for analysis of genomic and transcriptome data (Quinlan and Hall, 2010).

To corroborate the identity of a protein, phylogenetic analysis and sequence alignments can be conducted with known reference sequences for the protein of interest, using bioinformatics software (e.g. MEGA, BioEdit, etc.). Direct sequencing of the transcript of interest can also be performed to confirm the sequence and corroborate any changes from the reference sequence.

Detection of proteins from tick tissues via ELISA

ELISA (enzyme-linked immunosorbent assay) is used for the detection of proteins with a protein-specific antibody. In this study, ELISA was used to test the suitability of several commercially available antibodies against PLCs for the detection of tick PLCs in gut tissue before proceeding to downstream application such as confocal imaging. As reviewed by Shehab (1983), a protein sample is used to coat the bottom of a specially designed, polystyrene multi-well plate and an enzyme-labelled antibody for the antigen of interest is added and incubated to allow binding. The plates are subsequently washed multiple times to remove unbound antibody and then a substrate for the enzyme is added, allowing the colourimetric detection of a measurable signal. It is essential for the success of any ELISA experiment that the antibody is of high affinity and specificity and that the protein of interest is present in high enough concentrations to be detected at a minimum of a three times signal to noise ratio.

Identification of the Antigen 1 region involved in recognition of Bm86

Yeast-two-hybrid

To determine the region of Antigen 1 that interacts with Bm86, a yeast-two-hybrid approach was used. The yeast two-hybrid platform provides a transcriptional assay for the detection of protein-protein interactions *in vivo* in yeast. Specifically, they can be used to identify novel protein interactions, confirm putative interactions and/or define interacting protein domains (Clontech, 2010, 1999).

Yeast has been used successfully to date for the production of recombinant Bm86 for vaccination of cattle and was found to outcompete the protective ability of Bm86 produced via prokaryotic platforms (Sikhosana 2017, unpublished results). This is most likely due to the protein folding of Bm86 in yeast which allows for the generation of antibodies to protective epitopes found in native Bm86.

The interaction between Bm86 and Antigen 1 was identified using a yeast two-hybrid system with Bm86 as bait and screening of a cDNA library (created by Prof. C Maritz-Olivier using a mixture of RNA from various tissues) by Kiper in 2013 (unpublished data, invention disclosure approved, patent pending).

The principle of the yeast two-hybrid system is shown in Figure 2.1. Briefly, a chimaera is created using a protein of interest fused to the DNA-binding domain of the Galactose responsive transcription factor (GAL4). This is referred to as the bait molecule, and in this study, Bm86 was used as the bait. A suitable cDNA library is fused to the GAL4 transcriptional activation domain (prey) to create a vast pool of chimaeras, referred to as the prey molecules. If the bait protein interacts with a prey protein, it will bring the DNA binding and activation domains of the Gal4 into proximity allowing binding to the GAL4 promoters upstream of several reporter genes (Clontech, 2010, 1999).

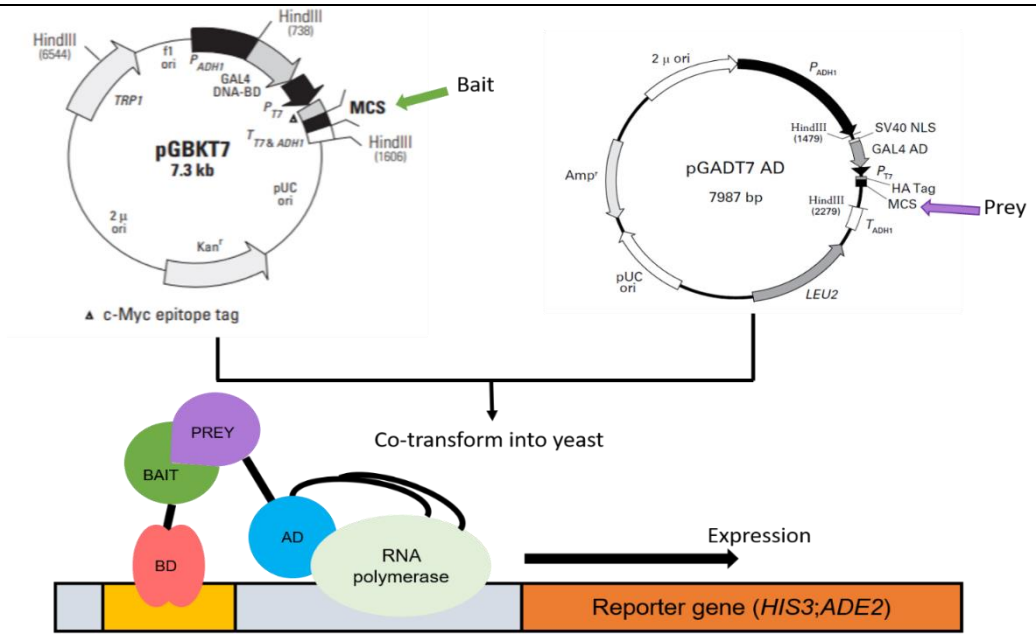


Figure 2.1: Principle of the yeast two-hybrid system. One protein of interest is expressed fused to the DNA binding domain (BD) of Gal4 (bait) while a library/another protein of interest is expressed fused to the activation domain (AD) of Gal4 (prey). If the bait and prey proteins interact, the DNA binding and activation domains are brought into proximity, resulting in transcriptional activation of reporter genes under control of the Gal4 promoter. Adapted from Clontech (2010)

The number of reporter genes activated depends on the system employed. In this study, the Clontech Matchmaker™ Gold Yeast-two-hybrid system is used (Clontech, 2010). In this system, the open reading frame of the bait, Bm86, was directionally cloned into the pGBKT7 plasmid, which already contains the coding sequence of the GAL4 DNA-binding domain. The prey plasmid pGADT7 was used for directional cloning of the various prey constructs, namely full-length Antigen 1 and truncations thereof. Plasmid maps are given in the appendix.

For the successful transformation of yeast cells, nutritional selection is conducted. In this study, recombinant colonies containing the *Bm86-GAL4 BD* and *TRP1* reporter genes were selected on media without Tryptophan. Similarly, prey constructs were selected using the Leucine reporter gene of the pGADT7 plasmid. All the latter is referred to as selection via single drop out media (SDO). If yeast contains both the pGBKT7 plasmid and pGADT7 plasmid (co-transformed), the selection is made using media lacking both Tryptophan and Leucine (termed double drop out media, DDO). To select positive protein-protein interacting clones, the expression of two additional reporter genes, for Histidine and Adenine, under control of the GAL4 promotor, is performed using media lacking Tryptophan, Leucine and Histidine (triple drop out

media, TDO). Final selection of strong interacting partners is conducted on TDO media that is further depleted of Adenine, so-called quadruple drop out media (QDO). The yeast strain used in this study is the Matchmaker® Gold Yeast Two-Hybrid yeast strain (Clontech Laboratories Incorporated.) and has a characteristic phenotype of turning a pink to red colour when growing in conditions where availability of Adenine is limited or depleted. This can include growth on media low in Adenine where the yeast cannot produce enough Adenine to overcome the media deficiency themselves, as well as when growing on QDO media that is completely depleted of Adenine. This phenotype is due to the insertion of the *ADE2* gene in the plasmids, allowing the yeast to grow on media low in Adenine but with an increasingly reddish colour as the Adenine, concentration produced decreases. Typically the red hue is observed when the protein-protein interactions are not strong enough to sufficiently allow binding to the GAL4 transcription factor to its promoter upstream of the *ADE2* gene, and as such is commonly used as a prediction of protein-protein interaction strength (Clontech, 2010).

MATERIALS AND METHODS

Transcriptome assembly

De novo *assembly of RNAseq data*

RNA sequencing data was assembled in collaboration with Dr N. Olivier at the Ion Torrent Sequencing Facility at the University of Pretoria. Following a quality check using FastQC (Andrews, 2010), reads were filtered, processed, and sequencing adaptors removed using Trimmomatic (Bolger *et al*., 2014). The filtered data were assembled using Trinity (Grabherr *et al*., 2013) and all transcripts shorter than 200 base pairs, as well as highly-homologous sequences were removed using CD-HIT (Fu *et al*., 2012; Li and Godzik, 2006). Open reading frames were predicted using TransDecoder (Haas *et al*., 2013) and the entire assembly was assessed for quality and completeness against the arthropod-specific BUSCO signature set (Simão *et al*., 2015).

Identification of open reading frames for Bm86, Antigen 1, G proteins and PLC pathway components

The relevant species sequences (Table 2.1) for the proteins of interest were obtained from Genbank (Benson *et al*., 2017). All assembled transcriptomes (corresponding to larvae, nymphs, salivary glands, gut and ovaries, respectively) were used to create a local custom nucleotide database for subsequent BLAST analyses (Altschul *et al*., 1997, 1990; Camacho *et al*., 2009). A custom command-line interface (CLI) was run for tBLASTn with the query protein sequence as input. An e-value cut-off of 1e-30 was applied for all results. The tBLASTn alignments with the maximum coverage, lowest e-value and longest read length were further considered.

Table 2.1: Query proteins searched for with BLAST against the *R. microplus* transcriptome assemblies. Only reviewed sequences were used where available from NCBI and Uniprot databases.

Protein of interest	Reference species used as BALST query	Accession number
Bm86	*R. microplus*	P20736
Antigen 1	*R. microplus*	No Reviewed Genbank sequence (Invention disclosure)
$G_{\alpha i}$	*D. melanogaster*	P20353
$G_{\alpha o}$	*D. melanogaster*	P16378
$G_{\alpha q}$	*D. melanogaster*	P23625
$G_{\alpha s}$	*D. melanogaster*	P20354
$G_{\alpha f}$	*D. melanogaster*	Q05337
$G_{\gamma\text{-}1}$	*D. melanogaster*	P38040
$G_{\gamma\text{-}e}$	*D. melanogaster*	Q9NFZ3
$G_{\beta\text{-}1}$	*D. melanogaster*	P26308
$G_{\beta\text{-}2}$	*D. melanogaster*	P29829
PLC_{ε}	*Caenorhabditis elegans*	G5EFI8
$PLC_{\delta\text{-}1}$	*H. sapiens*	P51178
$PLC_{\delta\text{-}3}$	*H. sapiens*	Q8N3E9
$PLC_{\delta\text{-}4}$	*H. sapiens*	Q9BRC7
$PLC_{\eta\text{-}1}$	*H. sapiens*	Q4KWH8
$PLC_{\eta\text{-}2}$	*Mus musculus*	A2AP18
$PLC_{\gamma\text{-}1}$	*H. sapiens*	P19174
$PLC_{\beta\text{-}1}$	*H. sapiens*	Q9NQ66
$PLC_{\beta\text{-}2}$	*M. musculus*	A3KGF7
$PLC_{\beta\text{-}3}$	*M. musculus*	P51432
$PLC_{\beta\text{-}4}$	*H. sapiens*	Q15147
$PLC_{\beta\text{-}21C}$	*D. melanogaster*	P25455
$PLC_{\beta\text{-}norpA}$	*D. melanogaster*	P13217
$PLC_{\beta\text{-}egl8}$	*C. elegans*	G5EBH0

The coordinates from the blast output file for the top, and the most representative hit was used, and a bed-format file created. The bed-file was used to extract the required nucleic acid sequence from the transcriptome assembly using the BEDtools getfasta package (Quinlan and Hall, 2010), taking into consideration the orientation of the blast-hit co-ordinates. The ExPasy translate web server tool (https://web.expasy.org/translate/) was used to predict the encoded protein sequence, which was checked to see if the correct full-length open reading frame sequence was predicted. The presence of the relevant full-length transcripts in the assemblies was evaluated using BLAST-analyses of the extracted sequence data. If full-length sequences were not extracted, the bed-file co-ordinates were modified to extend the extracted sequence in the 3' or 5' direction, taking note of the forward or reverse orientation of the blast hit to extend the sequence in the correct direction.

The most complete and correct sequences were compiled into separate nucleic acid and amino acid sequence fasta-files. This process was repeated for all transcriptome assemblies. The required amino acid fasta files from both GenBank and the transcriptome assemblies were subsequently concatenated into one file.

Sequence and phylogenetic analysis

New alignments were generated in MEGA-X V 10.1 (Kumar *et al*., 2018) with the MUSCLE algorithm (Edgar, 2004), on amino acid and nucleotide sequences. Neighbour-joining phylogenetic analyses were generated for G protein and PLC amino acid datasets using default settings and random bootstrap sampling using 10,000 replicates. Maximum Likelihood (ML) tree models were predicted for Bm86 datasets using jModelTest 2.1.10 (Darriba *et al*., 2015; Guindon and Gascuel, 2003) for nucleotide data and ProtTest 3.4.2 (Darriba *et al*., 2017; Guindon and Gascuel, 2003). Following the ML model prediction, ML trees were generated using PhyML 3.0 (Guindon *et al*., 2010) with 1 000 bootstrap replicates.

Images of alignments were generated in BioEdit software V 7.0.5.3 (Hall, 1999) and CLC Main Workbench V 8.0.1 (www.qiagenbioinformatics.com). Signal peptides were predicted using SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/), and GPI anchors were identified on predGPI using the GPI predictor tool (http://gpcr.biocomp.unibo.it/predgpi/pred.htm). Other domains were identified using SMART (http://smart.embl-heidelberg.de/), with a conserved domain search on the NCBI conserved domain site (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), on Prosite through the ExPaSy server (https://prosite.expasy.org/), and on the InterPro server of the European Bioinformatics Institute (https://www.ebi.ac.uk/interpro/).

Verification of the *de novo* assembled Bm86 sequence from larvae tissue of *R. microplus* using PCR and DNA sequencing

cDNA synthesis

cDNA was synthesised from three RNA samples used previously for RNA sequencing (as three biological repeats), using the Superscript IV cDNA kit (Invitrogen life technologies, USA) according to manufacturer's guidelines with minor adjustments. Briefly, Mixture A containing 3 µg RNA, 250 pmol poly-$dT_{19}VN$ primer and 500 pmol random nanomers in a volume of 10 µl and incubated for 10 minutes at 70°C to remove

RNA secondary structures, followed by immediate incubation in ice water for 10 minutes to allow primer annealing. Mixture B containing 6 µl of 5x first strand buffer, 100 mM DTT, 40 units RNasin RNase inhibitor (PROMEGA, USA), 10 mM dNTPs and 400 units of Superscript Reverse Transcriptase IV (Invitrogen life technologies, USA) was added to Mixture A with dddH_2O to a final volume of 30 µl and incubated for 20 hours at 42°C followed by enzyme inactivated at 80°C for 10 minutes before being held at 4°C. cDNA purification was subsequently done using the QI quick PCR purification kit (QIAGEN, USA) according to the manufacturer's instructions. Each sample was diluted with 5 volumes of the proprietary PB buffer (containing a high concentration of chaotropic salt), loaded onto a silica membrane and centrifuged at 16000 xg for 30 seconds. The flow-through was discarded, and 750 µl of PE buffer (proprietary composition) added to the column followed by centrifugation for 1 minute at 16000 xg. Again, the flow-through was discarded and columns dried via centrifugation for a minute at 16000 xg. To elute the cDNA, 30 µl DEPC-H_2O was added to the column and incubated at room temperature for 5 minutes before centrifugation at 16000 xg for 1 minute. The elution step was repeated once more, yielding a final volume of 60 µl. To determine whether there was salt and /or protein contamination, the cDNA purity and 'concentration' were determined using the Nanodrop-1000 system (Thermo Fisher Scientific Inc., USA).

PCR amplification of Bm86 from cDNA

Primers for regions of Bm86 of interest were designed by Bishop (2018, unpublished) using Oligo 7 primer design software (See Table 2.2). Each PCR reaction contained 30 ng cDNA with 10 pmol forward and reverse primers and 1X EconoTaq PLUS GREEN 2X Master Mix (Lucigen, USA) with dddH_2O to 25 µl total reaction volume. PCR reactions were mixed and centrifuged at max g for 30 seconds. PCR conditions were as follows: a hold of 94°C for 2 minutes, followed by 45 cycles of denaturation at 94°C for 30 seconds, annealing at 63°C for 30 seconds, and extension at 72°C for 1 minute per kb and ended with a final extension of 72°C for 10 minutes with a hold of 4°C. Correct sized bands were identified using electrophoresis in a 1% w/v agarose/TAE (40 mM Tris-acetate, 1 mM EDTA) stained with 0.5 µg/ml ethidium bromide (EtBr) in a MiniReadySub-Cell™ GT Cell (Bio-Rad) electrophoresis tank at 80 V with TAE running buffer and 1kb GeneRuler ladder (Thermo Fisher Scientific Inc., USA) with 5 µl of PCR product.

Table 2.2: Primers used for PCR amplification of Bm86 from cDNA.

Primer	Sequence (5' – 3')	Tm (°C)	Combined with:	Expected amplicon size (bp)
Bm86 Forward 1	ATGCGTGGCGTCGCTTTGTTCGTC	66	Reverse 1 (Primer set 2)	1947
			Reverse 2 (Primer set 3)	996
Bm86 Reverse 1	CGATGCTGCGGTGACTGAAGTAGC	66	Forward 1 (Primer set 2)	1947
			Forward 2 (Primer set 1)	1149
Bm86 Forward 2	GAAGACTGTCGTGTGCAGAAAGGA	63	Reverse 1 (Primer set 1)	1149
Bm86 Reverse 2	ATTGATGTTGACATTTGGGCCCGG	63	Forward 1 (Primer set 3)	996

DNA sequencing and analysis

A reaction containing 100 ng of PCR Product, 1 µl 5x sequencing buffer (BigDye Terminator v 1.1, v 3.1, Applied Biosciences), 2 µl 5x BigDye Terminator v 3.1 100RR (Applied Biosciences), 10 pmol gene-specific primer $ddd H_2O$ to a final volume of 10 µl was prepared. Amplification was conducted using an initial cycle at 80°C for 2 minutes and followed by denaturation at 96°C for 1 minute and 25 cycles of amplification (94°C for 30 seconds, 50°C for 5 seconds, 60°C for 75 seconds) with a final hold of 4°C. The amplified DNA was precipitated by the NaOAc/EtOH method where 10 µl water was added to the BigDye reaction (Mix B) and this transferred to a tube containing Mix A (3 µl 3 M Sodium Acetate pH 4.6, 62.5 µl absolute ethanol and $ddd H_2O$ to a final volume of 80 µl). The tubes were centrifuged for 45 minutes at 13 000 xg, the supernatant removed, and the DNA pellet washed twice with 250 µl ice-cold 70% ethanol. Pellets were dried *in vacuo* and submitted for sequencing at the ACGT sequencing facility at the University of Pretoria (RSA). Sequencing results were trimmed, assembled and then aligned on CLC Main Workbench V 8.0.1 software.

Preliminary, rapid evaluation of commercial antibodies against G proteins and PLCs in *R. microplus* gut tissue.

Protein extraction from R. microplus *gut tissue*

Due to the membrane location of all the proteins of interest, protein fractionation was conducted to enrich samples before ELISA. Membrane proteins were enriched for using the Alkaline Carbonate Extraction protocol from Schwab *et al.* (2000). Briefly, gut samples previously isolated and stored at -80°C in PBS containing 0.37 mg/ml protease inhibitor (cOmplete ULTRA, Roche, USA) were thawed on ice. Tissue was washed at least three times in 1 ml of 1X PBS and then in 1 ml of 100 mM NaCl with centrifugation at 1000 xg at 4°C for 2 minutes between wash steps. Tissue was then resuspended in 1 ml ice-cold 100 mM Na_2CO_3 (pH 11.5) containing 0.37 mg/ml protease inhibitor (cOmplete ULTRA, Roche, USA) and homogenized with at least 5 strokes of a homogenizer (T10 Basic Ultra-Turrax, IKA-Labortechnik) at speed setting 4 (14 500 rpm) with dispersal tool S10N-5G on ice. Following homogenization, samples were incubated on ice for 30 minutes and then centrifuged for 1 hour at 175 000 xg at 4°C. The cytosolic and peripheral protein portion (found in the supernatant) was transferred to a fresh 1.5 ml tube while the membrane pellet was resuspended in 1 ml 100 mM Na_2CO_3 (pH 11.5). Protein concentrations were determined with the Pierce™ BCA Protein Assay Kit (Thermo Scientific, USA).

ELISA

Protein fractions were immediately used to coat an ELISA plate. Briefly, 30 µg for the cytosolic fraction and 15 µg for the membrane fraction (25 µl) protein sample was loaded per ELISA well and the plate dried in a laminar flow under a 60W bulb. Once dried, samples were blocked overnight at 4°C with 200 µl blocking buffer (1x TBS; 20 mM Tris-HCl, 150 mM NaCl, pH 7.4-7.6 containing 1% casein w/v) per well. Following blocking, the plate was washed at least three times using blocking buffer. Primary antibodies were diluted in blocking buffer at a ratio of 1:300 for test samples and 1:1000 for controls (see Table 2.3 for a list of antibodies used) and incubated at 37°C for 1 hour followed by three washing steps. Next, plates were incubated with a horse-radish-peroxidase (HRP) conjugated secondary antibody diluted 1: 10 000 with blocking buffer at 37°C for 1 hour and washed as before. To allow colour development, wells were incubated with substrate buffer (0.05 M Citric acid with 0.25 M Na_2HPO_4, pH 5) and 0.04% (w/v) Ortho-Phenylenediamine (OPD)/ substrate buffer and 0.04% (v/v)

H_2O_2 / substrate buffer/OPD for 10-20 minutes before the addition of 2 N sulfuric acid to stop the reaction. Absorbance was read at 490 nm on a SpectraMax Paradigm multi-mode microplate reader (Molecular Devices Corp., USA). Protein was omitted in all negative control reactions. Antisera against Bm86 was used as a positive control for detection of membrane fractions, while anti-retinol dehydrogenase was used as a marker for the detection of the cytosolic fraction (Table 2.3).

Table 2.3: Antibodies used in ELISA assay.

Name of 1° Antibody	**Animal Protein target**	**Animal Raised in**	**Corresponding 2° Antibody (all produced by Abcam, UK)**
PLCγ1 (D9H10) XP® Rabbit mAb 5690[1]	Human PLCγ1	Rabbit	Goat Anti-rabbit IgG H&L (HRP)
PLCγ2 Antibody 3872[1]	Human PLCγ2	Rabbit	Goat Anti-rabbit IgG H&L (HRP)
PLCβ1 ANTIBODY (D-8): SC-5291[2]	Rat PLCβ1	Mouse	Goat Anti-Mouse IgG H&L (HRP)
PLCδ1 ANTIBODY (D-7): SC-393464[2]	Human PLCδ1	Mouse	Goat Anti-Mouse IgG H&L (HRP)
Gγ 1 ANTIBODY (1F8): SC-517057[2]	Human Gγ1	Mouse	Goat Anti-Mouse IgG H&L (HRP)
Gγ 2/3/4/7 (C-5): sc-166419[2]	Human Gγ2	Mouse	Goat Anti-Mouse IgG H&L (HRP)
Gα i-1 (B-11): sc-515658[2]	Rat Gαi-1	Mouse	Goat Anti-Mouse IgG H&L (HRP)
Gα q/11/14 (G-7): sc-365906[2]	Human Gα11	Mouse	Goat Anti-Mouse IgG H&L (HRP)
T3526 - Anti-Tubulin antibody produced in rabbit[3]	Human Tubulin α-1B	Rabbit	Goat Anti-rabbit IgG H&L (HRP)
Bm86[4]	*R. microplus* Bm86	Chicken	Goat Anti-chicken IgY H&L (HRP)
RDH[4]	*R. microplus* Retinol Dehydrogenase	Chicken	Goat Anti-chicken IgY H&L (HRP)

Antibodies were produced by Cell Signalling Technology®, USA ([1]), Santa Cruz Biotechnology, USA ([2]) and Sigma-Aldrich, USA ([3]) and the Ticks research group at the University of Pretoria ([4])

Yeast-Two-Hybrid

Amplification and directional cloning of Bm86 and Antigen 1 into yeast two-hybrid plasmids

Yeast codon-optimised synthetic gene constructs for both Bm86 and Antigen 1, as well as primers for amplification and directional cloning were already available (Table 2.4). These constructs and primers were utilised for the amplification of Bm86, full-length Antigen 1, as well as truncations of Antigen 1 for subsequent directional cloning into the pGBKT7 and pGADT7 plasmids, respectively. When designing the Antigen 1 truncations to be used, the encoded protein sequence was utilised for the conserved domain (CCD) BLAST search using the NCBI platform. Two BPTI/Kunitz-type inhibitory fold domains were identified, and primers were designed for the synthesis of two truncations, containing only the first Kunitz domain (396 bp fragment) and a construct without any Kunitz domains that contains only the first 73 amino acids from the N-terminus of the protein (219 bp fragment).

Table 2.4: Primer sequences used in the yeast two-hybrid study. Restriction enzyme cut sites are underlined. 'F' indicates a forward primer, while 'R' indicates a reverse primer. 'SP' indicates where a primer is used alone for sequencing purposes.

Name	Sequence 5'-3'	Tm (°C)	Expected amplicon (bp)
BM86_F_NDEI	GGAATTCCATATGGAGTCTTCCGTGTGTTCTGA	63.2	1893
BM86_R_XMAI	TCCCCCCGGGCGCAGGGAGGCAGCGGTCACAGA	75.6	
KUBP_F1_NDEI	GGGAATTCCATATGACCCCTGGCGGCGTCTG	68.3	Dependent on Reverse primer used
KUBP R1_BAMHI	CGGGATCCGTTAGTAGGCTTTCCAGACGAG	65.7	219
KUBP R2_BAMHI	CGGGATCCGTACATGTGGTAGTTGAGG	62.8	396
KUBP R3_BAMHI	CGGGATCCGCTCAGCAATGCAGATGGCTCG	69	558
T7_SP1 (5')	TTAATACGACTCACTATAGGGC	49	Dependent on insert
3'AD_SP2	AGATGGTGCACGATGCACAG	56	
3'BD_SP2	ATCATAAATCATAAGAAATTCGCC	58	

PCR was performed using 10 ng plasmid containing the full-length Antigen 1 transcript as a template, 10 pmol of the relevant forward and reverse primers to produce the three constructs of Antigen 1 and EconoTaq PLUS GREEN 2X Master Mix (Lucigen, USA) to a final volume of 25 µl. PCR was carried out with an initial denaturation hold at 94°C for 2 minutes, followed by 45 cycles of amplification (denaturation at 94°C for 30 seconds, annealing at 63 °C for 30 seconds, and extension at 72°C for 1 minute per kb) and a final extension of 72°C for 10 minutes. The same process was followed to amplify Bm86 from pAS2-1 using the Bm86 primers (BM86_F_NDEIandBM86_R_XMAI). Correct sized bands were identified using electrophoresis with a 1% w/v agarose/TAE gel stained with 0.5 µg/ml EtBr in TAE buffer. Two tubes of PCR product for each truncation was combined and cleaned up using the Wizard SV Gel and PCR clean-up kit™ (Promega, Germany) according to the manufacturer's guidelines and eluted in 30 µl $dddH_2O$. Concentrations and purity of all products were assessed spectrophotometrically on the Nanodrop-1000 (Thermo Fisher Scientific, USA).

Restriction enzyme digestion of PCR products and plasmids

PCR products and plasmids were digested with two restriction enzymes to afford directional cloning. Briefly, 1 µg of either plasmid or purified PCR product was combined with 1X of the appropriate restriction Buffer, 20 U of each restriction enzyme (i.e. NdeI and BamHI for Antigen 1 truncations and pGADT7 / NdeI and XmaI for Bm86 and pGBKT7 restriction enzymes) and water to a final volume of 50 µl. Reactions were mixed by pipetting, centrifuged briefly at 16 000 xg and incubated at 37°C for 5 hours. The reaction was inactivated at 65 °C for 20 minutes, the digested products analysed for complete digestion using 1 % agarose gel electrophoresis and then purified using the Qiagen PCR clean up kit ™ according to manufacturer's guidelines. Purified products were eluted in 30 µl water, and DNA concentration quantified spectrophotometrically on the Nanodrop-1000 (Thermo Fisher Scientific, USA). Samples were stored at 4°C until use.

Dephosphorylation of plasmids

Following restriction enzyme digestion, the purified plasmids were dephosphorylated to remove 5' phosphates and prevent self-ligation of the restricted plasmids. Briefly, 500 ng linearized plasmid DNA was combined with 1X thermosensitive alkaline

phosphatase buffer (Thermo Scientific, USA), 10 U thermosensitive alkaline phosphatase enzyme and water in a final volume of 20 µl. The reaction was incubated at 37 °C for 15 minutes, and the enzyme was inactivated at 75 °C for 10 minutes.

Ligation

Digested PCR products were combined in various ratios (3:1, 5:1 and 10:1) with 60 ng of the digested plasmid in a final concentration of 1X T4 DNA ligase buffer (New England Biolabs, USA), 0.2 pmol ATP and 20 U DNA ligase with water to a final volume of 20 µl. The components were mixed by pipetting and briefly centrifuged before incubation at 16°C overnight followed by enzyme inactivation at 65°C for 10 minutes. Ligation products were used directly for the transformation of competent bacterial cells via heat-shock transformation or electroporation (see below).

Preparation and transformation of electrocompetent E. coli

Electrocompetent cells were prepared for plasmid uptake as follows: A tip inoculated with *DH5α* stock was deposited in 20 ml Luria-Bertani broth (i.e. LB broth) and grown overnight at 30°C with shaking at 250 rpm. Two 1 litre flasks containing 250 ml LB broth each was inoculated with 5 ml from the overnight culture and grown at 37°C with shaking at 250 rpm. The optical density at 600 nm (OD_{600}) was determined spectrophotometrically at hourly intervals until an OD_{600} of 0.4-0.5 was reached. Cultures were then divided equally into ten 50 ml centrifuge bottles, incubated on ice for 20 minutes and centrifuged at 10 000 xg for 20 minutes at 4°C. The cell pellets were subsequently washed three times in 50 ml ice-cold $dddH_2O$, resuspension in 10 ml 10% (v/v) glycerol, incubated on ice for 60 minutes and pelleted via centrifugation at 10 000 xg for 10 minutes at 4°C. All supernatants were discarded, and all pellets combined into 1 ml 10% (v/v) glycerol and stored in 90 µl aliquots at -80°C.

For electroporation, 20 µg yeast tRNA was added to each ligation reaction, and DNA was precipitated using NaOAc/EtOH (see previous) and dissolved in a final volume of 10 µl water. For electroporation, electrocompetent cell stocks (90 µl) were thawed on ice, and 10 µl precipitated ligation added, mixed gently and loaded into a 0.1 cm gap Micropulser electroporation cuvette (Biorad, USA). Electroporation was performed at 2000 V for 4 milliseconds with the Electroporator 2510 (Eppendorf, USA). Following electroporation, 100 µl pre-warmed LB-glucose (LB broth containing 20 mM glucose)

was added directly to the cells in the cuvette and the contents transferred to 900 µl pre-warmed LB-glucose in a 2 ml Eppendorf tube. Transformed cells were then incubated at 37 °C for 1 hour with shaking at 250 rpm before plating cells onto LB/Amp plates (2% (w/v) agar/LB-broth with 100 µg/ml Ampicillin) and incubated overnight at 37°C.

Preparation and heat-shock transformation of chemically competent cells

Chemically competent cells were prepared for plasmid uptake as follows: A tip inoculated with *DH5α* stock was deposited in 5 ml LB Broth and grown overnight at 37°C with shaking at 250 rpm. LB broth (200 ml) was then inoculated with 1 ml of the overnight culture and grown until an OD_{600} of 0.3-0.4 was reached. The culture was then divided into 30 ml aliquots in pre-chilled 50 ml falcon tubes and centrifuged at 2000 rpm at 4°C for 10 minutes and the supernatant decanted. Cell pellets were resuspended in 15 ml ice-cold 50 mM $CaCl_2$ and centrifuged again for 5 minutes. The collected cells were then resuspended in ice-cold 50 mM $CaCl_2$:15% Glycerol (v/v) and incubated on ice for 1 hour before being dispensed into 100 µl aliquots for storage at -80°C.

For transformation, 100 µl chemically competent cells were thawed on ice, and 250 ng of pGBKT7 control or 5.5 µl ligation product test plasmid was added to the cells and gently swirled with a tip. The cell-plasmid mix was incubated on ice for 30 minutes, heat-shocked at 42°C for 90 seconds and then immediately incubated on ice for another 2 minutes. A 900 µl volume of LB-glucose was added to the cells, and the culture was incubated at 37°C for 1 hour with shaking. Finally, cells were collected via centrifugation and resuspended in 100 µl LB-glucose before plating the cells on LB/Kan plates (50 mg/ml Kanamycin in 2% (w/v) agar plates). All plates were incubated upside-down at 37°C overnight.

Colony PCR

Positive clones were identified via colony PCR using 10 pmol vector-specific forward (pGADT7_F or T7_SP1) and gene-specific reverse (KUBP R1_BAMHI PGADT7; KUBP R2_BAMHI PGADT7 and KUBP R3_BAMHI PGADT7 or BM86_R_XMAI) primers, respectively and 25 µl CloneID 1X Colony PCR Master Mix (Lucigen, USA). PCR conditions started with an initial denaturation at 98°C for 5 minutes followed by

45 cycles of amplification (denaturation at 98°C for 30 seconds, annealing at 62°C for 30 seconds and extension at 72°C for 2 minutes) and a final extension of 72°C for 10 minutes. Amplified products were analysed via agarose gel electrophoresis and recombinant colonies with the expected insert sizes were regrown at 37°C overnight in 50 ml falcon tubes containing 5 ml LB with Ampicillin or Kanamycin (as described previously).

Plasmid isolation from recombinant E. coli *clones*

Plasmids were isolated from overnight cultures with the PureYield™ Plasmid miniprep system (Promega, USA) according to manufacturer's guidelines. Briefly, 1.5 ml overnight culture was pelleted at max speed for 30 seconds, the supernatant discarded, and another 1.5 ml culture added and pelleted again. Cell pellets were resuspended in 600 µl water, and then 100 µl of 7x lysis buffer was added, the tubes inverted 6 times and finally incubated at room temperature for 3 minutes. Ice-cold neutralisation buffer was then added to the cell lysate, inverted 5 times and centrifuged at max speed for 4 minutes. The supernatant was transferred to a column in a collection tube and centrifuged for 15 seconds. Flow-through was discarded, and 200 µl Endotoxin Removal Wash added to the membrane and centrifuged for another 15 seconds. The flow-through was discarded, and 400 µl Column Wash Solution was added to the membrane, followed by centrifugation for 30 seconds. Columns were transferred to a clean 2 ml Eppendorf tube, and the plasmid DNA was eluted with 30 µl water. Plasmid concentrations were assessed spectrophotometrically on the Nanodrop-1000 (Thermo Fisher Scientific, USA) and visualised with electrophoresis using a 1% w/v agarose/TAE gel stained with 0.5 µg/ml EtBr in TAE buffer.

Preparation and sequential transformation of competent yeast cells

Glycerol cell stocks of Y2H-Gold yeast strains (provided with the Clontech Matchmaker® Gold Y2H system, Takara bio, USA) were steaked on YPDA plates (YPDA, 20 g/l peptone, 10 g/l yeast extract, 20 g/l agar, 20g/l dextrose, 0,03g/l adenine hemisulphate) and incubated upside down at 30°C for 3 days. For yeast transformation, 10 ml of YPDA (lacking agar) was inoculated with a single Y2H Gold yeast colony and grown overnight at 30°C with vigorous shaking. The overnight culture was then diluted to an OD600 of 0.2 and grown to an OD600 of 0.4-0.6. Cells were harvested at 12 000 xg for 30 seconds. Freshly grown log-phase cell pellets were resuspended in 1 ml 100 mM LiAc and incubated at 30°C for 5 minutes without shaking to make the cell membranes porous, pelleted at 12 350 xg for 30 seconds

and the supernatant discarded. The following was added to the cell pellet in this order: 240 µl PEG 4000, 36 µl 1 M LiAc, 25 µl 2 mg/ml heat-denatured Salmon Sperm DNA and 500 ng plasmid DNA made up to 50 µl final volume in dddH2O. The mixture was then vortexed for 1 minute and heat-shocked at 42°C for 25 minutes. Cells were again pelleted at 12 000 xg for 1 minute, the supernatant discarded, and the pellet resuspended in 200 µl water. The transformed pGBKT7-53, pGBKT7-LAM and Bm86 in pGBKT7 and pAS2-1 yeast cells were plated on SD/-Trp single drop out (SDO) solution, 6.7g/l Yeast Nitrogen Base without Amino Acids, 20g/l Agar, 100ml 10X Amino Acid Drop Out Solution (Clontech, Takara Bio, USA)) with glass beads, allowed to dry in a laminar flow and then incubated upside down at 30°C for 1-3 days. Untransformed yeast cells were used as a negative control on SD/-Trp plates. Colonies from the first SD/-Trp plates were replica plated on master SD/-Trp plates and SD/-Leu plates as a negative control. One colony was then resuspended in 10 ml SD/-Trp per intended transformation as with the bait transformation. Yeast with Bm86 in pGBKT7 and pAS2-1 were transformed with each of the plasmids for Antigen 1 truncations, as well as the full-length Antigen 1 plasmid. Yeast with pGBKT7-53 and pGBKT7-LAM were transformed with pGADT7-T. Cells were then plated on DDO medium (SD -Leu/-Trp) and incubated upside down 1-3 days at 30°C.

Screening of two-hybrid colonies mediating protein-protein interactions

To test for autoactivation, all control and test constructs were plated onto single and double drop-out plates containing various concentrations of 3-Amino-1,2,4-triazole (3-AT) and grown for 1 week at 30°C.

To screen for protein-protein interactions, co-transformed clones were plated onto SD/-Trp (SDO), then SD/-Trp/-Leu (DDO), SD/-Trp/-Leu/-His (TDO) and finally SD/-Trp/-Leu/-His/-Ade (QDO). Briefly, from DDO plates, the co-transformed clones were replica plated onto TDO with a sterile wooden stick and allowed to grow for 1-2 days. To eliminate false positives and background growth, positive cells from the first TDO were replica plated on fresh TDO containing 0, 2.5 and 5 mM 3-AT, and grown for 1-4 days. The second TDO streaks without 3-AT that grew were replica plated onto QDO following overnight growth. Quadruple drop out colonies were grown for 3-5 days before being replica plated again onto fresh QDO plates. This was done, to eliminate false positives which may have resulted in transference of residual TDO media during the initial replica plating from TDO. All plate incubations were done upside down at 30°C. In all plating, one replica plate was stored at 4°C as back up.

REFERENCES

Altschul, S., Madden, T., Schäffer, A., Zhang, J., Zhang, Z., Miller, W., Lipman, D., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andrews, S., 2010. FastQC: A quality control tool for high throughput sequence data. Babraham Bioinforma. https://doi.org/citeulike-article-id:11583827

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2017. GenBank. Nucleic Acids Res. 45, D37–D42. https://doi.org/10.1093/nar/gkw1070

Bolger, A. M., Lohse, M., & Usadel, B., 2014. Trimmomatic Manual: V0.32. https://doi.org/10.1093/bioinformatics/btu170

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Clontech, 2010. Matchmaker ™ Gold Yeast Two-Hybrid System User Manual.

Clontech, 1999. MATCHMAKER GAL4 Two-Hybrid System 3 & Libraries User Manual.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. Genome Biol. 17, 13. https://doi.org/10.1186/s13059-016-0881-8

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2017. Europe PMC Funders Group ProtTest 3 : fast selection of best-fit models of protein evolution. Bioinformatics 27, 1164–1165. https://doi.org/10.1093/bioinformatics/btr088.ProtTest

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2015. jModelTest 2: more models, new heuristics and high-performance computing Europe PMC Funders Group. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109.jModelTest

de Castro, M.H., de Klerk, D., Pienaar, R., Latif, A.A., Rees, D.J.G., Mans, B.J., 2016. *De novo* assembly and annotation of the salivary gland transcriptome of Rhipicephalus appendiculatus male and female ticks during blood feeding. Ticks Tick. Borne. Dis. 7, 536–548. https://doi.org/10.1016/j.ttbdis.2016.01.014

Deluca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W., Getz, G., 2012. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 28, 1530–1532. https://doi.org/10.1093/bioinformatics/bts196

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma,

F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883.Trinity

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883

Guindon, S., Dufayard, J., Lefort, V., 2010. Guindon *et al*. - 2010 - New algorithms and methods to estimate Maximum-Likelihood phylogenies assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Guindon, S., Gascuel, O., 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst. Biol. 52, 696–704. https://doi.org/10.1080/10635150390235520

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084

Haidula, T., 2016. RNA Sequencing and Analysis. The Namibian 2015, 5. https://doi.org/10.1101/pdb.top084970.RNA

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Kerfeld, C.A., Scott, K.M., 2011. Using BLAST to Teach “E-value-tionary” Concepts. PLoS Biol. 9, e1001014. https://doi.org/10.1371/journal.pbio.1001014

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096

Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158

Lobo, I., 2008. Basic local alignment search tool. Nat. Educ. https://doi.org/10.1016/S0022-2836(05)80360-2

Okonechnikov, K., Conesa, A., García-Alcalde, F., 2016. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics 32, 292–294. https://doi.org/10.1093/bioinformatics/btv566

Parra, G., Bradnam, K., Korf, I., 2007. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23, 1061–1067. https://doi.org/10.1093/bioinformatics/btm071

Parra, G., Bradnam, K., Ning, Z., Keane, T., Korf, I., 2009. Assessing the gene space in draft genomes. Nucleic Acids Res. 37, 289–297. https://doi.org/10.1093/nar/gkn916

Quinlan, A.R., Hall, I.M., 2010. BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ross, M.G., Russ, C., Costello, M., Hollinger, A., Lennon, N.J., Hegarty, R., Nusbaum, C., Jaffe, D.B., 2013. Characterising and measuring bias in sequence data. Genome Biol. 14, R51. https://doi.org/10.1186/gb-2011-12-2-r18

Schulz, M.H., Zerbino, D.R., Vingron, M., Birney, E., 2012. Oases: Robust *de novo* RNA-seq assembly

across the dynamic range of expression levels. Bioinformatics 28, 1086–1092. https://doi.org/10.1093/bioinformatics/bts094

Schwab, R.B., Okamoto, T., Scherer, P.E., Lisanti, M.P., 2000. Unit 5.4: Analysis of the association of proteins with membranes, in: Current Protocols in Cell Biology. John Wiley & Sons, Inc., pp. 5.4.1-5.4.17.

Shehab, Z.M., 1983. The enzyme-linked immunosorbent assay (ELISA). Infect. Dis. Newsl. 2, 91–92. https://doi.org/10.1016/S0278-2316(83)80010-3

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

CHAPTER 3: RESULTS

De novo transcriptome assembly for the South African strain of *R. microplus*

RNA sequencing data from larvae, nymphs and three adult tissues from a South African strain of *R. microplus* (provided by ClinVet, Bloemfontein, SA) was already available in the Tick and Tick-borne diseases group at the University of Pretoria. The transcriptome assembly was performed in collaboration with Dr N. Olivier at the Ion Torrent Sequencing facility at the University of Pretoria. A summary of the assembly statistics per tissue and life stage is given in Table 3.1.

Table 3.1: Summary statistics of the *de novo* transcriptome assembly of a South African *R. microplus* strain.

Assembly data set	Total Contig size (nt)	Number of contigs	Longest contig (bp)	Shortest contig (bp)	Mean contig length (bp)	Median contig length (bp)
Gut	129,837,586	141 772	23 692	201	916	1 731
Salivary Gland	186,509,941	216 791	18 769	201	860	422
Ovary	180,540,522	198 865	17 329	201	908	436
Larvae	237,411,162	300 261	22 772	201	791	407
Nymphs	228,085,412	278 539	18 548	201	819	414
Total	485,380,458	735 719	24 811	201	660	373

From Table 3.1 it is evident that, of the adult tissues, the gut has the smallest transcriptome assembly with a size of 129 837 586 bp, although it does contain the longest contiguous sequence (23 692 bp) — the combined size of the adult tissue transcriptome assemblies' amount to 496 888 049 bp. The larval transcriptome assembly is 47% the size of the combined adult total, while the nymph transcriptome assembly is 46% compared to the adult total. During the assembly process, all contigs ≤200 bp were removed. The mean contig lengths in all transcriptomes are approximately equivalent while the median contig length in the gut is more than 50% longer than the median contig length in any other transcriptome assembled (see comparisons to other *Ixodidae* in Chapter 4). The GC content was calculated as 46.8%, and the A/T and G/C content for each transcriptome equate to 100% (results not shown) indicating that no anomalous nucleotides were read and included into the assembly.

The individual reads were combined into one transcriptome assembly, referred to as the total transcriptome assembly. In this assembly identical sequence reads were collapsed into one representative sequence, and as such it represents only the most abundant sequences overall. The total transcriptome is 98% the size of the combined adult tissue assemblies. As the comparison with the combined adult tissues, the larva and nymph assemblies are 49% and 47% the size of the total transcriptome assembly respectively.

The completeness of the transcriptome assemblies for the larva, nymph, adult gut, ovary and salivary glands, as well as the total transcriptome, were assessed using the BUSCO dataset that is specific for arthropods. A comparison of the *R. microplus* dataset with that of *R. decoloratus*, *I. scapularis* and *I. ricinus* is given in Table 3.2.

Table 3.2: Completeness assessment of *R. microplus de novo* transcriptome assembly using the BUSCO arthropod dataset in comparison to other *Ixodid* transcriptome assembly BUSCO completeness assessments.

Assembly	South African *R. microplus*						Other Ixodidae transcriptome assemblies		
Species	*R. microplus*						*R. decoloratus**	*I. scapularis*[#]	*I. ricinus*[#]
Life stage	Larva	Nymph	Adult			Total	Total	N/A	N/A
Organ	N/A	N/A	Gut	Ovary	Salivary Gland				
% complete BUSCO	95%	94%	93.2%	93.7%	94.3%	94%	95%	85.2%	94.8%
Single copy genes	38%	35.5%	42%	33.3	37.5%	25.7	71.8%	83.5%	35.6%
Duplicate copy genes	57%	58.6%	51.1%	60.4%	56.8%	68.3%	23.2%	1.7%	59.2%
Fragmented genes	4.4%	4.9%	4.2%	3.9%	4.1%	5.3%	1.6%	9.9%	1.6%
Missing genes	0.56%	1%	2.6%	2.3%	1.6%	0.75%	3.6%	4.9%	3.6%

**BUSCO analysis for the *R. decoloratus* transcriptome. (Baron *et al*., 2018)

[#]BUSCO analysis for *I. scapularis* and *I. ricinus* transcriptome. Data received from Prof. Ben Mans (Agricultural Research Council)

The *R. microplus* adult gut, ovary and gland tissues, as well as the larval and nymph life stage transcriptomes were successfully assembled with a mean BUSCO completeness score of over 94%, with only 1.5% BUSCO genes missing from the assemblies. Duplicated genes make up a mean of 59% across all transcriptome sets. Transcripts that are classified as missing or fragmented are often artefacts of the

assembly process, indicating that the assembly and/or sequencing read pool did not include the total representative transcript and/or gene sequences in the RNA samples.

The respective RNA sequencing reads were subsequently mapped back to the respective transcriptome assemblies, and QC performed using the Samtools Flagstat tool to confirm that all the reads do map back to the assemblies. Results show 100% of the reads mapping back to their respective transcriptome (results not shown).

All the data confirm that the RNA sequence assemblies for the South African *R. microplus* transcriptome assembly sets are representative of the core gene content for arthropods and are suitable for further annotation and analysis.

Identification of sequences for Bm86, Antigen 1, G-proteins and PLCs

The transcriptome assemblies were used to generate local BLAST databases for similarity analyses. Sequences for Bm86 (accession number P20736) and Antigen 1 (In-house sequence), as well as various PLC and G proteins, were used in a tBLASTn analysis against the assembled *R. microplus* transcriptomes to identify putative homologous sequences.

Global Bm86 sequence variation

The most representative Bm86 sequences from each life stage and tissue transcriptome assembly plus the sequences from GenBank were used to build two Maximum Likelihood (ML) phylogenetic trees (Chapter 2), i.e. one based on the amino acid sequence alignment (Figure 3.1) and the other for the nucleotide sequence alignment (Figure 3.2). The bootstrap consensus tree was inferred from 1000 replicates. The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test is shown next to the nodes.

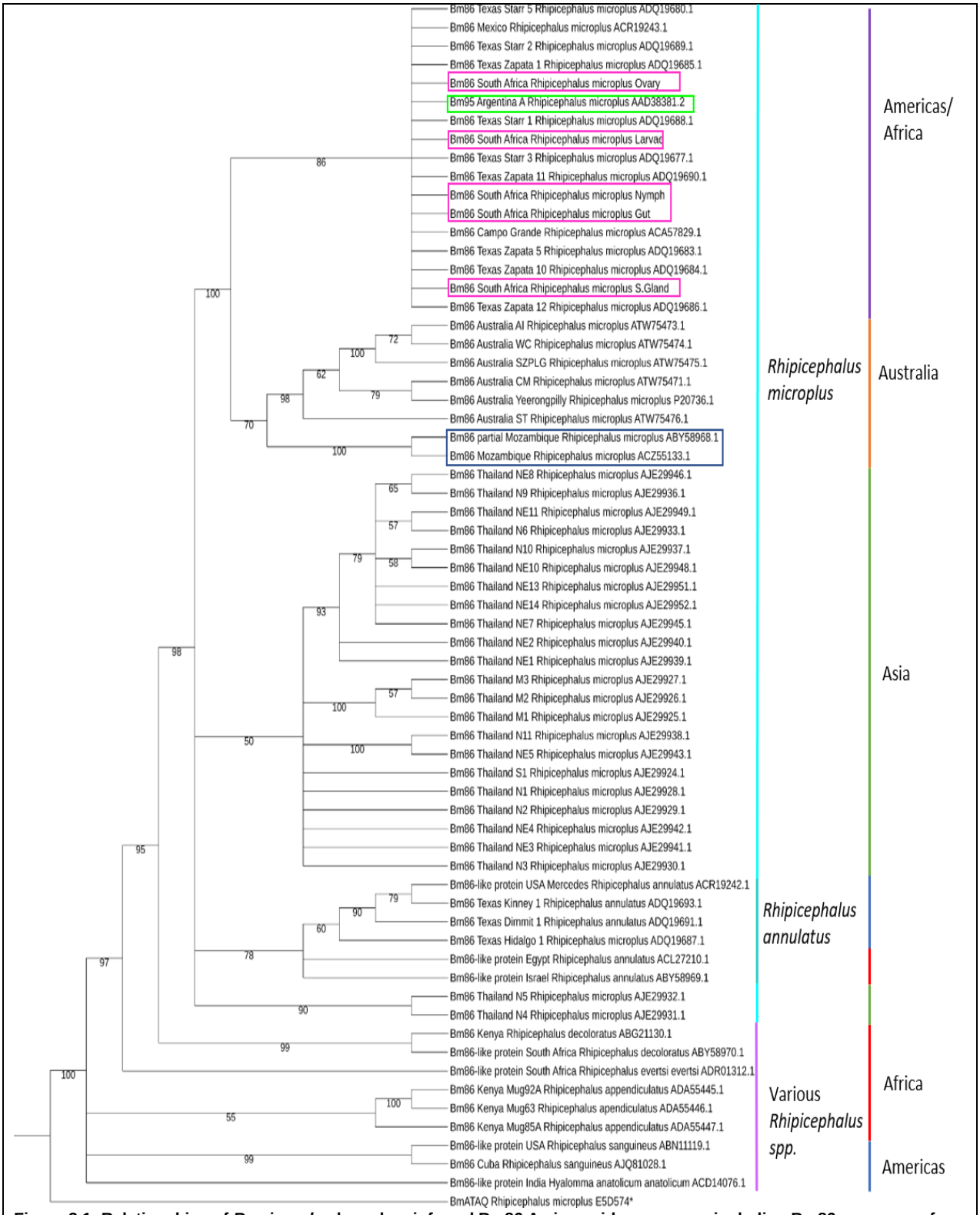


Figure 3.1: Relationships of *R. microplus* based on inferred Bm86 Amino acid sequences, including Bm86 sequences from the transcriptome assemblies of South African laboratory strain of *R. microplus*. While sequences from each species tend to group together and within their region, the translated sequences extracted from the RNA sequence assemblies, indicated in the pink boxes, groups with American sequences from the Texas outbreak strains and not the Mozambique sequence, indicated in the blue box which groups with the Australian sequences. The green box indicates Bm95. The Maximum Likelihood tree was inferred using the FLU+G+F model with a Gamma shape parameter = 1.336 and with the BmATAQ sequence as an outgroup (indicated by asterisk "*"). Nodes supported by less than 50% bootstrap (1 000 replicates) were collapsed.

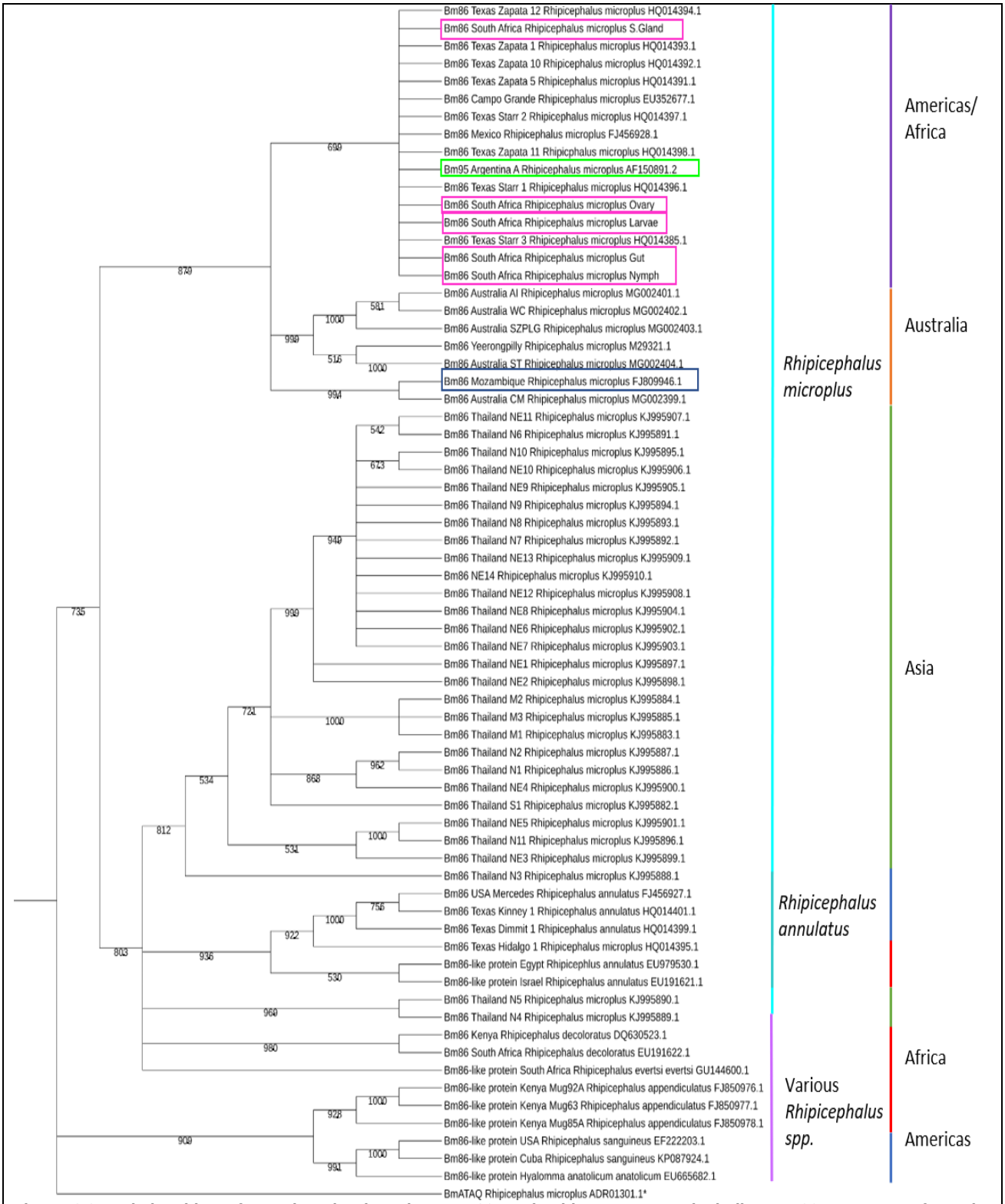


Figure 3.2: Relationships of *R. microplus* based on Bm86 nucleotide sequences, including Bm86 sequences from the transcriptome of a South African laboratory strain of *R. microplus*. While sequences from each species tend to group together and within their region, the sequences extracted from the RNA sequence assembly, indicated in the pink boxes, groups with American sequences from the Texas outbreak strains and not the Mozambique sequence, indicated in the blue box, which groups with the Australian sequences. The green box indicates Bm95. The Maximum Likelihood tree was inferred using the GTR + G model with a Gamma shape parameter = 1.055 and with the BmATAQ protein as an outgroup (indicated by an asterisk, "*"). Nodes supported by less than 50% bootstrap (1 000 replicates) are collapsed.

The ML trees indicate that the amino acid and nucleotide trees are congruent in that they depict the formation of the same clades. Sequences from each tick species tended to group together and within their species, as well as within their region of origin in both trees. The sequences extracted from the respective South African transcriptome assemblies (Figure 3.1 and 3.2, magenta boxes), all group with the American *R. microplus* Bm86 sequences from the Texas outbreak strains, as well as the Argentinean A sequence of Bm95 (Figure 3.1 and 3.2, green box). The Bm95 Argentinean A sequence is the only confirmed Bm95 sequence available in non-redundant databases (García-García *et al*., 2000). The Mozambican Bm86 groups with the Australian sequences in both ML trees. The South African and American sequences group with the Australian *R. microplus* sequences in both ML trees.

The phylogenetic trees also showed the Bm86 sequences from different tissues, and life stages are distinct, and thus alignments of these sequences were analysed.

Bm86 sequence variation within the South African Rhipicephalus microplus *strain.*

An alignment of the extracted sequences for Bm86, and the first reported and reviewed sequence for Bm86 from Australia (Yeerongpilly strain, accession number P20736) and Bm95 (Argentinean-A strain) is shown in Figure 3.3. Sequences from the total transcriptome were not included in the alignment or tree analysis as they are already represented in the individual transcriptomes. The percent identity (ID) for the alignment is given in a Matrix in Figure 3.4, where each sequence is compared to each other.

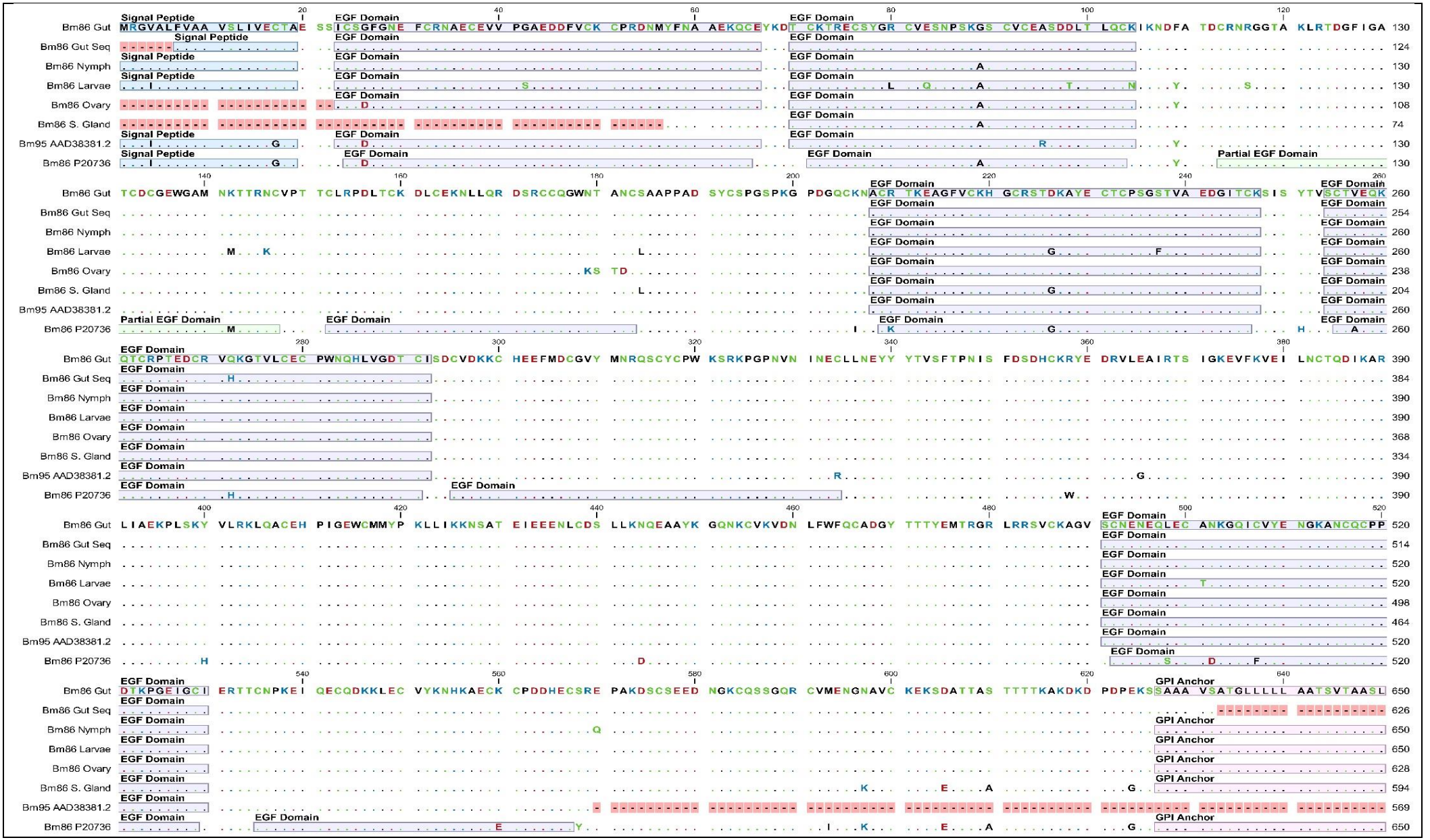

Signal Peptide
EGF Domain
Partial EGF Domain
GPI Anchor
Bm86 Gut
Bm86 Gut Seq
Bm86 Nymph
Bm86 Larvae
Bm86 Ovary
Bm86 S. Gland
Bm95 AAD38381.2
Bm86 P20736

Figure 3.3: Alignment of Bm86 amino acid sequences extracted from the RNA transcriptome assemblies of a South African laboratory strain, Sanger sequenced transcripts using cDNA from the South African laboratory strain and Bm95 (Q9Y0V1). Dots: Conserved amino acids in comparison to the top sequence. Purple dashes: Missing amino acids. Highlighted residues: Ambiguous amino acids identified via Sanger sequencing. Pure black text among dots: changes that are tissue/ stage-specific in comparison to the top sequence. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar. EGF domains predicted by Nijhof *et al*. (2010) is shown on the Yeerongpilly sequence. For the rest of the sequences, only the predicted domains from this study are shown.

		1	2	3	4	5	6	7	8
Bm86 P20736	1		96,15	92,77	96,15	95,38	92,92	88,77	84,62
Bm86 Gut	2	96,15		96,15	99,69	97,69	95,54	90,31	86,46
Bm86 Gut Seq	3	92,77	96,15		95,85	94,00	93,48	88,20	87,97
Bm86 Nymph	4	96,15	99,69	95,85		97,69	95,54	90,31	86,31
Bm86 Larvae	5	95,38	97,69	94,00	97,69		94,00	89,23	84,77
Bm86 Ovary	6	92,92	95,54	93,48	95,54	94,00		92,83	82,92
Bm86 S. Gland	7	88,77	90,31	88,20	90,31	89,23	92,83		77,85
Bm95 AAD38381.2	8	84,62	86,46	87,97	86,31	84,77	82,92	77,85	

Figure 3.4: Percent identities of Bm86 amino acid sequences extracted from the RNA transcriptome assemblies of a South African laboratory strain, Sanger sequenced transcripts using cDNA from the South African laboratory strain and Bm95 (Q9Y0V1). Colours are from a minimum in dark purple to maximum in red.

In Figure 3.3, several amino acid missense mutations were found in all life stages and tissues of the South African strain sequences for Bm86. All but ten of these changes occur within the EGF domains predicted by Nijhof *et al.* (2010). The Argentinean Bm95 (accession nr.: AAD38381.2, García-García *et al.,* 2000) is seen to have 19 amino acid differences, of which 12 results in alteration of the physicochemical property of the changed amino acid, and the protein lacks a GPI anchor when compared to the Yeerongpilly Bm86 (P20736) sequence. Additionally, the Bm95 sequence also has the lowest ID score to any sequence in Figure 3.4.

Comparison of the Yeerongpilly Bm86 sequence and the South African gut Bm86 (96.15% ID) shows a 25 amino acid difference of which only 1 was unconfirmed with Sanger sequencing by another member of our research group. This Sanger sequencing validated sequence is indicated as “Bm86 Gut Seq” in Figure 3.3. Only one conserved mutation was found at residue 89 (in the second EGF domain), where a glycine in the gut amino acid sequence was changed to alanine in all the other South African sequences.

When the South African sequences from the different life stages and tissues are compared to that from the gut sequence, more amino acid differences are evident. The Bm86 sequence in the ovary (95.54% ID to the gut) contains 7 amino acid differences compared to the gut (amino acid residues 26, 89, 109, 179, 180, 181, 182) with 4 appearing to be unique to the ovary sequence in this alignment (residues 179,180,181,182). Additionally, 1 amino acid difference was found in the Bm86 Yeerongpilly (92.92% ID) and Bm95 Argentinean-A (82.92% ID) sequences (amino acid residue 26), as well as 1 that is shared with all other sequences barring the gut and Bm95 (residue 89) and finally 1 amino acid change that is shared with all sequences except the gut and salivary gland (residue 109).

The Bm86 sequence from the salivary glands shows 7 amino acid differences to that from the gut (90.31% ID) amino acid residues 89, 184, 226, 597, 605, 610, 624). Only 1 alteration is shared with all other sequences barring the gut and Bm95 (77.85 % ID) at residue 89. The amino acid change at residue 184 is also seen in the larval transcriptome sequence (89.23% ID) for Bm86 as is that found at residue 226 which is also seen in the Yeerongpilly Bm86 (88.77% ID) sequence. All other amino acid differences in the salivary gland Bm86 sequence also appear in the Yeerongpilly Bm86 sequence (residues 597, 605, 610, 624).

The sequence from the nymph stage differs from that from the adult gut sequence with only 2 amino acid differences, and the sequences are 99.69% identical. The first

occurs at position 89, as with all other sequences except Bm95 (86.31% ID) and the gut. The second occurs at position 570, which is unique to the nymph stage. The nymphal amino acid differences remain to be corroborated via Sanger sequencing.

The sequence identified in the larval transcriptome contains 15 amino acid changes compared to that of the gut sequence (residues 4, 42, 80, 83, 89, 98, 104, 109, 116, 142, 146, 184, 226, 237, 501) and they are 97.69% identical. This comparison shows the most amino acid changes than any other Bm86 sequence comparisons extracted from any of the transcriptome assemblies here. The amino acid changes at residues 4, 42, 80, 83, 98, 104, 116, 146, 237, and 501 are unique to the larval Bm86 sequence (10 in total). The change at residue 89 is shared with all sequences except the gut and Bm95 (84.77% ID). The F to Y difference at residue 109 is shared with the ovaries (94% ID), Bm95 and the Yeerongpilly Bm86 (95.38% ID). The change seen at residue 184 is also only seen here and in the salivary gland (89.23% ID) while that at residue 226 is also seen in the salivary gland and the Yeerongpilly Bm86.

The protein domains for each Bm86 sequence were also predicted, and as seen in Figure 3.3, a signal peptide region is present in larvae, nymphs and adult gut tissues. However, none was found in the ovary and salivary gland Bm86 sequences.

In 2010, Nijhof *et al*. predicted 9 full EGF domains and 1 partial EGF domain (Nijhof *et al.* (2010). In the data from the assembly, only 5 of the full domains were identified in all sequences, including Bm95 (Figure 3.3). The partial EGF domain as well as domain 3,6, and 9 that were predicted by Nijhof *et al.* (2010) (shown on the reference Bm86 sequence in the alignment) were not corroborated here, and this requires further investigation. The Bm86 sequence from the adult salivary gland transcriptome is the only sequence missing the first EGF domain, but this requires further validation with Sanger sequencing. In all sequences, the GPI anchor was identified by domain prediction, except for in Bm95 and the partial Sanger sequence.

Verification of the larval Bm86 sequence in the South African strain

The larval sequence showed the most amino acid variation and was therefore selected for further validation by Sanger sequencing. Complementary DNA was synthesised from total RNA isolated for the same three samples used for RNA sequencing and used for subsequent PCR amplification and Sanger sequencing. First, the RNA quality was evaluated (Figure 3.5) and concentrations determined.

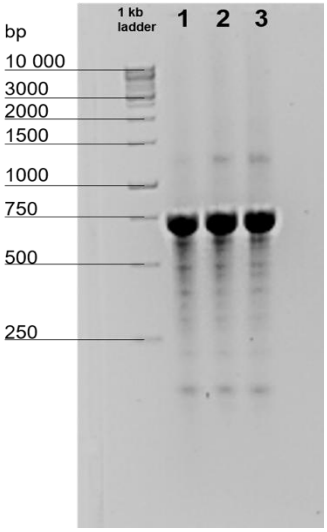


Figure 3.5: 2% Agarose gel of three Larval RNA samples used for cDNA production for Bm86 larval sequencing. 1 kb Thermo Scientific GeneRuler used for a size reference. Lane 1: RNA sample 1, Lane 2: RNA sample 2 Lane 3: RNA sample 3.

The RNA quality for all three samples was deemed appropriate to continue with cDNA synthesis and PCR. From the PCR results (Figure 3.6), it is evident that all 3 primer pairs (see Chapter 2) and the positive actin control produced amplicons of the expected sizes. The PCR products were subsequently purified and used for Sanger sequencing.

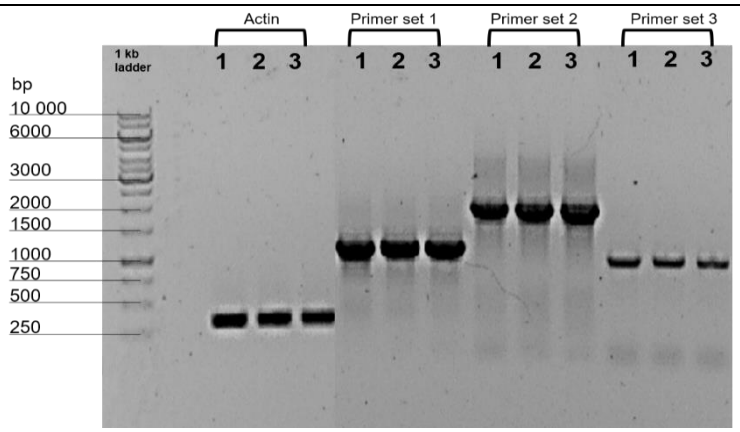


Figure 3.6: 1% Agarose gel of PCR amplicons of Bm86 from larval cDNA using three primer sets. 1 kb Thermo Scientific GeneRuler used as a size standard. Numbers correspond to the RNA/cDNA sample number used in Figure 3.5.

The PCR products were sequenced, translated into the corresponding amino acid sequences and aligned (Figure 3.7). From these results, 15 amino acid differences were observed via Sanger sequencing.

The final sequence identified in the larval transcriptome contains 15 amino acid changes compared to that of the gut sequence (residues 4, 42, 80, 83, 89, 98, 104, 109, 116, 142, 146, 184, 226, 237 and 501 in Figure 3.3). However, only 11 of these changes were present in the Sanger sequence (residues 42, 80, 83, 89, 98, 104, 109, 116, 142, 146 and 237 in Figure 3.7) leaving 4 uncorroborated (residues 4, 184, 226, 501). From the results, it is evident that all the confirmed amino acid differences occur in the N-terminal region (amino acids 42-146) of the South African Bm86 larval sequence. It is also interesting to note that the Sanger sequencing identified 4 residue changes not seen in the larval transcriptome assembly sequence but observed in the ovary transcriptome sequence (residues 179, 180, 181, 182).

The presence of two variants of the sequence was also observed in the Sanger sequencing results due to heterologous sites being present where two possible nucleotides were called at a single site (this is true for every amino acid change observed in the transcriptome assembly). These changes are illustrated in the nucleotide alignment (Appendix 2), and the frequencies at which each change in the sequence occur are summarised in Table 3.3.

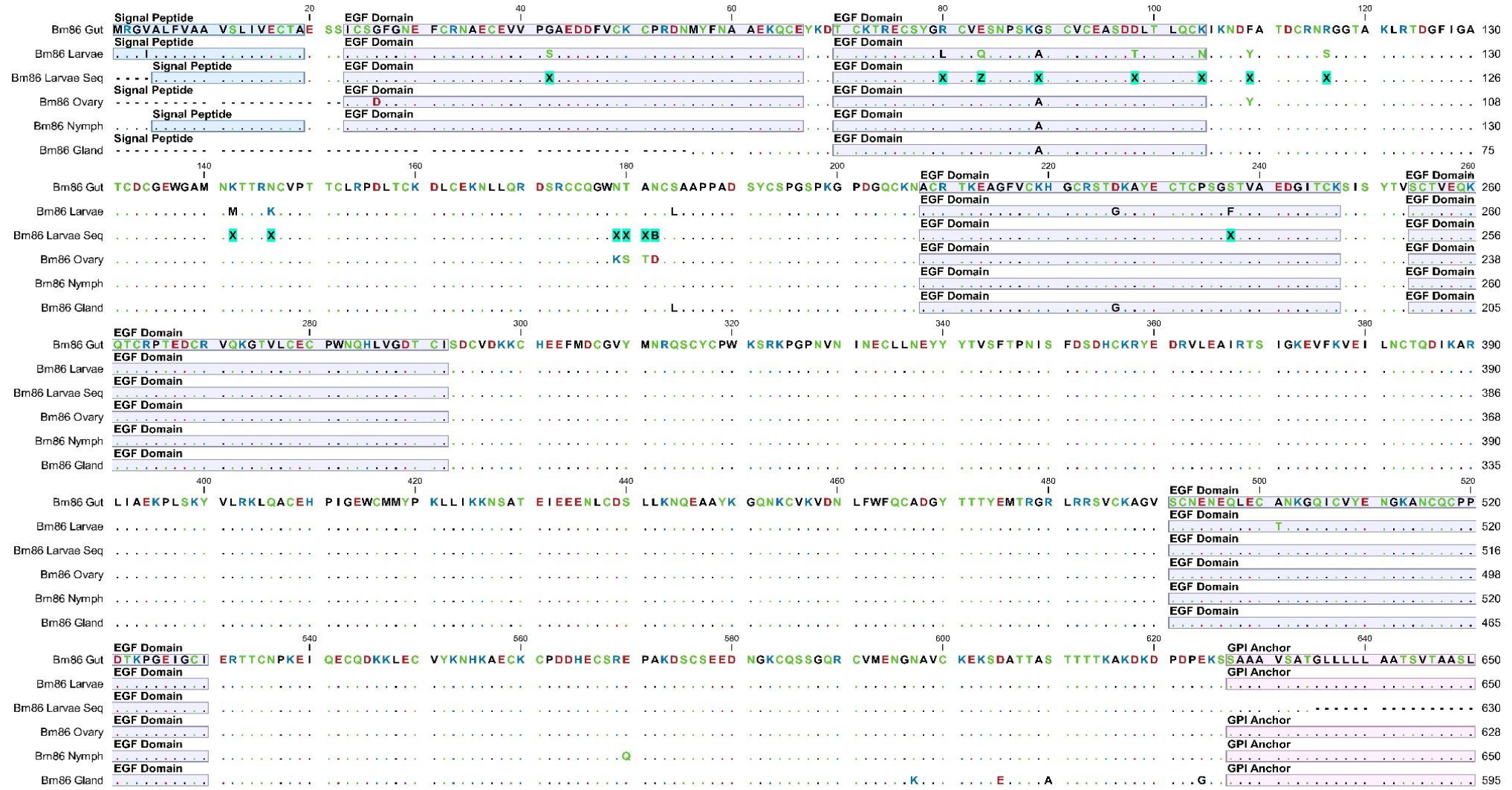


Figure 3.7: Amino acid alignment of Bm86 sequences from the transcriptome assembly and Sanger sequencing from larval cDNA. Dots represent conserved amino acids, while dashes represent missing amino acids in a position. Amino acid changes confirmed by sanger sequencing are highlighted in bright blue. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar

Table 3.3: A summary of missense mutations corroborated by Sanger sequencing for the larvae sequence.

Summary of amino acid changes and their relevant missense mutation				**Summary of missense nucleotide changes**		
Symbol [a]	Amino Acid difference	Frequency	Nucleotide in sequence [b]	Nucleotide IUPAC symbol	Nucleotide change	Frequency of nucleotide difference
B	N or D	1	R^1	Y	C or T	1
X^{42}	G or S	1	R^1	R	A or G	4
X^{80}	R or L	1	K^1	W	A or T	4
X^{89}	G or A	1	S^1	S	G or C	2
X^{98}	D or T	1	R^1 and M^2	K	T or G	1
$X^{104;146;179}$	K or N	3	M^1	M	C or A	3
X^{109}	Y or F	1	W^1			
X^{116}	R or S	1	M^1 and W^3			
X^{142}	M or K	1	W^1			
X^{180}	T or S	1	W^1			
X^{181}	A or T	1	R^1			
X^{237}	F or S	1	Y^1			
Z	E or Q	1	S^1			
Total		15				15

[a] Superscripts indicate residue number

[b] Superscript indicates the position in the codon

It is evident from the summary in Table 3.3 that the most frequent amino acid change is a lysine (K) to asparagine (N) which occur three times (position 104, 146 and 179). All the other amino acid changes occur only once. In all the cases, the K to N change resulted from adenine to cytosine in the nucleotide sequence in the first position of the codon. The most frequent nucleotide changes are for A or T and A or G, which occurs 4 times (W and R). In two positions, two nucleotide changes in a single codon occurred (position 98 and 116).

Antigen 1 Sequence Variation

As done for Bm86, Antigen 1 was used to find hits in the assembled transcriptome. The sequence from the gut transcriptome appeared truncated as does the sequence from the larvae transcriptome, and thus no conclusions can be made about these hits. However, another member of the tick research group performed Sanger sequencing from PCR products for Antigen 1 from the gut and found 5 amino acid changes in the N-terminal region. The transcriptome sequences for Antigen 1 from the salivary gland displays 6 amino acid changes and that from the nymph life stage only 4, again all occurring in the N-terminal region. Antigen 1 was not identified in the transcriptome derived from the ovaries. In all cases, the sequences are highly conserved in the annotated domain regions as with Bm86. As the sequences of Antigen 1 are under intellectual property protection at the University of Pretoria, these sequences have been omitted in this thesis.

Heterotrimeric G proteins

In this study, we only focussed on the sequences from adult gut tissue as this will pave the way forward for the design of future experiments to validate our hypothesis that Bm86 functions in a G protein-mediated PLC pathway. Several hits for heterotrimeric G proteins in the gut were identified, aligned and a neighbour-joining phylogenetic tree constructed which include all known Acari and reviewed arthropod G protein subunit sequences (Figure 3.8). The analysis involved 51 amino acid sequences from which all ambiguous positions were removed. There was a total of 559 positions in the final dataset. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed.

From the phylogenetic tree, it is evident that each subunit of the heterotrimeric G proteins was detected in gut, namely α, β and γ. As G protein classes are defined based on the sequence and function of their alpha subunits (Neves *et al*., 2002) we further classified these into separate subtypes namely: G_s, G_i, G_o, G_q and an unclassified subtype.

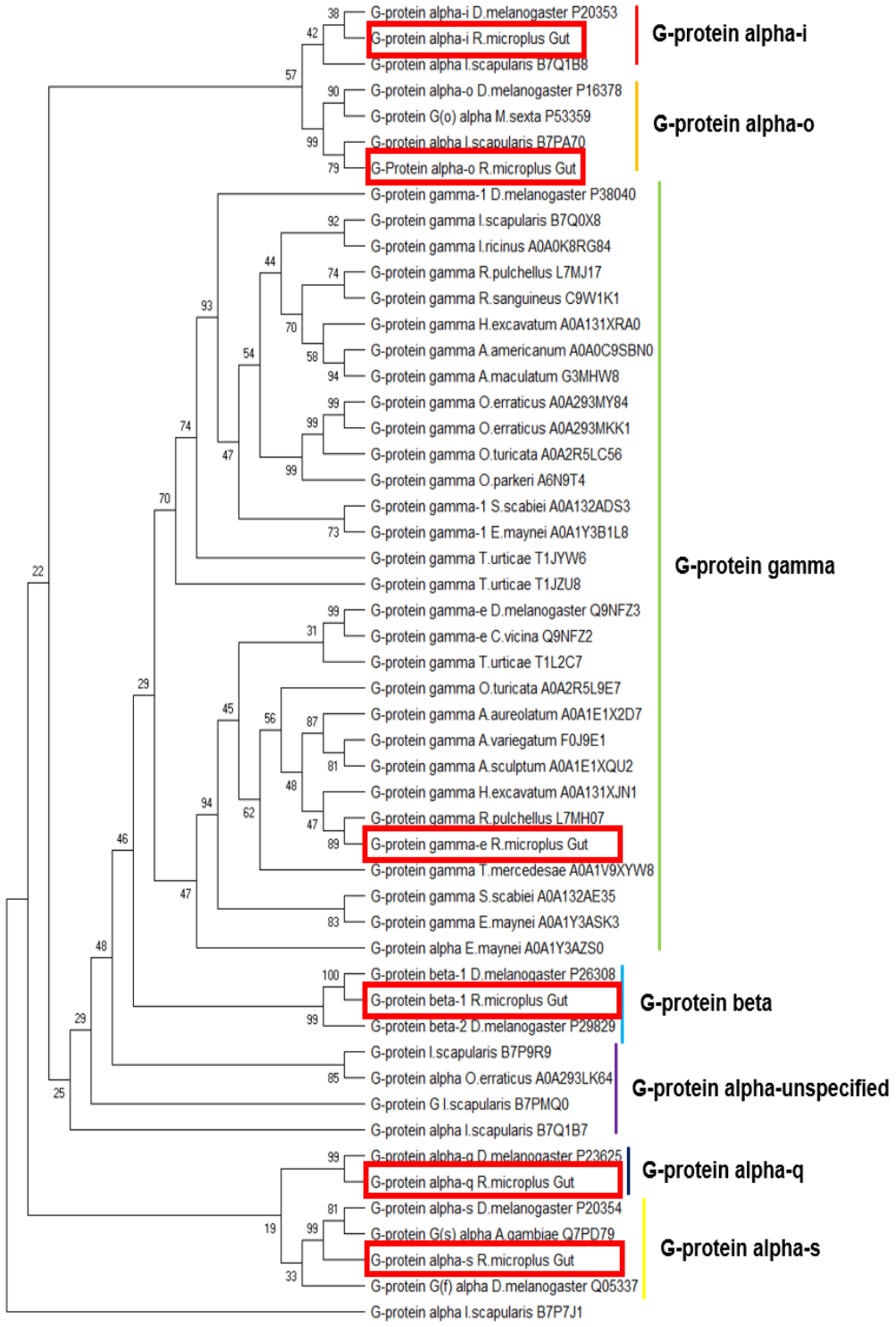


Figure 3.8: Relationships of heterotrimeric G proteins from the *R. microplus* gut transcriptome assembly with those of validated arthropod sequences and unreviewed G-protein sequences from Acari. The relationships were inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated protein sequences clustered together in the bootstrap test is shown next to the nodes. The bootstrap consensus tree was inferred from 10 000 replicates.

Following the phylogenetic analysis, each representative *R. microplus* amino acid sequence was aligned with a reference sequence for comparison of their domain architecture in order to further corroborate the putative functions of these predicted proteins (Figures 3.9-3.20).

In Figure 3.9 and 3.10, the sequence and domain architecture for the $G_{\alpha i}$ hit from the *R. microplus* gut transcriptome is compared with that of $G_{\alpha i}$ found in *D. melanogaster* (accession nr.: P20353, the only reviewed arthropod sequence on Uniprot). In total, 60 amino acid differences are evident (9%) with 27 of these being a change in amino acids that share the same chemical characteristics and are unlikely to affect protein function. However, 33 amino acid differences were observed that could alter protein folding and/or function such as a change involving the introduction of cysteine residues at positions 120, 142 and 354. Therefore, future functional assays will be vital to fully validate the protein function in *R. microplus*.

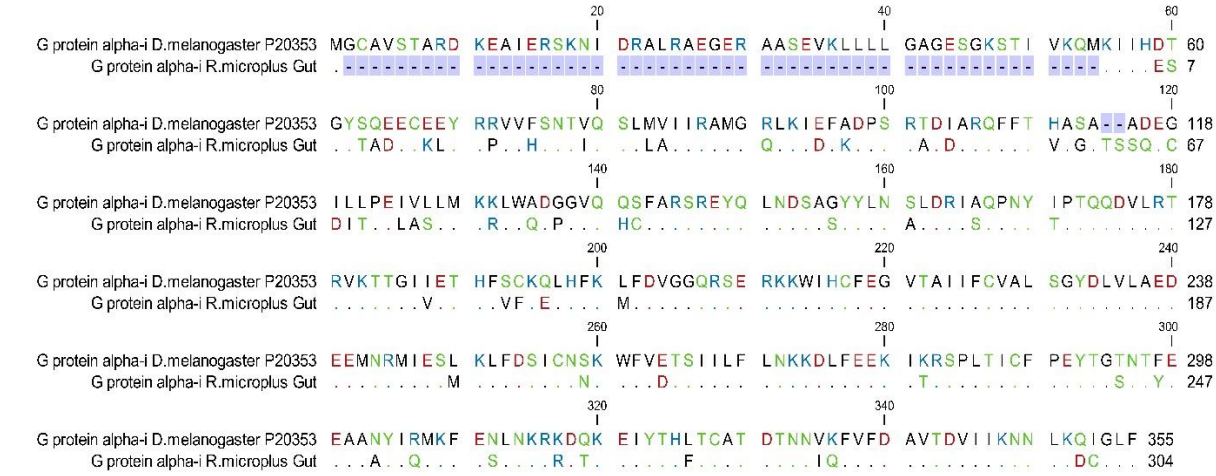


Figure 3.9: Amino acid alignment of $G_{\alpha i}$ from the assembled transcriptome and *D. melanogaster* P20353. Dots represent conserved amino acids, while dashes represent missing amino acids in the N-terminal (purple box). Amino acid substitutions are indicated by the corresponding single letter abbreviation for the changed amino acid. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar.

With regards to the domain analyses, all expected domains from the reference sequence for $G_{\alpha i}$ are present in the *R. microplus* transcript (Figure 3.10), supporting the hypothesis that the extracted sequence is indeed a $G_{\alpha i}$ subtype. The only functional site not seen in the extracted sequence is the GTP/Mg^{2+} binding site. This missing site may not have been picked up in the domain prediction search due to the missing amino acids at the beginning of the extracted sequence (seen the purple box in Figure 3.10) which may be an artefact of the assembly. Further sequence validation is, therefore, critical.

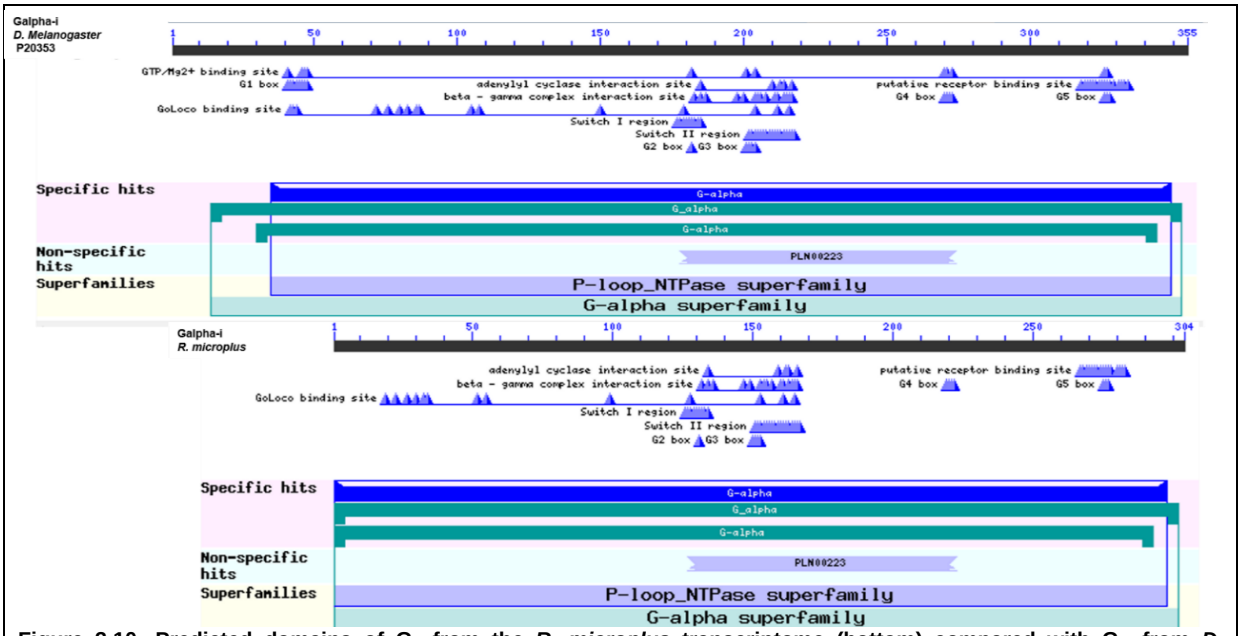


Figure 3.10: Predicted domains of $G_{\alpha i}$ from the *R. microplus* transcriptome (bottom) compared with $G_{\alpha i}$ from *D. melanogaster* P20353 (top). Domains were predicted using the NCBI: conserved domain database search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Triangles indicate specific binding sites common to all G proteins.

In Figure 3.11 and 3.12, the sequence for $G_{\alpha o}$ from the *R. microplus* gut transcriptome is compared with that of $G_{\alpha o}$ found in *D. melanogaster* (accession nr.: P16378, the only reviewed arthropod sequence on Uniprot). The *R. microplus* sequence lacks a single amino acid on the C-terminal that may be attributed to the nature of the RNA assembly. The extracted sequence is 91.8% conserved when compared to the reference sequence used, although containing 29 amino acid alterations. Of these changes, 15 are a change in amino acids that share the same chemical characteristics and are unlikely to affect protein function. However, 15 amino acid differences were observed, although no cysteine residues were introduced that could alter protein folding and/or function; amino acids which have other properties (like different polarities) may still affect the final protein in innumerable ways.

Figure 3.11: Amino acid alignment of $G_{\alpha o}$ from the *R. microplus* transcriptome with the $G_{\alpha o}$ from *D. melanogaster* P16378. Dots represent conserved amino acids, while dashes represent missing amino acids in the C-terminal (purple box). Amino acid substitutions are indicated by the corresponding single letter abbreviation for the changed amino acid. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar.

All expected domains from the $G_{\alpha o}$ *D. melanogaster* reference sequence are present in the *R. microplus* supporting that the extracted sequence is indeed a $G_{\alpha o}$ subtype. The presence of a GEM1 domain in the $G_{\alpha o}$ sequence provide further support that both $G_{\alpha o}$ and $G_{\alpha i}$ subtypes are present, as a GEM1 domain is only present in $G_{\alpha o}$.

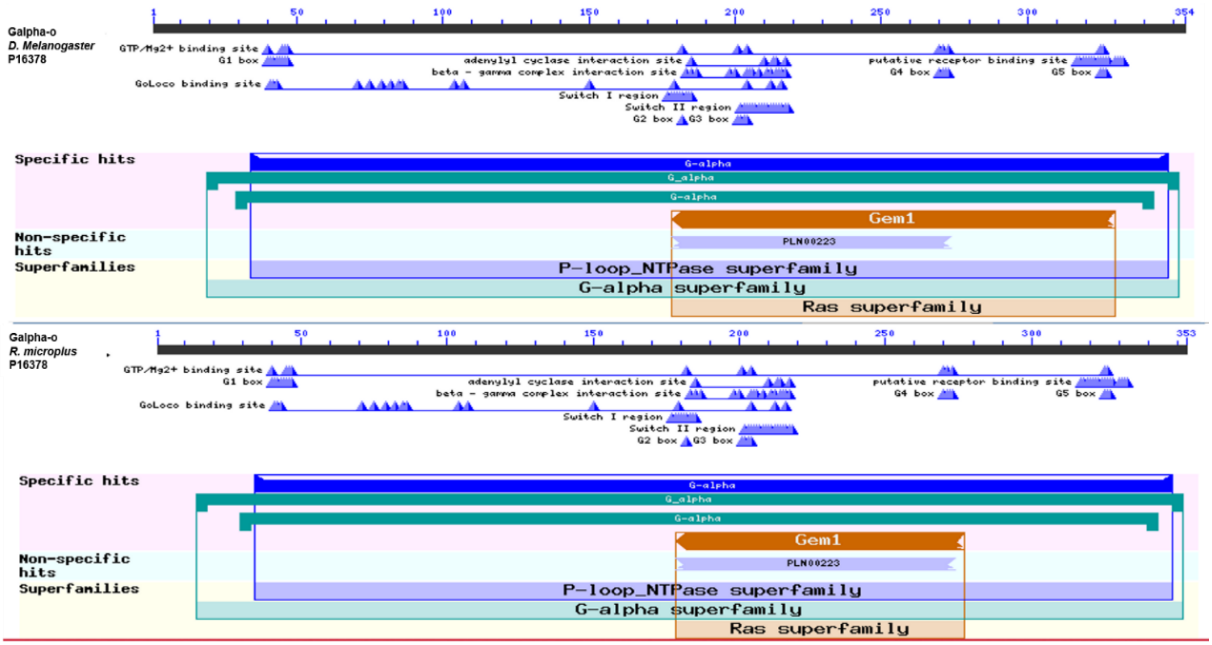


Figure 3.12: Predicted domains of $G_{\alpha o}$ from the transcriptome (bottom) compared with $G_{\alpha o}$ from *D. melanogaster* P16378 (top). Domains were predicted using the NCBI: conserved domain database search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Triangles indicate specific binding sites such as the GTP binding site common to all G proteins.

In Figure 3.13, the sequence for a putative $G_{\alpha q}$ from the *R. microplus* gut transcriptome is compared with that of the $G_{\alpha q}$ of *D. melanogaster* (accession nr.: P23625). In total, 65 amino acid changes (a change of 18%) in the *R. microplus* sequence are evident. From these changes, 29 are indicated to be a change in amino acids that share the same chemical characteristics and are unlikely to affect protein function. However, 35 amino acid differences were observed with different chemical properties. Only one Cysteine residue at position 324 is introduced, and one is changed from Cysteine to an Alanine that could alter protein folding and/or function. However, any amino acid changes resulting in different chemical properties may affect the final product.

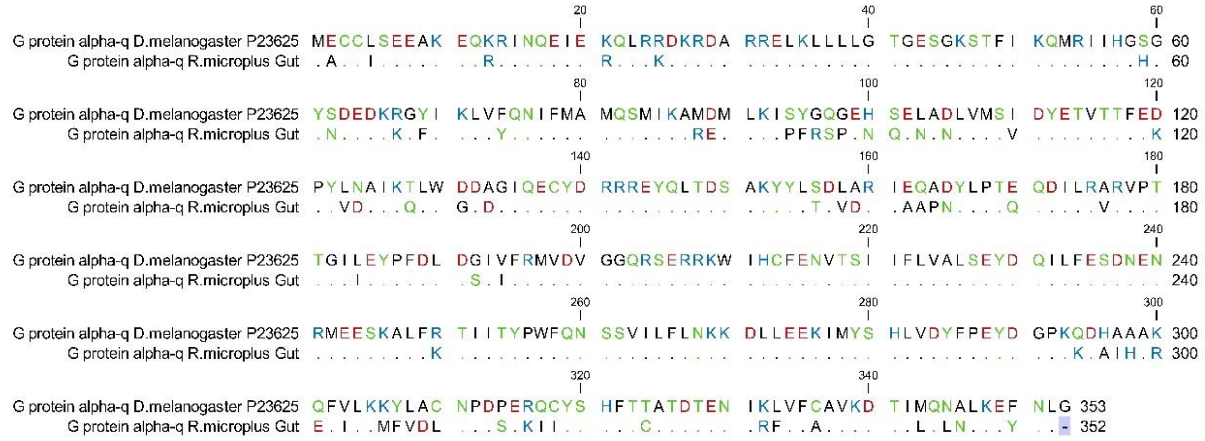


Figure 3.13: Amino acid alignment of $G_{\alpha q}$ from the transcriptome with $G_{\alpha q}$ from *D. melanogaster* P23625. Dots represent conserved amino acids, while dashes represent missing amino acids in the C-terminal (purple box). Amino acid substitutions are indicated by the corresponding single letter abbreviation for the changed amino acid. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar

In Figure 3.14, despite the missense mutations observed in Figure 3.13, all expected domains, aside from a small GTP domain from the reference sequence for $G_{\alpha q}$ are seen in the *R. microplus* sequence for $G_{\alpha q}$, supporting that the extracted sequence indeed encodes for a $G_{\alpha q}$ subtype.

Figure 3.14: Predicted domains of $G_{\alpha q}$ from the transcriptome (bottom) compared with $G_{\alpha q}$ from *D. melanogaster* P23625 (top). Domains were predicted using the NCBI: conserved domain database search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Triangles indicate specific binding sites such as the GTP binding site common to all G proteins.

In Figure 3.15 and 3.16, the sequence for $G_{\alpha s}$ from the *R. microplus* gut transcriptome is compared with that of $G_{\alpha s}$ found in *D. melanogaster* (accession nr: P20354). The extracted sequence is 85% conserved when compared to the reference sequence from another species, except for the consecutive deletion of 3 amino acids in positions 14-14 and 311-313 in the *R. microplus* sequence and 55 amino acid alterations. 22 of these changes being a change in amino acids that share the same chemical characteristics and are unlikely to affect protein function. However, 33 amino acid differences were observed introducing amino acids with different chemical properties that could alter protein folding and/or function despite there being no cysteine residues introduced.

In Figure 3.16, it is evident that despite the amino acid changes in the *R. microplus* sequence, all expected domains for $G_{\alpha s}$ are present.

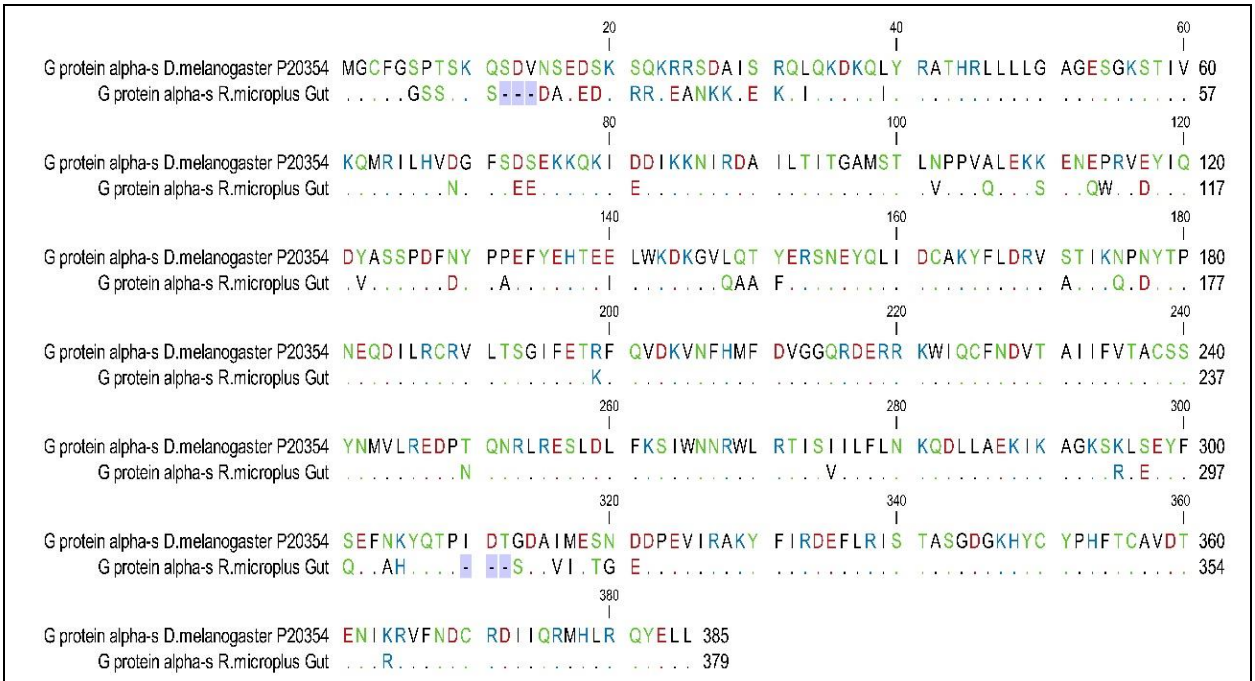


Figure 3.15: Amino acid alignment of $G_{\alpha s}$ from the transcriptome with $G_{\alpha s}$ from *D. melanogaster* P20354. Dots represent conserved amino acids, while dashes represent missing amino acids in the C-terminal (purple box). Amino acid substitutions are indicated by the corresponding single letter abbreviation for the changed amino acid. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar

Figure 3.16: Predicted domains of $G_{\alpha s}$ from the transcriptome (bottom) compared with $G_{\alpha s}$ from *D. melanogaster* P20354 (top). Domains were predicted using the NCBI: conserved domain database search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Triangles indicate specific binding sites such as the GTP binding site common to all G proteins.

In Figure 3.17, the sequence for G_γ from the *R. microplus* gut transcriptome is compared with that of G_γ found in *D. melanogaster* (accession nr.: Q9NFZ3). The extracted sequence is only 77% conserved, showing many amino acid alterations (17) given the small size of the protein (generally 74 amino acid in length). The majority (12) of these changes are a change in amino acids that share the same chemical characteristics and are unlikely to affect protein function. Only 5 amino acid differences were observed that could alter protein folding and/or function, and no cysteine residues were altered. However, the protein still contains all the expected domains and binding sites for a G_γ protein (Figure 3.18).

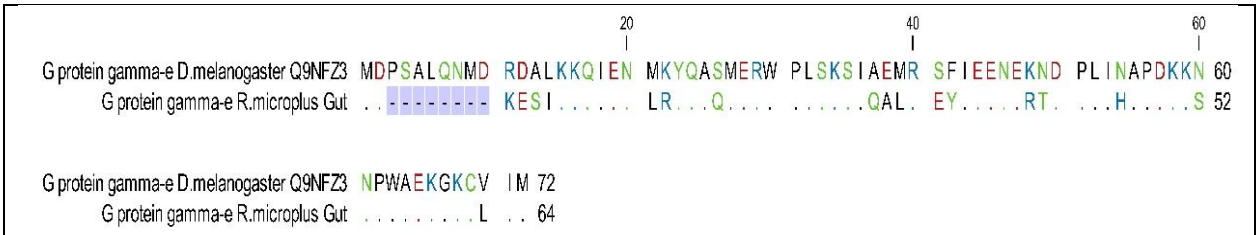


Figure 3.17: Amino acid alignment of G_γ from the transcriptome with G_γ from *D. melanogaster* Q9NFZ3. Dots represent conserved amino acids, while dashes represent missing amino acids in the N-terminal (purple box). Amino acid substitutions are indicated by the corresponding single letter abbreviation for the changed amino acid. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar

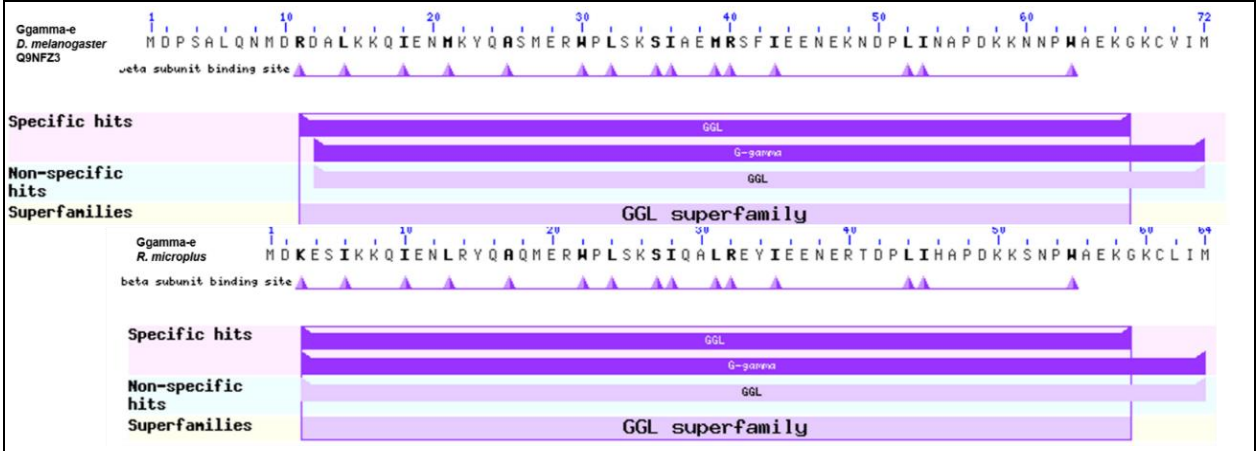


Figure 3.18: Predicted domains of G_γ from the transcriptome (bottom) compared with G_γ from *D. melanogaster* Q9NFZ3 (top). Domains were predicted using the NCBI: conserved domain database search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Triangles indicate specific binding sites such as the GTP binding site common to all G proteins.

In Figure 3.19 and 3.20, the sequence for G_β from the *R. microplus* gut transcriptome is compared with that of G_β found in *D. melanogaster* (accession nr.: P26308). The alignment in Figure 3.17 shows 40 amino acid changes (11%). More than half of these changes (i.e. 25 amino acids) share the same chemical characteristics and are unlikely to affect protein function. However, 15 amino acid differences were observed changing the chemical property of the amino acid at a residue, and 3 alterations of cysteine

residues are observed (residue 177, 195 and 196). All the latter might alter protein folding and/or function.

Again, the *R. microplus* sequence misses a single C-terminal amino acid which can be attributed to the nature of the RNA assembly and therefore needs to be validated with additional sequencing.

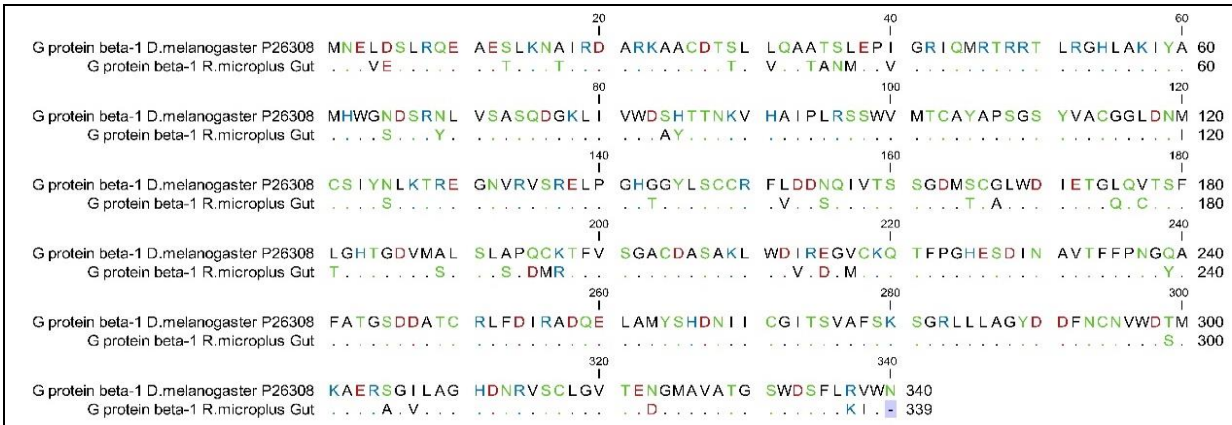


Figure 3.19: Amino acid alignment of G_β from the transcriptome with G_β from *D. melanogaster* P26308. Dots represent conserved amino acids, while dashes represent missing amino acids in the N-terminal (purple box). Amino acid substitutions are indicated by the corresponding single letter abbreviation for the changed amino acid. Amino acid abbreviations are coloured according to polarity: Green: neutral, polar; Black: neutral, nonpolar; Red: acidic, polar; Blue: basic polar

Figure 3.20 shows the highly conserved domain architecture for the *D. melanogaster* and *R. microplus* G_β proteins, which provide confidence that the extracted sequence is indeed a G_β subunit.

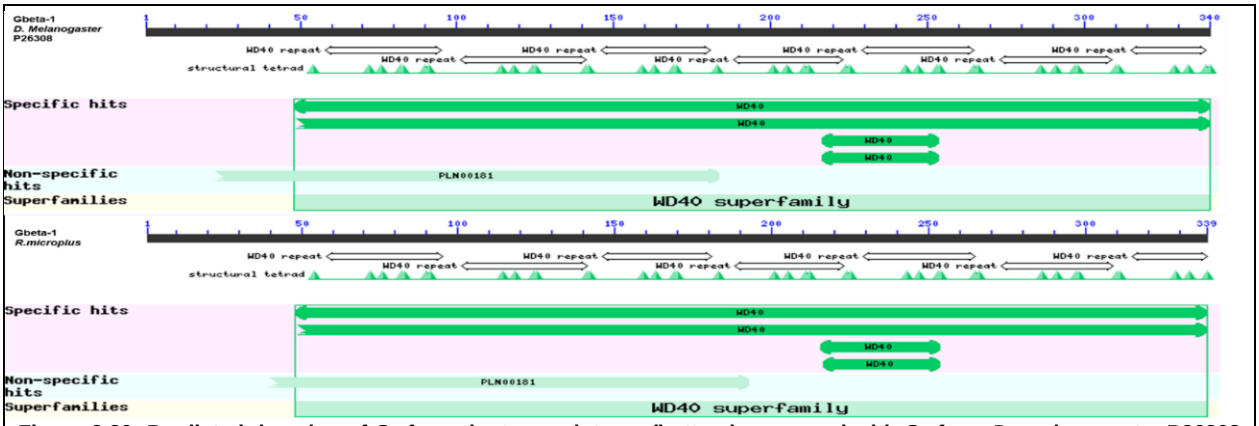


Figure 3.20: Predicted domains of G_β from the transcriptome (bottom) compared with G_β from *D. melanogaster* P26308 (top). Domains were predicted using the NCBI: conserved domain database search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Triangles indicate specific binding sites such as the GTP binding site common to all G proteins.

For all proteins discussed here, future functional assays will be vital to validate the protein function/s in *R. microplus* fully.

Phospholipase C proteins (PLCs)

Possible sequences for PLC proteins were identified, aligned and used for the construction of a phylogenetic tree that includes all known Acari and reviewed arthropod PLC sequences (Figure 3.21). The analysis involved 56 amino acid sequences. There was a total of 3 466 positions in the final dataset.

Currently, PLCs are classified into six isotypes (β, γ, δ, ε, ζ, η) according to their structure (Rhee and Bae, 1997). Within these isotypes, there are multiple subgroups within the PLC_{β}, PLC_{γ} and PLC_{δ} classes (Bamji-Mirza and Yoa, 2011). In this study, only a Phospholipase C_{γ} and a putative PLC_{β}-21C were identified from the gut transcriptome assembly. Possible PLC_{ε}, $PLC_{\beta4}$ and a $PLC_{\delta/\eta}$ were identified in the total transcriptome, indicating that they are present in other tick tissues and/or life stages.

With regards to the PLC_{ε} from *R. microplus*; it groups at a node with a bootstrap value of 100 in the phylogenetic tree (Figure 3.21). In the tree, the PLC_{ε} from *Ixodes scapularis* groups as an outgroup with the PLC_{ε} identified here. This may indicate these PLC_{ε} proteins share architecture and sequence that is unique to ticks (see Chapter 4), which remains to be corroborated.

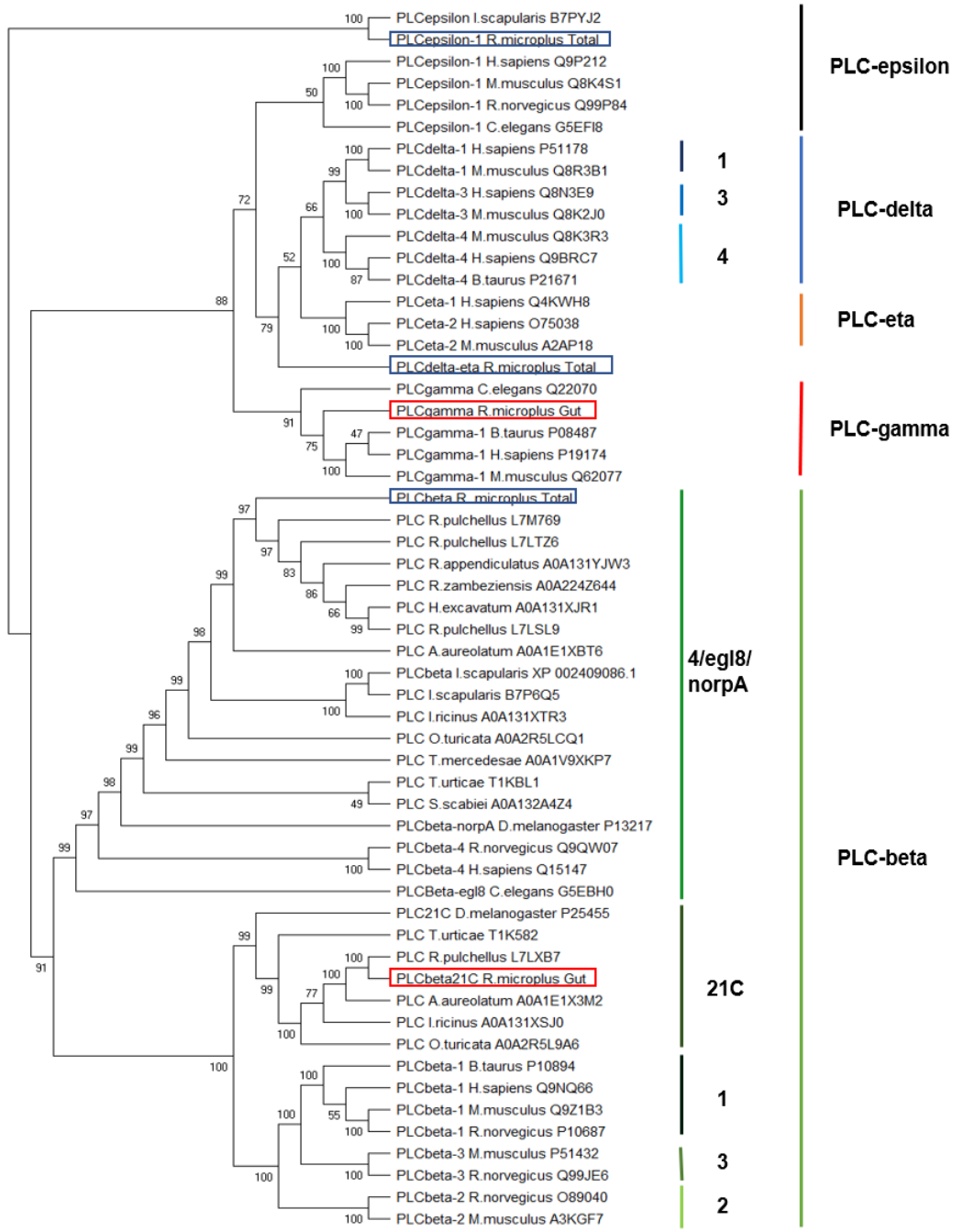


Figure 3.21: Relationships of PLC proteins from the *R. microplus* transcriptome assembly with those of arthropods and putative, unreviewed Acari PLC sequences. The phylogenetic tree was inferred using the Neighbour-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10 000 replicates) are shown next to the branches. Analyses were conducted in MEGA X.

The predicted domain structures for the PLC sequences extracted from the *R. microplus* transcriptome were compared with reference sequences as a means of corroborating the protein identity and putative function. Figures 3.22-3.26 illustrate the domains found by a conserved domain search on the NCBI conserved domain site (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), as well as by Prosite through the Expasy server (https://prosite.expasy.org/). As many PLC proteins contain more than 1 000 amino acids, only the domain architecture is shown and not the amino acid sequence alignments.

Figure 3.22 compares PLC$_\varepsilon$ found in the *R. microplus* total transcriptome assembly with that of PLC$_\varepsilon$ from *C. elegans* (accession number G5EFI8). The *R. microplus* sequence hit was found to be shorter than the reference sequence, and only two of the expected domains could be detected. An additional 'FERM' domain was detected in *R. microplus*, but this needs to be re-evaluated once a full-length sequence is available.

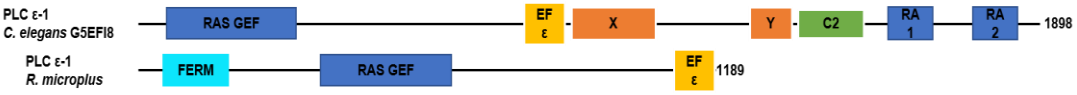


Figure 3.22: A comparison of a reviewed sequence domain structure for PLC$_\varepsilon$ from *C. elegans* G5EFI8 with that of PLC$_\varepsilon$ from the total transcriptome assembly of *R. microplus*. The RAS GEF and RA 1 and 2 domains are indicated in blue while the EF$_\varepsilon$ hand domain is indicated in yellow. The catalytic X and Y box is in orange, and the Ca^{2+} binding site C2 is indicated in green. In aqua, a FERM domain is indicated on the *R. microplus* PLCε sequence. Amino acid residue number is indicated to the right.

Figure 3.23 compares a reviewed sequence domain structure for PLC$_\eta$ from *H. sapiens* (accession number Q4KWH8) and PLC$_\delta$ from *H. sapiens* (accession number P51178) with the potential PLC$_\delta$ or $_\eta$ extracted from the *R. microplus* total transcriptome assembly. Again, the hit from the transcriptome is not full length and contains only 437 amino acids. As such, it is not possible to classify the PLC subtype, as it only contains an X-Y box and a C2 domain. Further sequencing is needed to confirm the full domain complement of this sequence.

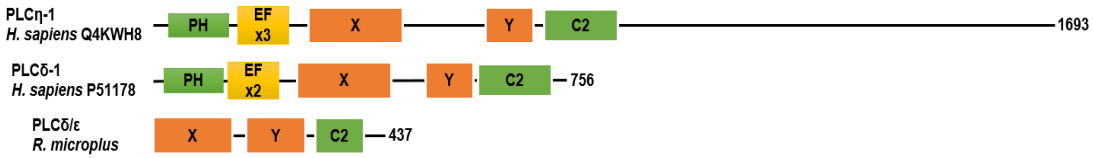


Figure 3.23: A comparison of a reviewed sequence domain structure for PLC$_\eta$ from *H. sapiens* Q4KWH8 and PLC$_\delta$ from *H. sapiens* P51178 with the potential PLC$_\delta$ or $_\eta$ extracted from the *R. microplus* total transcriptome. The EF-hand domains are indicated in yellow. The catalytic X and Y box is in orange while the PH and the Ca^{2+} binding site C2 are indicated in green. Amino acid residue number is indicated to the right.

Figure 3.24 illustrates the domain structure of a PLC$_\gamma$ from *H. sapiens* (accession number P19174) as a representative of the general structure of a PLC$_\gamma$ as there is no reviewed sequence for PLC$_\gamma$ for arthropods yet. It was not possible to extract the full sequence from the gut transcriptome assembly, and so it lacks the first two domains. The protein does, however, contain the unmistakable central PLC$_\gamma$ domain architecture.

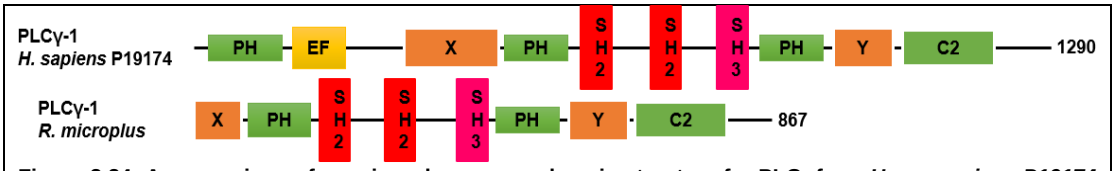


Figure 3.24: A comparison of a reviewed sequence domain structure for PLC$_\gamma$ from *Homo sapiens P19174* with the potential PLC$_\gamma$ extracted from the *R. microplus* gut transcriptome assembly. The EF-hand domains are indicated in yellow. The catalytic X and Y box is in orange while the PH and the Ca^{2+} binding site C2 are indicated in green. In red are the SH2 domains and in pink the SH3 domain. Amino acid residue number is indicated to the right

Figure 3.25 compares the domain structure for a validated PLC$_\beta$ from *H. sapiens* (accession number Q15147) and *D. melanogaster* (accession number P13217) with the putative PLC$_\beta$ extracted from the *R. microplus* total transcriptome. The protein extracted from the transcriptome is again truncated and lacks the PLC beta C domain, which again will need additional sequencing and domain analyses.

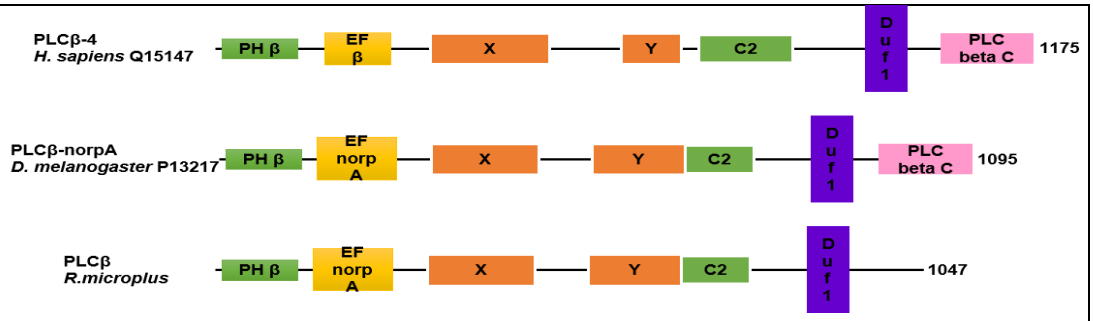


Figure 3.25: A comparison of reviewed sequence domain structure for PLC$_\beta$ from *H. sapiens* Q15147 *and D. melanogaster* P13217 with the potential PLC$_\beta$ extracted from the *R. microplus* total transcriptome. The EF-hand domains are indicated in yellow. The catalytic X and Y box is in orange while the PH and the Ca^{2+} binding site C2 are indicated in green. In purple a DUF1 domain and in pink a PLC$_\beta$ C terminal domain. Amino acid residue number is indicated to the right

In Figure 3.26, the final PLC identified in the *R. microplus* gut transcriptome assembly, a potential PLC$_\beta$-21C, is compared to the reviewed PLC$_\beta$-21C from *D. melanogaster* (accession number P25455). Here, all domains were successfully identified.

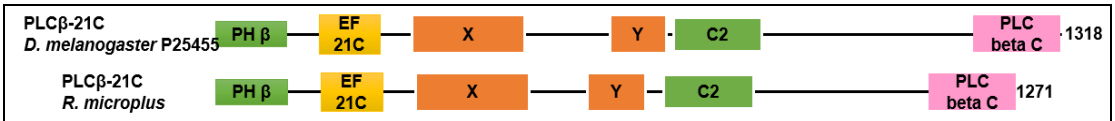


Figure 3.26: A comparison of reviewed sequence domain structure for PLC$_\beta$-21C from *D. melanogaster* P25455 with the potential PLC$_\beta$-21C extracted from the *R. microplus* gut transcriptome assembly. The EF-hand domains are indicated in yellow. The catalytic X and Y box is in orange while the PH and the Ca^{2+} binding site C2 are indicated in green. In pink a PLC$_\beta$ C terminal domain. Amino acid residue number is indicated to the right.

The identification of these proteins was done *in silico*. Given the numerous amino acid alterations seen in the G proteins and the truncated amino acids noted in the PLC analyses thus far, the existence of these proteins *in vivo* requires validation. Before in vivo testing can be undertaken, prior validation via *in vitro* methods is recommended.

Preliminary, rapid ELISA screening of commercial antibodies

In order to verify the presence of proteins identified via RNA-sequencing analysis, a protein extract from *R. microplus* midgut tissues was prepared. Protein concentrations higher than 500 µg/ml were obtained for each cytosolic/peripheral and membrane protein fraction which was deemed appropriate to proceed with further analysis.

During ELISA, a signal was obtained with an initial OD_{490} greater than 2 in the membrane fraction for the Bm86 control. Similar results were obtained for tubulin (OD_{490} greater than 1) and retinol dehydrogenase (OD_{490} greater than 0.5) controls in the cytosolic/peripheral protein portion. However, weak signals were observed (OD_{490} less than 0.3) for all the commercial antibodies evaluated (Figure 3.27). The commercial antibodies used are readily available and directed against human and rat protein homologs. These antibodies appear to be too specific to effectively detect the tick proteins screened for. Therefore, the tick proteins may not present the same epitopes that can be detected by antibodies directed at mammalian proteins. As such, tick-specific antibodies will be required for further validation, before proceeding to any downstream experimentation.

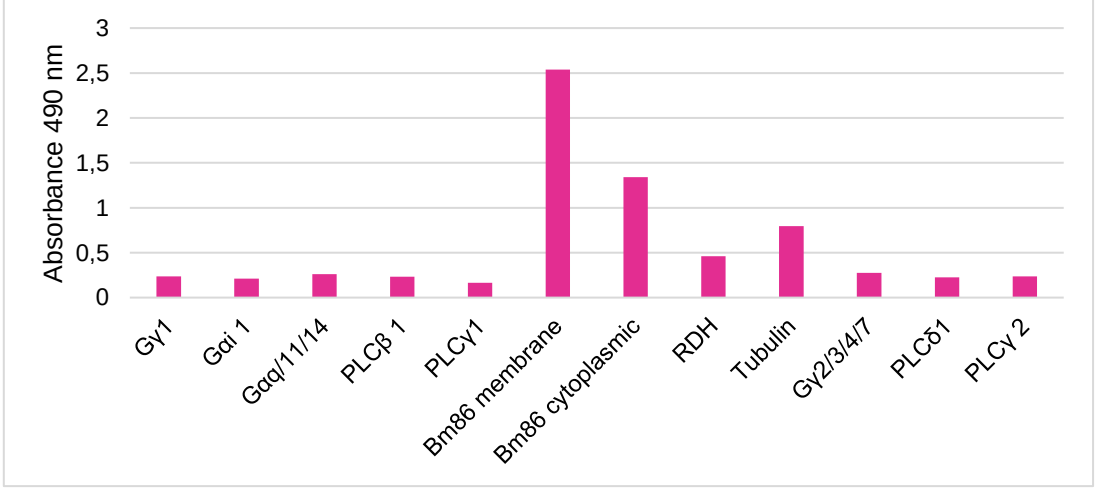


Figure 3.27: Evaluation of commercial antibodies for the identification of G protein and PLCs in *R. microplus*.

Yeast two-hybrid mapping of the region involved in the interaction between Bm86 and Antigen 1

PCR amplification of inserts

DNA fragments for directional cloning into the yeast-two-hybrid plasmids were PCR amplified from already available constructs Antigen 1 was amplified as three different lengths termed truncation 1 (T 1, only the N-terminal of antigen 1 without any Kunitz domain), truncation 2 (T 2, with only the first domain) and the full-length open reading frame (T 3). All bands obtained were of the expected sizes (Figure 3.28).

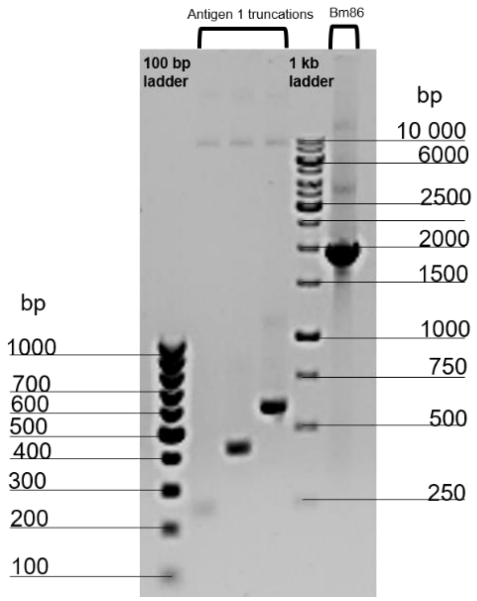


Figure 3.28: Electrophoresis image of the PCR amplification of inserts for yeast-two-hybrid plasmids. Lanes correspond to (1) 100 bp markers, (2) Truncation 1 (T1), (3) Truncation 2 (T2), (4) full-length Antigen 1 (T3), (5) 1 kb marker and (6) full-length Bm86.

Restriction enzyme digestion of PCR product and plasmid for ligation

The vector plasmid pGADT7 and the PCR amplified inserts of Antigen 1 (constructs T1 to T3) were restriction digested with NdeI and BamHI, while the pGBKT7 vector plasmid and the PCR amplified Bm86 were digested with NdeI and XmaI, for directional cloning. The results of the digestions are presented in Figure 3.29. Only one band of the expected size is observed for the digested plasmids.

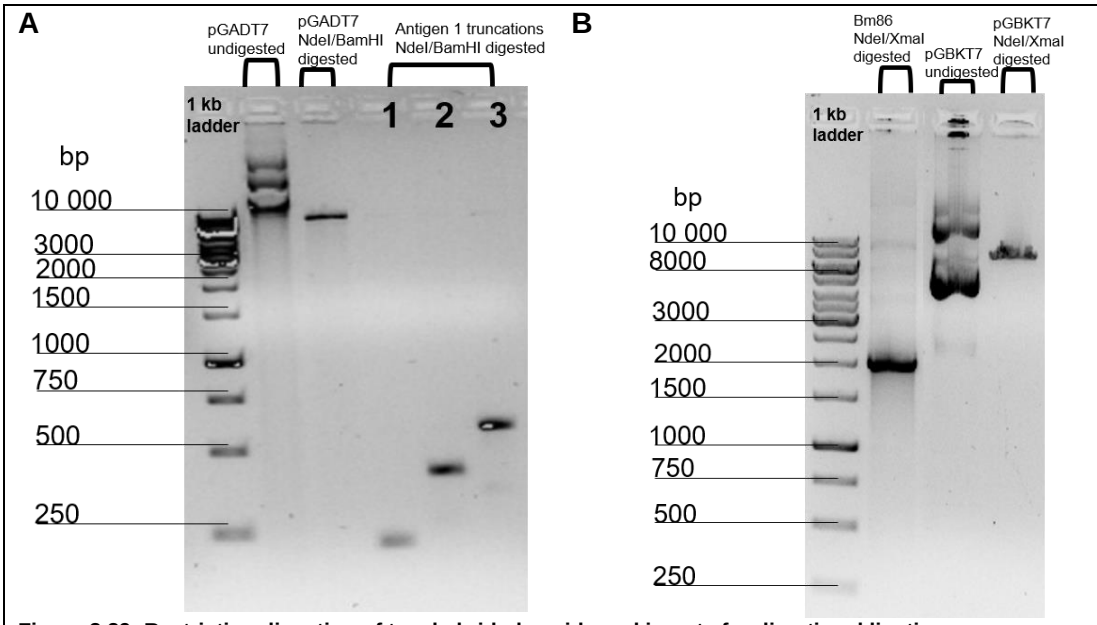


Figure 3.29: Restriction digestion of two-hybrid plasmids and inserts for directional ligation.
Panel A corresponds to the restriction digestion Antigen 1 and the pGADT7 plasmid with NdeI and BamHI. All lanes are labelled with headings. Lanes marked as 1-3 correspond to the three Antigen 1 truncations. Panel B corresponds to the restriction digestion of Bm96 and the pGBKT7 plasmid with NdeI and XmaI. All lanes are labelled with headings.

Transformation of DH5α *cells*

pGADT7 and pAS2-1 plasmids were transformed into electrocompetent DH5α *E. coli* cells by electroporation and plated on LB plates containing Ampicillin. pGBKT7 plasmids were transformed into chemically competent DH5α *E. coli* cells via chemical transformation and plated on LB plates containing Kanamycin. All the transformation efficiencies were greater than 1 X 10^8. Colony PCR was used to verify recombinant clones which were grown overnight and used for plasmid extraction and DNA sequence verification.

Verification of constructs by DNA sequencing

The sequences of the pGADT7-Antigen 1 constructs determined from automated Sanger sequencing, were aligned using the CLC Main Workbench v. 8.0.1. Plasmid sequence was identified upstream and downstream of the insert as expected along with the expected restriction enzyme sites in the beginning and end of the insert, confirming the insert orientation, as well as the maintenance of insert sequence (results not shown due to IP policy of University of Pretoria). pGBKT7-Bm86 and pAS2-1-Bm86 sequences were verified, and no mutations were detected (results not shown).

In order to confirm that no changes in the vector backbone occurred, each construct was analysed using restriction fragment length polymorphism mapping (RFLP). All restriction digested fragments resulted in the expected size bands (Table 3.4 and Figure 3.30).

Table 3.4: Expected band sizes following RFLP mapping of constructs.

Plasmid	Enzyme	Buffer	Expected sizes (kbp)	Expected number of bands
pGADT7	*Hind III*	CutSmart NEB	7.1; 0.8	2
pGADT7-T	*XhoI/EcoRI*	NEB 2.1	8.0; 2.0	2
pGADT7-T1	*BtgI*	CutSmart NEB	5.8; 1.2	2
pGADT7-T2	*BtgI*		5.8; 1.4; 1.2	3
pGADT7-T3	*BtgI*		5.8; 1.5; 1.2	3
pGBKT7	*Hind III*		4.9; 1.5; 0.9	3
pGBKT7-Bm86	*ApaLI*		3.5; 2.6; 2.1; 1.0	4
pGBKT7-53	*BamHI/EcoRI*	NEB 3.1	7.3; 1.0	2
pGBKT7-LAM	*BamHI/EcoRI*		7.3; 0.57	2

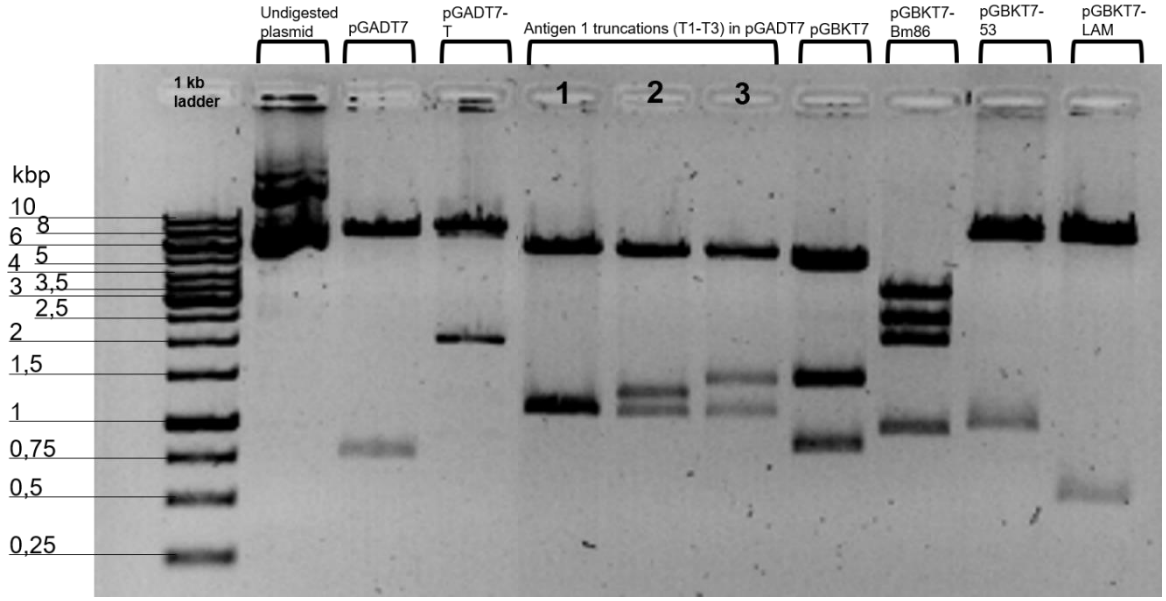


Figure 3.30: Restriction fragment length polymorphism analysis of plasmid constructs for the yeast-two-hybrid study. Each lane has a descriptive header as to which plasmid is shown. All plasmids were digested as described in the methods and show the expected band size indicating that the plasmids are valid and ligated as expected.

Transformation of yeast cells and evaluation of autoactivation

The bait plasmids (pGBKT7 and pAS2-1) contain the *TRP1* gene which confers the ability for autonomous tryptophan production in transformed yeast. Y2H Gold cells were transformed separately with pGBKT7 bait plasmids containing Bm86, the pGBKT7-LAM negative control, the pGBKT7-53 positive control (Clontech Matchmaker® Gold Y2H system), and a pAS2-1 construct also containing Bm86. Transformed cells were selected on single drop out media (SDO, -Trp) and a transformation efficacy of at least $1x10^6$ CFU was achieved for all plates obtained.

The prey pGADT7 plasmid contains the *LEU2* gene which confers the ability for autonomous leucine production in transformed yeast. Cells transformed with the Bm86 bait plasmids were sequentially transformed with the prey plasmid constructs (pGBKT7 with T1, T2 or T3), as well as the negative and positive controls provided. Co-transformed cells were selected on double drop out media (DDO, -Trp/-Leu), with a transformation efficacy in the order of $1x10^5$ CFU for all plates obtained.

Interaction between proteins produced by the insert sequences in the bait and prey plasmids in co-transformed cells allows the activation of reporter genes under the control of the *GAL4* promotor. This activation enables the yeast to autonomously produce histidine and adenine, and therefore grow on media lacking these amino acids (in addition to tryptophan and leucine). However, autoactivation of the reporter gene for histidine may occur and results in the growth of false positives. This was assessed by the inclusion 3-amino-1,2,4-triazole (3-AT) at increasing concentrations within dropout media.

To determine background growth (i.e. autoactivation) of the histidine reporter gene, the yeast that was transformed only with the bait plasmids was grown on media lacking tryptophan and histidine. The concentration of 3-AT at which minimal growth of transformants is observed after a week is the level of 3-AT at which growth of double transformants can be considered a genuine positive interaction (Figure 3.31).

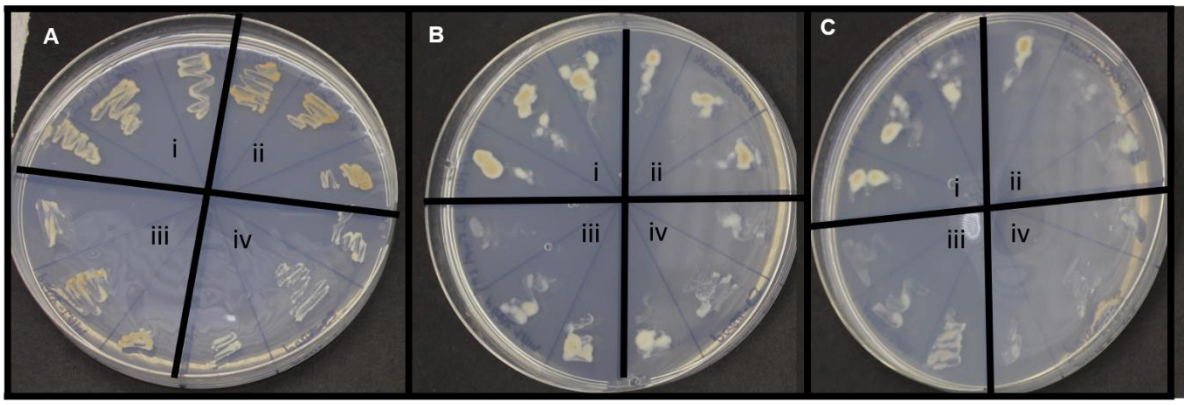


Figure 3.31: Evaluation of autoactivation of baits. In panel A, cells were grown in the absence of 3-AT on -Trp/-His media. In panels B and C cells were grown on -Trp/-His media containing 2,5 mM and 5 mM 3-AT, respectively. In all cases, the quarters indicated by (i) correspond to pAS2-1-Bm86 transformed Y2H Gold yeast, (ii) indicates pGBKT7-Bm86 transformed Y2H Gold yeast, (iii) indicates the negative control pGBKT7-LAM transformed Y2H Gold yeast and (iv) indicates the positive control pGBKT7-53 transformed Y2H Gold yeast.

From Figure 3.31.A, it is evident that there is background autoactivation for all the constructs in the Y2H Gold yeast strain. It is reduced with the addition of 2.5 mM (Figure 3.31.B) and 5 mM 3-AT (Figure 3.31.C) for all construct except for pAS2-1-Bm86. As such, pAS2-1 constructs were excluded from further analysis. Further screening of interactions was done using both concentrations of 3-AT on triple drop-out media.

Screening for protein-protein interactions

Co-transformed cells were replica plated on SD/-Trp/-Leu/-His (TDO) to screen for protein-protein interactions. In Figure 3.32, the control constructs provided with the kit is shown, and it is evident that the positive control pGBKT7-53 + pGADT7-T grows at the same level on TDO with all concentrations of 3-AT as expected while the negative control pGBKT7-LAM + pGADT7-T shows background growth without 3-AT and no growth at higher levels of 3-AT.

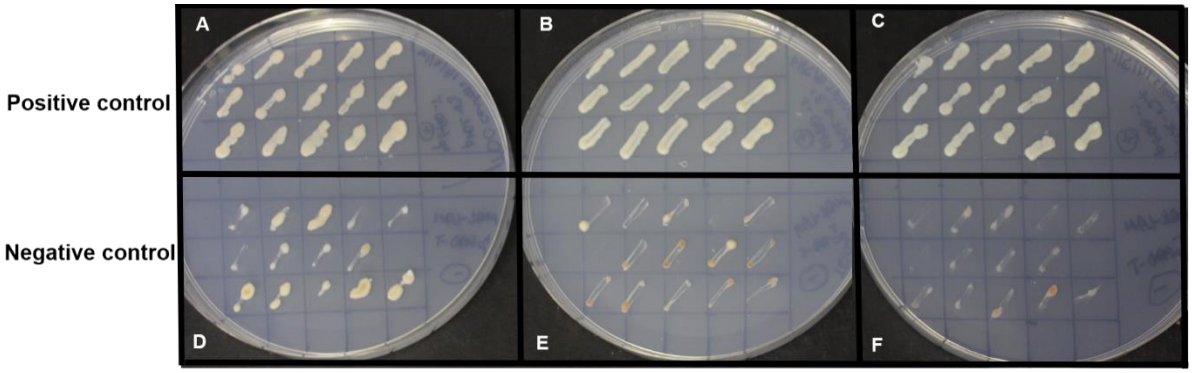


Figure 3.32: Y2H Gold yeast co-transformed with control plasmids provided with the kit on triple drop out medium (SD -Trp/-Leu/-His). In panel A, pGBKT7-53 and pGADT7-T on TDO with no 3-AT is shown. In panel B pGBKT7-53 and pGADT7-T on TDO with 2.5 mM 3-AT is shown. Panel C shows pGBKT7-53 and pGADT7-T on TDO with 5 mM 3-AT. Panel D shows pGBKT7-LAM and pGADT7-T on TDO with no 3-AT added. In panel E pGBKT7-LAM and pGADT7-T on TDO with 2.5 mM 3-AT. Panel F indicates pGBKT7-LAM and pGADT7-T on TDO with 5 mM 3-AT. Each streak represents a biological repeat (i.e. a separate colony selected) of the experiment.

Screening of Bm86 against full-length Antigen 1 (Figure 3.33) shows growth at all concentrations of 3-AT, indicating that these transformed cells express the reporter genes for the presence of both plasmids (Leu and Trp) and autonomously produce histidine at high enough levels to overcome the 3-AT inhibition. This corroborates that Bm86 interacts with the full-length Antigen 1.

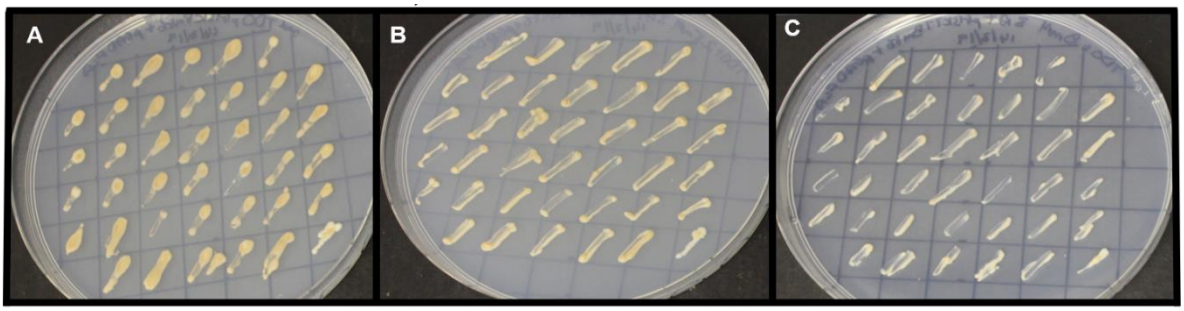


Figure 3.33: Y2H Gold yeast co-transformed with Bm86 and Antigen 1 (full-length (T3)) on triple drop out medium (SD -Trp/-Leu/-His). Panel A shows pGBKT7-Bm86 and pGADT7-Antigen 1 (Full length) on TDO with no 3-AT. Panel B shows pGBKT7-Bm86 and pGADT7-Antigen 1 (Full length) on TDO with 2.5 mM 3-AT. In panel C pGBKT7-Bm86 and pGADT7-Antigen 1(Full length) on TDO with 5 mM 3-AT is shown.

The interactions between Bm86 and truncation 2 is shown in Figure 3.34. Growth is evident at all levels of 3-AT tested as with the full-length Antigen 1 and as such indicates that the interaction of Antigen 1 with Bm86 is not dependent on the second BPTI/Kunitz domain or the C-terminal portion of Antigen 1.

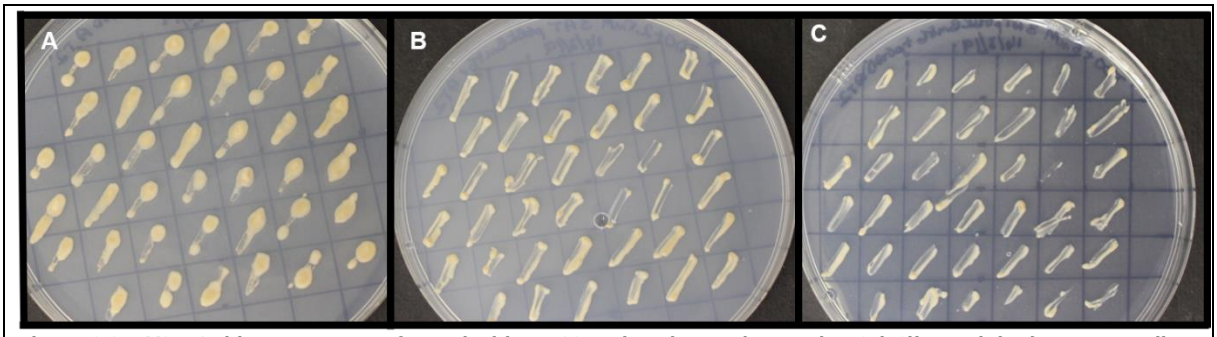


Figure 3.34: Y2H Gold yeast co-transformed with Bm86 and Antigen 1 (truncation 2 (T2)) on triple drop out medium (SD -Trp/-Leu/-His). Panel A shows pGBKT7-Bm86 and pGADT7-Antigen 1 (T2) on TDO with no 3-AT. Panel B shows pGBKT7-Bm86 and pGADT7-Antigen 1 (T2) on TDO with 2.5 mM 3-AT. Panel C shows pGBKT7-Bm86 and pGADT7-Antigen 1(T2) on TDO with 5 mM 3-AT.

Lastly, in Figure 3.35 screening of Bm86 interaction with truncation 1 shows growth at all levels of 3-AT tested, as with the full-length Antigen 1, indicating that that the interaction of Antigen 1 with Bm86 is not dependent on either BPTI/Kunitz domain or the C-terminal portion of Antigen 1. It is therefore proposed that Bm86 likely interacts only with the N-terminal region of Antigen 1.

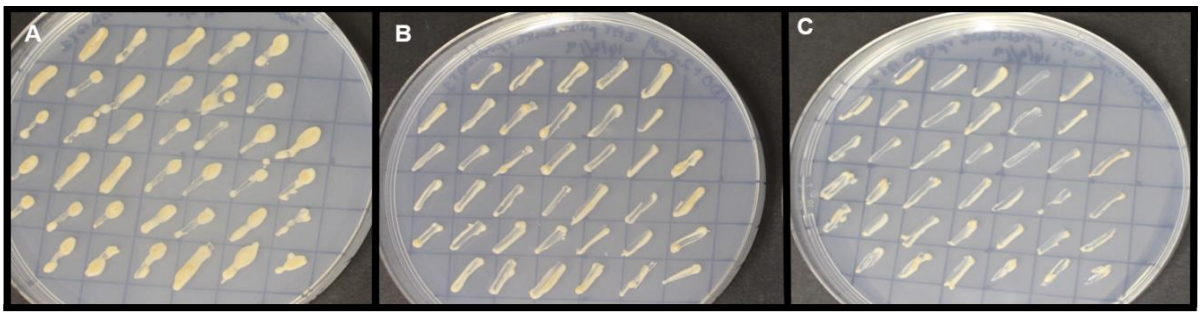


Figure 3.35: Y2H Gold yeast co-transformed with Bm86 and Antigen 1 (truncation 1 (T1): only the N-terminal) on triple drop out medium (SD -Trp/-Leu/-His). Panel A shows pGBKT7-Bm86 and pGADT7-Antigen 1 (T1) on TDO with no 3-AT. Panel B shows pGBKT7-Bm86 and pGADT7-Antigen 1 (T1) on TDO with 2.5 mM 3-AT. Panel C shows pGBKT7-Bm86 and pGADT7-Antigen 1(T1) on TDO with 5 mM 3-AT.

Clones that grew on TDO plates were transferred to SD/-Trp/-Leu/-His/-Ade quadruple drop out media (QDO). As QDO is the most stringent selection media, it selects for stable and strong interactions which will grow as white colonies. Cells exhibiting a reddish-pink appearance on QDO media indicate Adenine depletion and that the yeast is therefore not autonomously producing adenine in high enough amounts to be considered a strong-interaction. The latter is indicative of transient interactions that involve protein interactions that are formed and broken easily, which are commonly found in many aspects of cellular function (Acuner Ozbabacan *et al*., 2011).

In Figure 3.36, the control constructs provided with the kit are shown. It is evident that the positive control (Figure 3.34.A) has a strong interaction based on the white colonies observed on QDO plates, while the negative control shows no growth, indicating no interaction.

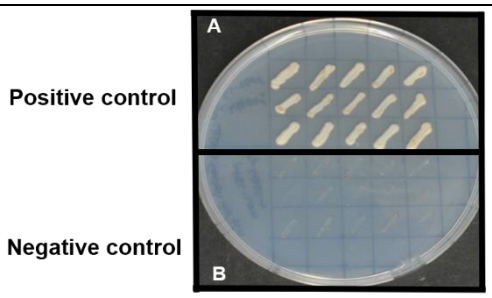


Figure 3.36: Y2H Gold yeast co-transformed with control plasmids provided with the kit on quadruple drop out medium (SD -Trp/-Leu/-His/-Ade). In panel A, the positive control pGBKT7-53 and pGADT7-T on QDO are shown. In panel B, the negative control pGBKT7-LAM and pGADT7-T on QDO are shown.

In Figure 3.37, the test constructs are shown on quadruple drop out medium. Again, growth is uniform across all truncations of Antigen 1, although the yeast displays a reddish colour in all instances. The latter indicates that the interaction of Bm86 with Antigen 1 may be transient.

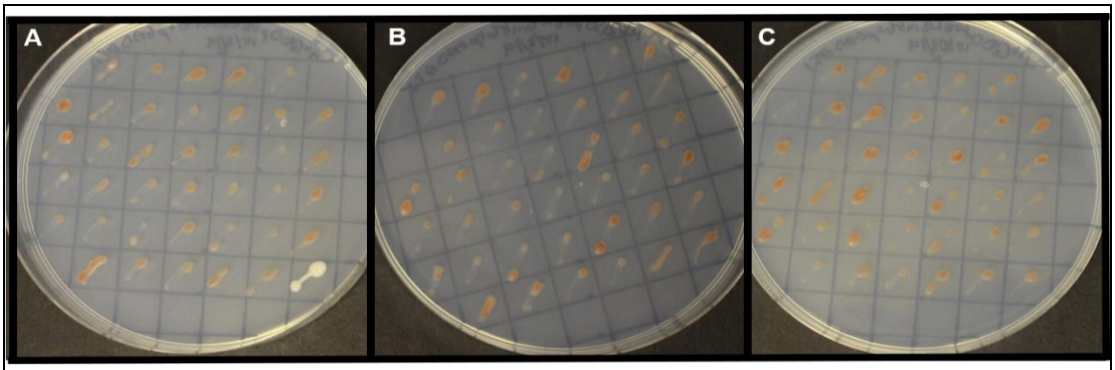


Figure 3.37: Y2H Gold yeast co-transformed with pGBKT7-Bm86 and the pGADT7-Antigen 1 (T3-T1) on quadruple drop out medium (SD -Trp/-Leu/-His/-Ade). Figure A shows pGBKT7-Bm86 and pGADT7-Antigen 1 (Full length) on QDO. Figure B shows pGBKT7-Bm86 and pGADT7-Antigen 1 (T2) on QDO. Figure C shows pGBKT7-Bm86 and pGADT7-Antigen 1 (T1).

A summary of the data is shown in Table 3.5. From the data, it is evident that the interaction across the truncations is stable, pointing towards the N-terminal region of Antigen 1 being implicated in interaction with Bm86. The interaction drops off on QDO, which may indicate that the interaction with Bm86 is transient or weak. This all remains to be further validated in future by the testing of a construct lacking the N-terminal region.

Table 3.5: Growth of transformed yeast on selective media. N/A indicates where the test is inapplicable. Blocks labelled "++++" (in dark green) indicates the highest growth, decreasing to blocks labelled with "+" (in reddish-brown) and no growth (red blocks labelled with "- ").

Plasmid combination	**Test type**	**Selective Growth Media**									
		-Trp	**-Trp/-His**			**-Leu**	**-Trp / -Leu**	**-Trp/-Leu/-His**			**-Trp/ -Leu/ -His/ - Ade**
			3-AT Concentration (mM)					**3-AT Concentration (mM)**			
			0	*2,5*	*5*			*0*	*2,5*	*5*	
pGBKT7-53	+ Control Bait	++++	++	+	-	-	N/A				
pGBKT7-53 + *pGADTT7-T*	+ Control	++++	N/A			++++	++++	++++	++++	++++	++++
pGBKT7-LAM	- Control Bait	++++	+++	++	+	-	N/A				
pGBKT7-LAM + *pGADT7-T*	- Control	++++	N/A			++++	++++	+++	++	+	-
pGBKT7-Bm86	Test Bait 1	++++	++	+	-	-	N/A				
pGBKT7-Bm86 +pGADT7-Antigen 1 Full length (T3)	Test 1	++++	N/A			++++	++++	++++	+++	+++	+
pGBKT7-Bm86 +pGADT7-Antigen 1 Truncation 1 (T1)	Test 2	++++				++++	++++	++++	+++	+++	+
pGBKT7-Bm86 +pGADT7-Antigen 1 Truncation 2 (T2)	Test 3	++++				++++	++++	++++	+++	+++	+

REFERENCES

Acuner Ozbabacan, S.E., Engin, H.B., Gursoy, A., Keskin, O., 2011. Transient protein-protein interactions. Protein Eng. Des. Sel. 24, 635–648. https://doi.org/10.1093/protein/gzr025

Bamji-Mirza, M., Yoa, Z., 2011. Phospholipases, AOCS Lipid Library. https://doi.org/10.21748/lipidlibrary.39190

Baron, S., Barrero, R.A., Black, M., Bellgard, M.I., van Dalen, E., Maritz-Olivier, C., 2018. Differentially expressed genes in response to amitraz treatment suggests a proposed model of resistance to amitraz in *R. decoloratus* ticks. Int. J. Parasitol. Drugs Drug Resist. 8, 361–371. https://doi.org/10.1016/j.ijpddr.2018.06.005

García-García, J.C., Montero, C., Redondo, M., Vargas, M., Canales, M., Boue, O., Rodríguez, M., Joglar, M., Machado, H., González, I.L., Valdés, M., Méndez, L., 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. https://doi.org/10.1016/S0264-410X(99)00548-4

Neves, S.R., Ram, P.T., Iyengar, R., 2002. G Protein Pathways. Science (80-.). 296, 1636–1639. https://doi.org/10.1126/science.1071550

Nijhof, A.M., Balk, J.A., Postigo, M., Rhebergen, A.M., Taoufik, A., Jongejan, F., 2010. Bm86 homologues and novel ATAQ proteins with multiple epidermal growth factor (EGF)-like domains from hard and soft ticks. Int. J. Parasitol. 40, 1587–1597. https://doi.org/10.1016/j.ijpara.2010.06.003

Rhee, S.G., Bae, Y.S., 1997. Regulation of Phosphoinositide-specific Phospholipase C isozymes. J. Biol. Chem. 272, 15045–15048. https://doi.org/10.1074/jbc.272.24.15045

CHAPTER 4: DISCUSSION

The complete transcriptomes for various life stages and tissues of *R. microplus*

The smallest tick genome available currently is within the same range as the largest mosquito genome, approximately 2-2.5 Gbps (Ullmann *et al*., 2005). The *R. microplus* genome is an extensive 7.1 Gbp genome with up to 70% being repetitive (Ullmann *et al*., 2005). Barrero *et al.* 2017 have published a draft genome assembly for *R. microplus,* which is only 40.1% complete, based on BUSCO analysis. However, the genome may be improved upon as methods for sequencing of longer reads become available. This will be vital in the assembly of complex repetitive regions.

In this study, the transcriptomes for *R. microplus* adult gut, ovary and salivary gland tissues, as well as the larval and nymph life stages, were assembled. These were fed on Holstein Friesian cattle. As tick larvae do not attach simultaneously to the host upon infestation (it can be up to 72 hours before attachment), ticks collected at a specific time will represent an array of feeding stages, which correspond to differences in their individual transcriptomes on a molecular level. This concept is supported by a recent publication by Perner *et al*., (2018) which indicated that even transcriptomes from single salivary glands differ. As such, the larvae and nymphs were collected and verified to be in a specific life stage prior to RNA isolation. The transcriptomes will, therefore, be representative of the genes expressed in the pool of larvae, nymphs or adult tissues collected at the specific time. Apart from the diversity within the individual transcriptomes, natural diversity within the population will also be present in the assembly. The transient nature of transcriptomes is highlighted in the transcriptome data from the USDA (kindly provided by Dr Felix Guerrero, Texas, USA). This was *de novo* assembled from an assortment of *R. microplus* life stages, tissues and an extensive list of conditions, for example; transcriptomes from ticks treated with acaricides and ticks on cattle that are prevented from feeding, to name but a few. All the latter is essential when working towards a complete transcriptome as gene expression is a highly regulated, time-dependent and stimuli-responsive process.

The assembly from the South African strain of *R. microplus* ranged from 93.2% to 95% BUSCO completeness, with the lowest percentage of missing genes recorded for larvae (0.56%) and the highest for gut tissue (2.6%). An assembly is classified as ‘complete’ when the transcript lengths are within two standard deviations of the BUSCO group mean length (Simão *et al*., 2015), and these transcriptome assemblies

are well within this range. There is also an average, across the transcriptome sets, of approximately 70% that correspond to duplicated transcripts, echoing the highly repetitive nature of the *R. microplus* genome.

As the specific repertoire of genes that are expressed at any moment is continuously fluctuating, we did not attempt comparative gene expression analyses[1] but identified specific genes and their transcript sequences, such as Bm86, Antigen 1 and members of the proposed PLC pathway. This approach, despite the experimental limitations, did provide insight into specific coding sequence variations which can now be further investigated.

Bm86 sequence variation

The Bm86 sequences found in this study cluster with the American Texas outbreak strain sequences reported by Freeman *et al.* 2010 (Figure 3.1). Most of the available online sequences used for comparison are derived from gut tissue while the specific life stage/tissue for some sequences are not reported. Despite the limitations, these sequences clustered together per geographical area. The grouping of the South African sequences with those of American origin is not surprising, given previous phylogenetic studies done by our group using non-coding gene regions (ITS2) and mitochondrial *R. microplus* genes (COI). Those field samples from South Africa grouped with American and Brazilian sequences (Oberholster, 2014, unpublished).

Historically *R. microplus* is known as the Asiatic tick; however, it was first reported to be displacing the endemic *R. decoloratus* in South Africa in the early 1960s. It was introduced with the mass importation of cattle from various regions around the globe due to massive cattle losses from disease outbreaks across the country (Oberholster, 2014, unpublished). It is thus not apparent which strains of *R. microplus* were introduced into South Africa at which time points (Oberholster, 2014, unpublished). But, all current data points towards a closer relationship with American *R. microplus* tick populations. This is a focus of future research using additional phylogenetically informative genes.

Current analyses indicated that Bm95 groups with the Bm86 sequences from America. This observation is corroborated by a recent study on the sequence variation of Bm86 in Mexico (Martínez-Arzate *et al*., 2019). Unfortunately, these Mexican sequences will

[1] This assessment of the assembled data is already underway within the research group.

not be publicly available until 2021, and so we were not able to include it in our analyses. Once these sequences are available, it would be prudent to reconstruct these trees to further corroborate the South African sequence grouping. The Thailand sequence analyses done to date made use of neighbour-joining trees that did not include Bm95 (Kaewmongkol *et al*., 2015). However, their groupings do largely coincide with those seen in the ML trees in this study.

Previously, it has been shown that the Bm86 gene is differentially expressed in all life stages as well as being found in the gut and ovary tissues, with the highest expression in the gut (Bastos *et al*., 2010). However, the expressed sequences for Bm86 in each life stage and tissue had not previously been investigated. Many common single amino acid missense mutations are evident from our data that occur in both the domain-coding parts of the sequences and the linker regions between domains (Figure 3.3). With regards to the sequences obtained, the gut and nymph sequences share 99,69% identity while the salivary gland sequence share only 90,31% identity. The latter diversity needs to be validated as only partial sequence data was obtained from the salivary gland transcriptome. Noteworthy is the observation that the sequence for the South African Bm86 is different from that of the original Yeerongpilly sequence used for the initial production of the Bm86 vaccine (Rand *et al*., 1989). Variations between the ovary and salivary gland were observed with 92.83% identity shared between the two sequences. Despite the 97.69% identity shared between the larvae and gut transcriptome sequences, closer inspection of the amino acid sequences indicate that the larvae transcriptome sequence contains 10 unique amino acid changes with only 1 of these also being present in the Yeerongpilly sequence (a glycine in position 226) which was not corroborated via Sanger sequencing. cDNA sequencing confirmed 9 of the unique amino acid changes (Figure 3.7).

Different alleles for Bm86 in *R. microplus* have been reported (García-García *et al*., 2000; Nijhof *et al*., 2009; Sossai *et al*., 2005) and may offer a possible explanation for the heterozygosity detected during Sanger sequencing data of Bm86 amplified from larval cDNA. However, the possibility of these nucleotide changes corresponding to standard variation within the strain cannot be excluded. As such additional sequencing from this South African strain as well as South African field samples will be vital before conclusions can be drawn.

The variation detected in larvae is of particular interest when considering the proposed function of Bm86 in tick feeding (Bastos *et al*., 2010). When *R. microplus* larvae attach to the softer, thinner skin regions of the host (due to their relatively shorter mouthparts),

it is likely that the bovine epidermis is not fully penetrated (Moorhouse and Tatchell, 1966). During the initial stages of feeding, larvae attach approximately 2 times every 8 hours and spend at least half their time attached (Roberts, 1971). Seifert *et al.* in 1968 found that it was possible to analyse the dietary content of each life stage of *R. microplus* and indeed found that the proportion of red blood cells in a tick increases with each life stage and time spent feeding. Initially, the meal is composed mainly of plasma and in the larval life-stage would appear to be made up almost entirely of plasma. The same group also noted that ticks may initially feed on tissue fluids, and that skin cells may contain extravascular plasma or lymph before capillary permeability is possible. In 1975, Kemp *et al*. (1975) found that it is possible to feed larvae to engorgement on a bovine serum alone and that the amount of protein present was not limiting (although some protein is necessary, for larval growth and development). Trentelman *et al.* (2017) found that it is possible to feed *R. australis* larvae to engorgement using artificial feeding. When larvae were fed on serum from Bm86 vaccinated cattle, it resulted in a 47% reduction in larval engorgement; emphasising that the Bm86 vaccination of cattle has a limited effect on larvae. Based on our hypothesis that Bm86 acts as a signalling molecule, which is activated in response to a specific signal, we propose that the changes observed in the N-terminal 240 amino acids may be involved in binding of unique signalling molecules, which would be a factor present in plasma/lymph in larvae but a different factor in nymphs and adult life stage due to the different content of the bloodmeal.

Differences in the sequences of Bm86 have been proposed as a significant driver for the variation observed with Bm86 protection after vaccination in different geographical areas (García-García *et al*., 1999). As even a single amino acid change in an antigen can drastically reduce the efficacy of a vaccine, understanding the diversity of an antigen remains vital for vaccine production. An additional example of the latter is a malaria vaccine, where it was found that when the antigen sequence slightly differed to that of the native antigen target found in a particular region, a reduction in vaccine efficacy from 50% to less than 35% resulted (Leach *et al*., 2015).

The molecular mechanism giving rise to the extensive sequence variation within Bm86 remains to be elucidated. García-García *et al*. (1999) studied variation in Bm86 from cDNA and referred to their findings as variation from a single locus, which remains to be validated once a genome is available. Variation resulting from multiple copies of the Bm86 gene, given the repetitive nature of the *R. microplus* genome, remains to be investigated. The possibility of more than one gene can be analysed via Southern blotting. A vast number of additional mechanisms that could give rise to sequence

variation, such as RNA editing (i.e. C→U or A→I deamination), alternative splicing, differential use of polyadenylation sites and the role of small regulatory RNAs may also provide additional insights (Figure 4.1).

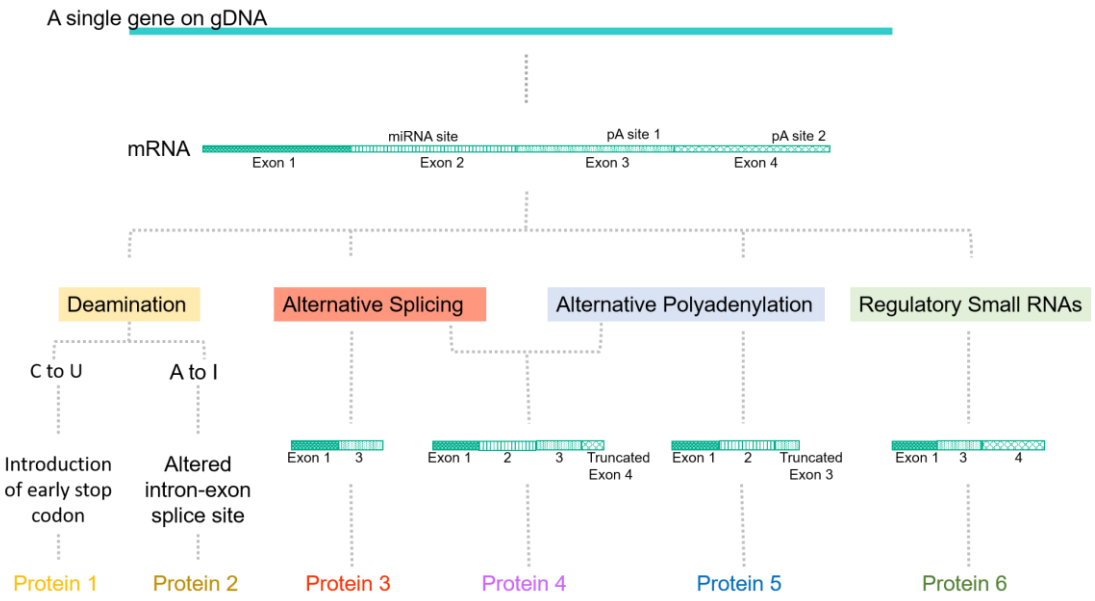


Figure 4.1: Possible mechanisms of RNA editing giving rise to various forms of a protein from a single gene. Adapted from (Latchman, 2015).

The role of alternative splicing is well-known in generating diversity in all eukaryotic organism. It has also been suggested as a significant driving factor in generating the diversity of Acetylcholine esterase (AChE) in *R. microplus* and other tick species (Reviewed by Lees and Bowman, 2007) but remains to be validated. The ovary and salivary gland sequences for Bm86 in our current data appear to lack a signal peptide. Two possible molecular reasons for this could be alternative splicing or the use of different polyadenylation signals (Figure 4.1), which is the typical mechanism used in creating secreted and membrane-bound immunoglobulins (Latchman, 2015). Truncation of Bm86 sequences has also been reported for Mozambique samples that displayed one full length and one isoform that lacks 22 amino acids Nijhof *et al*. (2009). Bm95 lacking the GPI anchor has been reported by García-García *et al*. (2000). In this case, there was a T to G change which resulted in an amber stop codon and subsequent formation of a truncated protein lacking both the customary C-terminal and GPI anchor.

The larval sequences contain evidence for 10 transversions, each occurring on the first nucleotide position of a codon resulting in the formation of a missense mutation (Figure

3.7 and Table 3.3). In one position (amino acid 237) a transition from C to T was observed, which is most likely due to RNA editing (deamination of a C to U on the mRNA, which upon sequencing is read as a T). On an additional 4 sites, transitions of A to G give evidence of RNA editing (A to I deamination). As such, we propose RNA editing as a driving factor in generating Bm86 diversity (Table 3.3). The diversity in Bm86 highlights (a) the need for tailored vaccines for a region or (b) targeting more than one life stage / tissue in a tick based on the unique sequences expressed. As such, it is prudent to conduct country-wide and life stage / tissue-specific sequence analyses of Bm86 in South Africa.

Antigen 1 sequence variation and mapping of its interaction region with Bm86

Antigen 1 is a novel *R. microplus* protein discovered by this research group that contains two BPTI/Kunitz binding domains. In an initial yeast two-hybrid study, it was found to have potential protein-protein interactions with Bm86 (Kiper 2013, unpublished) and antigenic properties, making it attractive for inclusion as a vaccine antigen. BPTI/Kunitz domains are known as protease inhibitors with a wide array of functions in several metabolic pathways. In ticks, Kunitz proteins are well known as anti-haemostatics (Maritz-Olivier *et al*., 2007), but an array of additional functions can also be ascribed to Kunitz-domain containing proteins. Based on transcriptome data, Antigen 1 is predicted to be a secreted protein. Its biological function, life stage and tissue distribution, as well as sequence diversity, remains unknown.

In this study, truncated sequences were obtained for Antigen 1 from the transcriptome assemblies of both larvae and gut tissue. However, the full-length open reading frame of Antigen 1 was since amplified from cDNA and validated via Sanger sequencing (Bishop, 2018, unpublished data). This stresses the importance of validating predicted sequences from transcriptome data. Comparisons between the Sanger sequence and all the transcriptome sequences showed changes in amino acids in the N-terminal of Antigen 1. No changes were observed in the predicted Kunitz domains.

The variation in the N-terminal region of Antigen-1 is particularly interesting when viewed in conjunction with the yeast-two hybrid findings of this study; where full-length and truncated versions of Antigen-1 were evaluated for binding to Bm86. All constructs successfully interacted with Bm86, pointing towards the N-terminal 219 amino acids of Antigen-1 binding to Bm86 (as removal of the Kunitz domains did not disrupt interaction). As Bm86 is a GPI-anchored protein (at the C-terminal), we propose that the interaction may be between the N-terminal region of Bm86 (which lacks EGF

domains and is highly variable) and the N-terminal of Antigen-1 (Figure 4.2). This hypothesis will be further tested by means of additional deletion studies. In summary, we propose that Antigen 1 is interacting with Bm86, but upon recognition of a cellular signal, it dissociates from Bm86. This will allow Bm86 dimerization and subsequent signalling via the PLC pathway, while Antigen 1 is released into the extracellular environment to fulfil other biological roles.

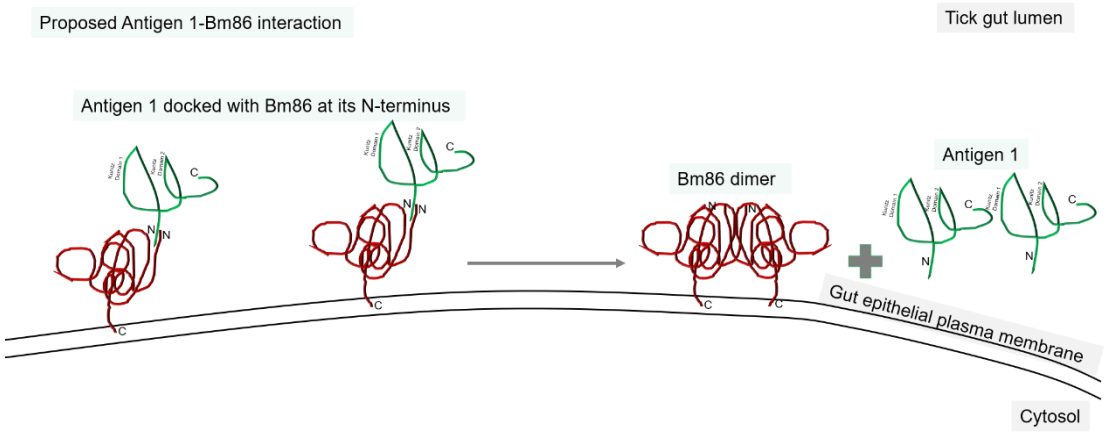


Figure 4.2: Proposed interaction site of Antigen 1 with Bm86 and proposed mechanism of interaction. It is proposed that (1) in the native state the antigen 1 (a double loop conformation predicted based on two Kunitz-like domains) is bound and interacting with Bm86, but upon some signal dissociates from Bm86, (2) allowing Bm86 to dimerize on the plasma membrane while Antigen 1 is released to fulfil other biological roles.

Despite the inherent limitations of the study, it is interesting that the regions of sequence variation in both Antigen 1 and Bm86 correspond to the predicted regions of interactions. Further understanding of the essential interacting regions of Antigen 1 and Bm86 will allow for the design of novel vaccines and/or therapeutics capable of interrupting this interaction. Our current findings (based on the pink-red phenotype of yeast colonies on QDO media, Figure 3.37) point towards a transient interaction between Antigen-1 and Bm86. Validation of our current findings using different biochemical technologies are underway, such as isothermal titration and biocore.

Identified Heterotrimeric G proteins and Phospholipase C sequences

Sequences for heterotrimeric G proteins in the gut were identified. Some Phospholipase C coding sequences were identified in the gut transcriptome, while additional sequences could only be detected in the total transcriptome. This points

towards the expression of these PLCs in other tick tissues and or life stages. In this study, we only focussed on the sequences from adult gut tissue.

Heterotrimeric G proteins

Heterotrimeric G proteins typically consist of three subunits, namely α, β and γ. The α subunit binds and hydrolyses GTP while the β and γ subunits form a dimer known as βγ (Reviewed by Neer 1995). Heterotrimeric G proteins often function as signal transduction molecules, communicating signals from membrane receptors to intracellular effectors (Neves *et al*., 2002). The families of heterotrimeric G proteins are divided based on the homology of their α subunit. The α subunit provides specificity for the proteins' receptor and effector combination and is usually implicated in the activation of second messengers involved in the signalling cascade. For a list of G proteins found in arthropods to date and their respective interaction partners, please refer to Table 1.3 (Chapter 1).

In the phylogenetic tree for the G protein sequences (extracted from the gut RNA transcriptome and sequences from the NCBI and UniProt) it is evident that sequences for each G protein subunit as well as the i, o, q and s families are present in *R. microplus*. Each of the *R. microplus* sequences also groups with the expected subunit (Figure 3.8). In all cases, the domain architecture identified the same functional domains as those present in the reference sequences, which corroborate the presence of functional G proteins in the gut of *R. microplus*.

In Figure 3.9 and 3.10, the sequence for the $G_{\alpha i}$ protein from the *R. microplus* gut transcriptome is compared with that of $G_{\alpha i}$ found in *D. melanogaster.* As noted in Table 1.3, $G_{\alpha i}$ in arthropods, such as *Drosophila*, is involved in adenylate cyclase modulation in neuronal cell division, differentiation and interacts with $G_{\beta 1}$, Loco, Rapsynoid and the G protein-coupled receptor Moody (Granderath *et al*., 1999; Yu *et al*., 2005). The domains mediating each of these functions of $G_{\alpha i}$ are seen in the domain architectures in this study (Figure 3.9). The only domain that was not detected in the *R. microplus* sequence is the GTP/Mg^{2+} binding site. This missing site may not have been picked up by the domain prediction software due to the missing amino acids which most likely is an artefact of the assembly. Further sequencing is, therefore, necessary to confirm the full coding sequence and domain architecture of the $G_{\alpha i}$ protein.

Sequence data for the $G_{\alpha o}$ from the *R. microplus* gut transcriptome was compared with that of $G_{\alpha o}$ found in *D. melanogaster* (Figures 3.11 and 3.12). $G_{\alpha o}$ is involved in the

Wnt/frizzled and planar/frizzled pathway transduction as well as adenylate cyclase modulation in neuronal cell division and differentiation and interacts with $G_{\beta 1}$; $G_{\gamma 1}$; Frizzled; Axin; Loco; Rapsynoid; GPCR Moody and PLC_{β} 21C (which is discussed later in this chapter) (Dahdal *et al*., 2010; Egger-Adam and Katanaev, 2010; Katanaev *et al*., 2005; Katanayeva *et al*., 2010). The *R. microplus* $G_{\alpha o}$ has all the sites and domains found in the reference $G_{\alpha i}$ with the addition of a GEM1 Domain, which is part of the Ras superfamily. Proteins that contain the GEM1 domain are GTPase proteins that have been found in the ERMES (ER-mitochondria encounter structure) tethering complex and play a role in the endoplasmic reticulum-mitochondrial exchange of phospholipids (Kornmann *et al*., 2011).

In Figures 3.13 and 3.14, the sequence for $G_{\alpha q}$ from the *R. microplus* gut transcriptome is compared with that of $G_{\alpha q}$ found in *D. melanogaster.* $G_{\alpha q}$ in arthropods is involved in the activation of PLC_{β} and visual and olfactory transduction and is noted to interact with $G_{\beta 2}$; PLC_{β} 21C; retinal degeneration A and Frazzled (Elia *et al*., 2005; Hardie *et al*., 2002; Hiramoto and Hiromi, 2006; Kain *et al*., 2008).

In Figures 3.15 and 3.16, the sequence for $G_{\alpha s}$ from the *R. microplus* gut transcriptome is compared with that of $G_{\alpha s}$ found in *D. melanogaster*. $G_{\alpha s}$ in arthropods is involved in adenylate cyclase activation, the mediation of CHIP/LBD complexes in the wing, sensory brittle development and the dopamine receptor signalling pathway. It interacts with $G_{\beta 1}$; $G_{\gamma 1}$; Fasciclin-2; Dunc; CHIP and Ribosomal proteins S6,13, L26 in *Drosophila* (Bronstein *et al*., 2010; Dahdal *et al*., 2010; Giot *et al*., 2003; Katanayeva *et al*., 2010; Wolfgang *et al*., 2004). With regards to the alpha subunits for G_f (Bausek and Zeidler, 2014; Quan *et al*., 1993) and Concertina (Fuse *et al*., 2013; Giot *et al*., 2003; Nikolaidou and Barrett, 2004) these were not detected in the gut transcriptome of *R. microplus*. These subfamilies of G proteins have to date only been found in *Drosophila* (see Table 1.3, Chapter 1) and are likely unique to this species.

All G_{α} subunits must function in concert with a G_{β} and a G_{γ} subunit. In the *R. microplus* gut transcriptome assembly, one of each of the latter was identified and shown to group with the corresponding families (Figure 3.8).

Comparison of the sequence for G_{γ} from the *R. microplus* gut transcriptome with the G_{γ} of *D. melanogaster* indicated that despite differences on sequence level, the overall domain architecture of the G_{γ} proteins appears conserved. Similarly, comparison of the sequence for G_{β} from the *R. microplus* gut transcriptome with the G_{β} found in *D. melanogaster*, we found that all the expected domains and binding sites of a G_{β} subunit

are present. The sequences identified here are the only $G_{\gamma/\beta}$ subunits found in the gut transcriptome assembly.

As noted in Table 1.2, $G_{\beta/\gamma}$ in arthropods are involved in all processes that require a G_{α} subunit. Together they ligand-specifically activate effectors, such as PLC, PLA2, PI3K and more (Carty *et al.* 1990; Linder *et al.* 1990; Reviewed by Kaziro *et al.* 1991 and Neer 1995).These proteins are integral components of all G proteins and as such are necessary for any G protein functions and thus expected to be implicated in our hypothesis.

The identification of G proteins in this study verifies the presence of heterotrimeric G proteins and for the first time reports on the various subunits present in the *R. microplus* gut. It also opens the window to design specific future experiments, with specific subunits in mind, to verify our current model of a Bm86 signalling pathway.

Heterotrimeric G proteins, as they are involved in multiple signalling pathways, are also promising drug targets for signal interceptor drugs that could be the next generation of tick control therapeutics. As compounds that act on G proteins have different forms of selectivity (Höller *et al*., 1999), the possibility of designing tick-specific therapeutics remains a viable option. G proteins and specifically the α-subunit, as drug targets have been proposed previously, mainly for the treatment of human parasites (Chahdi *et al*., 1998; Freissmuth *et al*., 1999; Höller *et al*., 1999; Ja *et al*., 2006; Kimple *et al*., 2011).

Two mechanisms of signal inhibition via G protein antagonists (not by receptor antagonists) of interest to tick control are shown in Figure 4.3. While it would be possible to block the GTP binding pockets, this method would not be feasible in tick control as GTP binding pockets are highly conserved even across species and the risk of off-target effects in the host at this site is too high (Freissmuth *et al*., 1999). However, there are two remaining options. Figure 4.3 A indicates that the receptor-G protein interface can be targeted; in nature, insect venoms have been seen to act in this way. One compound, Mastoparan (Wasp venom), evolved to activate $G_{\alpha i/o}$ but can be modified to inhibit $G_{\alpha s}$ proteins instead (Freissmuth *et al*., 1999). This indicates that it may be possible to synthesise or extract similar compounds to agonise or antagonise G proteins with low molecular weight compounds (Chahdi *et al*., 1998; Höller *et al*., 1999).

G protein inhibitors also include receptor derived and related peptides as well as G protein-derived peptides. These derivatives can cause the uncoupling of the receptor

and G proteins if they, for example, encompass the amino acid residues implicated in the receptor-G protein interaction (Rasenick *et al*., 1994). In addition, some non-peptide antagonists of G proteins can also be employed, such as Suramin that is used to treat *Trypanosoma spp.* and *Onchocerca volvulus,* which causes African sleeping disease and river blindness, respectively (Voogd *et al*., 1993). Suramin acts by suppressing the release of GDP from the G_{α}-subunit (the rate-limiting step in G protein functionality) (Freissmuth *et al*., 1999). Another non-peptide example is Lithium, which has been noted to have antimanic and antidepressant effects, also acts biologically to directly inhibit G proteins (Avissar *et al*., 1988). While compounds that can affect the G protein-effector interaction site remain elusive, especially considering that membrane permeability would be essential, it is not unreasonable to consider this site as a potential target as well.

These examples illustrate that it is possible to design novel drugs which can act at select sites or target selected G proteins in a stimulatory or inhibitory manner. It would be of interest to investigate this path for the control of ticks. G proteins and PLCs are essential in a myriad of pathways in arthropods (see Table 1.3 and 1.4). If the G proteins in ticks were inhibited, then all the pathways involving these proteins would be halted. Having so many cellular processes affected are unlikely to be circumvented, and tick survival in the face of such a drug would be minimal.

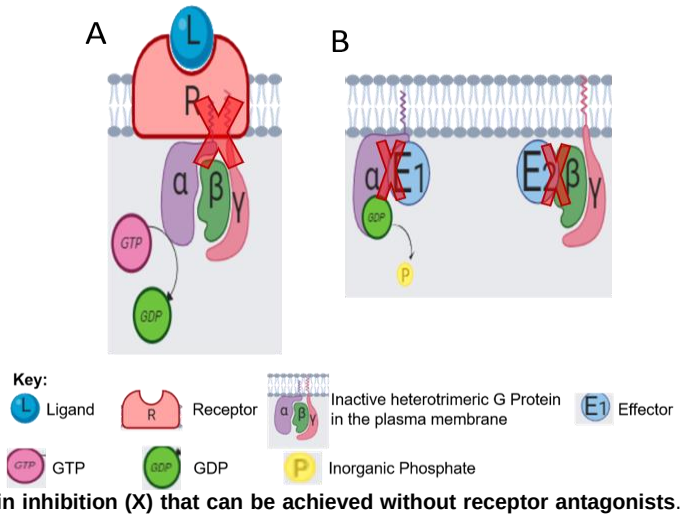


Figure 4.3: G protein inhibition (X) that can be achieved without receptor antagonists. In A the inhibition of receptor and G-protein coupling, either general or class-specific, is shown. In B the inhibition of G protein complex subunit-effector coupling and/or dissociation is indicated

L: Ligand; E: Effector; R: Receptor; αβγ: G protein subunits. Adapted from Freissmuth *et al.* (1999) using Bio Render (https://biorender.com).

Phospholipase C proteins

Phospholipase C refers to a class of multidomain, soluble protein families under cell surface receptor control. This group of effector proteins characteristically consists of a shared set of protein domains arranged around an X- and a Y-box region that forms a catalytic α/β-barrel, (Williams and Katan, 1996). The pleckstrin homology domains (PH) act as membrane tethering devices which mediate regulatory signals to the PLC, for example via coupling with $G_{\beta\gamma}$ and facilitating the binding of PLC to a phospholipid (Wang *et al*., 2000). The X and Y structural domains are responsible for the catalytic activity of PLC while the EF-hand motif (EF) is a helix-turn-helix structural domain that has a role in Ca^{2+} binding. The Ca^{2+}-dependent phospholipid-binding domain (C2) may possess multiple Ca^{2+} binding sites (Bamji-Mirza and Yoa, 2011; Suh *et al*., 2008).

Some members of the family also contain unique identifier domains. PLC_{γ} contains Src homology (SH) 2, Src homology 3, and a split PH domain which is involved in protein-protein interactions (Bamji-Mirza and Yoa, 2011). The simplest PLC family, found in prokaryotes, consist only of the catalytic α/β-barrel (Heinz *et al*., 1996). The largest family, PLC_{ε}, contains two more protein domains than any other known PLC protein family (Shibatohge *et al*., 1998).

In Figure 3.21, a phylogenetic tree of identified PLC sequences from the *R. microplus* transcriptome as well as reviewed sequences and putative Acari PLC sequences is illustrated. Phospholipase C_{γ} and a possible PLC_{β}-21C were identified in the gut while possible PLC_{ε}, $PLC_{\beta 4}$ and a $PLC_{\delta/\eta}$ were identified in the total transcriptome, indicating that they are present in other tick tissues and/or life sages but not in the gut. However, this needs to be validated. Each family of PLC found will be discussed individually.

No sequences were identified in the *R. microplus* transcriptome for PLC_{ζ}. This PLC has only been found in testes to date and as such may only be present in male ticks. (Bamji-Mirza and Yoa, 2011; Saunders *et al*., 2002). As only female ticks were used in this study, this hypothesis needs to be validated for *R. microplus.* PLC_{ζ} plays a role in fertilisation and is the smallest known mammalian PLC (Saunders *et al*., 2002).

PLC_{ε} has been identified in arthropods, but to date, no reviewed sequence for this PLC from the arthropods occur in public databases. Phospholipase C_{ε} is the largest known PLC family member and the only PLC that contains a RAS GEF domain and two RA domains which facilitate interaction with the Ras family small G proteins (Suh *et al*., 2008). In *C. elegans*, PLC_{ε} has a role in the control of ovulation and the regulation of epidermal morphogenesis (Kariya *et al*., 2004; Vázquez-Manrique *et al*., 2008). In humans PLC_{ε} has the highest expression in the colon and endometrium and is

associated with nephrotic syndrome type 3. The role of this PLC in ticks is not known, and further studies are needed to elucidate this.

The Phospholipase C_ε sequence identified from the transcriptome appears to have a FERM domain, which is unique. The FERM domain contains a PH-fold subdomain and is found in several cytoskeletal-associated proteins that are localised to the plasma membrane and cytoskeleton interface. FERM domains have a role in regulating the binding of PIP_2 (Hamada *et al*., 2000). The presence of the domain may be as a result of the incorrect assembly of some regions or may be a unique domain in the *R. microplus* PLC_ε. Also, it was not possible to extract a full-length sequence for this protein as the transcript was not fully sequenced, additional sequencing for this protein will need to be done to confirm the presence of the FERM domain and the length of the protein to fully identify the sequence as truly being PLC_ε.

This is the first time a PLC that has a putative match to either a PLC_δ or a PLC_η has been described in arthropods. PLC_δ and PLC_η are the most similar of any two PLC family members as they contain the same domains and simple organisation. PLC_δ is the smallest of the two and is considered the fundamental PLC, while PLC_η has a longer C terminal sequence without any additional known domains (Suh *et al*., 2008). PLC_δ contains two EF-hand domains, while PLC_η has three. Also, the X and Y domains are closer together in PLC_δ than in PLC_η. In humans, these two proteins are found predominantly in the brain and testis with PLC_η also being found in the eye and pancreas (Bamji-Mirza and Yoa, 2011). PLC_η may have a role in neural network maintenance and localised to the plasma membrane without extracellular stimuli but are activated by G protein-coupled receptor (GPCR) stimulation. PLC_δ has a role in the cell cycle, skin homeostasis, placental development, is the most sensitive to Ca^{2+} and associated directly with G_α (Bamji-Mirza and Yoa, 2011; Cockcroft, 2006; Nakamura *et al*., 2005; Suh *et al*., 2008). The sequence extracted for these proteins from the RNA transcriptome of *R. microplus* is not full length, being only 437 amino acids long. It was not possible to differentiate which PLC this may be as it contains only an X-Y box and a C2 domain, however it may be a PlC_δ based on the proximity of the X and Y region to each other and the lack of an extended C terminal region. Further sequencing will need to be done to confirm the length and full domain complement of this sequence.

In general, all PLCs have the usual PLC PH and EF domains followed by the X domain part of the catalytic X-Y box. However, in all PLC_γ identified to date, the X-Y box contains four additional internal domains, namely two SH2 and one SH3 domain which

is flanked by a split PH domain, this architecture is unique to PLC_{γ}. The Y domain is then followed by a C2 domain as with all other PLC proteins.

The unique catalytic region of PLC_{γ} has been associated with the activation of this PLC downstream of tyrosine kinase activity associated receptors and GPCRs. Specifically, $PLC_{\gamma\text{-}1}$ has been seen to be activated in response to polypeptide growth factors, usually containing EGF domains and which bind to receptor tyrosine kinases. It is interesting to recall here that Bm86 contains multiple EGF domains. The SH2 domain mediates binding to phosphorylated tyrosine residues within the receptor (Kamat and Carpenter, 1997) while the SH2, SH3 and split PH domains have all been implicated in the protein-protein interactions of this PLC. $PLC_{\gamma\text{-}1}$ plays a role in cell migration, proliferation and the regulation of some ion channels (Suh *et al*., 2008). In the *R. microplus* transcriptome, a PLC_{γ} has been identified. Although it was not possible to extract the full sequence (it lacks the N-terminal two domains) the protein does contain the unmistakable PLC_{γ} domain structure. Given that this PLC is known to interact directly with EGF domain-containing receptors as well as GPCRs, and that Bm86 is an EGF domain-containing protein spanning the cell membrane, and that this PLC is found in the gut, it is possible that the PLC involved in the proposed signalling pathway for Bm86 could be this PLC_{γ}, with or without a mediating G protein. This will only be corroborated with further studies.

PLC_{β} is a large family of PLC proteins that includes $PLC_{\beta 1\text{-}4}$. $PLC_{\beta 4}$ was the PLC_{β} clade where the *R. microplus* transcriptome PLC_{β} grouped and is also the clade into which the PLC_{β} NorpA from *D. melanogaster* groups. As such, this PLC_{β} family is used for comparison here. PLC_{β} all have a PLC_{β} specific PH and EF-hand domains and follow the usual PLC domain organisation with the PH, EF-hand and X/Y catalytic region domains followed by a C2 domain, however, the PLC_{β} group has an additional PLC_{β} C domain at the C terminal end of the proteins. The $PLC_{\beta 4}$ also contains a DUF1 domain between the C2 and PLC_{β} C terminal domain. PLC_{β} functions are activated in response to receptors from the rhodopsin superfamily of transmembrane proteins; this superfamily of receptors often contains several transmembrane spanning segments. The $PLC_{\beta 4}$ respond to G protein subunits and can also be activated by phosphatidic acid and drugs such as azacytidine, which is a DNA methyltransferase inhibitor currently approved for the treatment of Myelodysplastic Syndrome and is under investigation as a treatment for other haematological diseases. The C-terminal region containing the unique PLC_{β} domain is required for localisation of this protein to the cytoplasm and without this section acts in the nucleus. As shown in Figure 3.25, this region is missing, and as such, we propose that in *R. microplus,* this protein may act

in the nucleus. However, it was not possible to extract the full-length sequence of the protein from the transcriptome, and this is a possible reason that the C-terminal PLC_{β} domain appears to be missing. Further sequencing is necessary to confirm this result.

In *Drosophila*, there are at least two types of NorpA (No receptor potential A) proteins. The first is a PLC_{β} involved in phototransduction in the retina that interacts with transient receptor potential (TRP) via a scaffold protein, inactivation-no-afterpotential D (INAD), for gating of the TRP channel in photoreceptor cells (Bloomquist *et al*., 1988; Shieh *et al*., 1997). The transcriptome of the eye has not been investigated, and the presence of this form of the protein is unknown. The second form is a splice variant of the former that is only expressed in the fly body and thus may function in other signalling pathways (Kim *et al*., 1995). It is likely the latter that has been detected in the gut transcriptome assembly here and may be involved in PLC signalling pathways in the gut. However, this hypothesis remains to be confirmed.

PLC_{β}-21C is considered a $PLC_{\beta\text{-}2}$, and like $PLC_{\beta\text{-}2}$, does not contain the DUF1 domain that other PLC_{β} family members contain. However, in Figure 3.21, PLC_{β}-21C clusters as a separate group within the PLC_{β} family and the EF-hand domain is identified in a domain search as specifically a 21C EF-hand domain. PLC_{β}-21C is involved in olfactory transduction in *Drosophila* and responds to the $G_{\alpha q}$ subunit, which is activated by the heterodimerization of receptors (Kain *et al*., 2008). PLC_{β}-21C also responds to $G_{\alpha s}$ and $G_{\alpha o}$ in a pathway regulating behavioural rhythms in *Drosophila* (Dahdal *et al*., 2010). The role of this PLC_{β} in *R. microplus* is yet to be elucidated but may prove to have a role in similar pathways as that in *Drosophila*.

Should the proteins identified in this study be further corroborated, then the initially proposed hypothesis for the mechanism of Bm86 would have been further corroborated and, with the new understanding of the Bm86 and Antigen 1 interaction, can be updated as illustrated in Figure 4.4 below. These findings are particularly important in the design of tick control strategies such as vaccine and acaricide development; as they provide insight into the mechanism of two vaccine antigens as well as identify various potential novel drug targets in the biology of *R. microplus*.

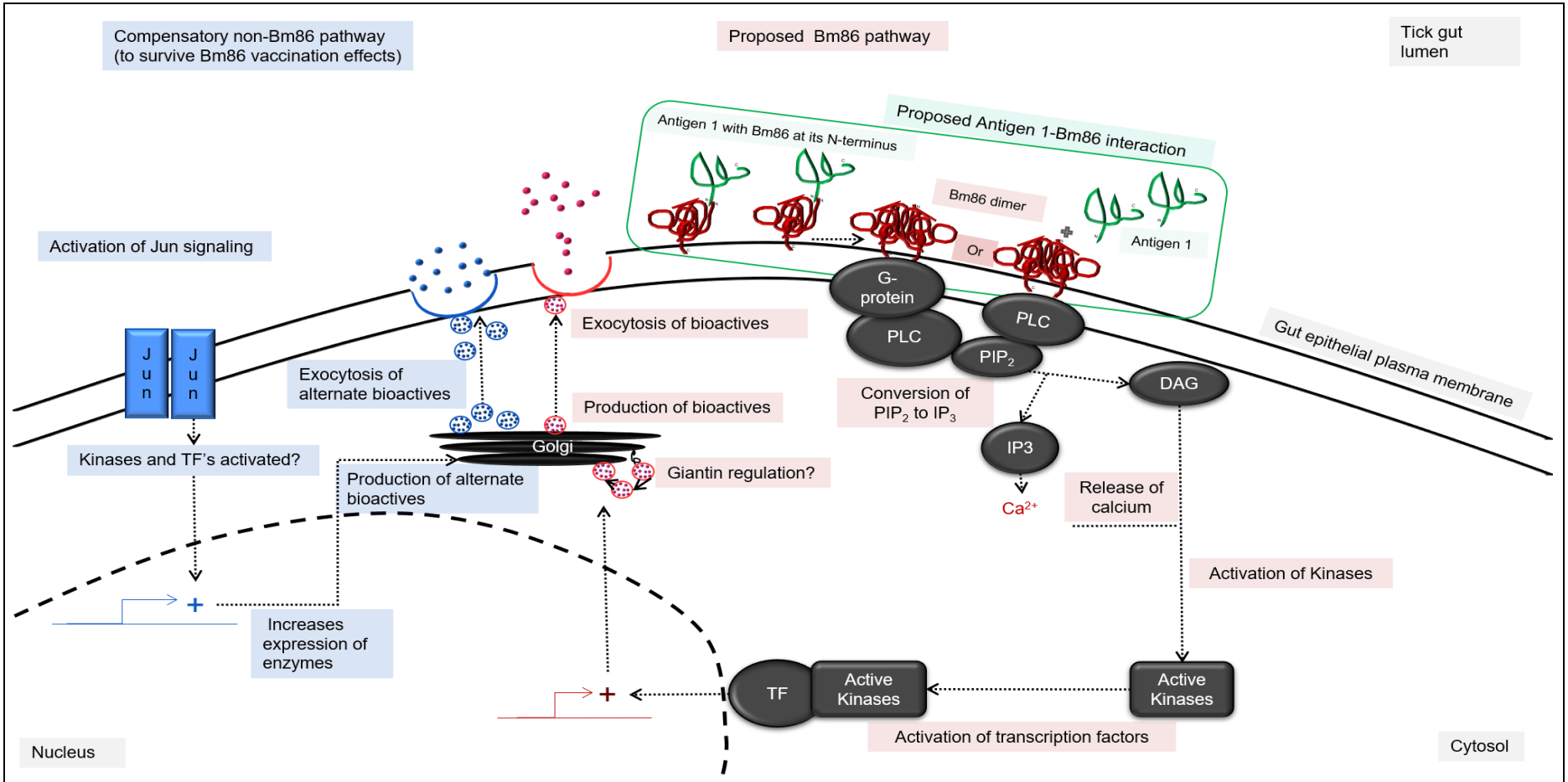


Figure 4.4: Updated proposed mechanism of Bm86 signalling. In its native state Bm86 interacts with Antigen 1 as illustrated in green. Signalling is initiated in red, beginning with a ligand binding which triggers dimerization of Bm86, leading to the production of IP_3 and release of Ca^{2+} ions which activate kinases, which in turn activate transcription factors to initiate transcription of bioactive molecules. This normal pathway is proposed to be inhibited by antibodies imbibed by the tick when feeding on a Bm86 vaccinated animal. To survive it is proposed that ticks may employ a Jun/Jun or Jun/Fos compensatory signalling pathway, leading to the increased production and exocytosis of Serine carboxypeptidases and other secreted peptides as illustrated in blue.

CONCLUDING SUMMARY

Ticks are considered one of the most economically significant parasites in the global livestock industry. The ectoparasitic hematophagous arthropods cause adverse effects through both direct (i.e. feeding habits) and indirect (i.e. disease transmission) mechanisms. Traditional tick control methods, such as acaricides, have not proven consistently effective, nor are they sustainable in the long term. However, vector directed vaccines are a promising method of tick control. To date only one protective antigen, Bm86, has been commercialised; however, it has variable efficacy, and despite a previous study which proposed a possible role for Bm86 in Phospholipase C signalling (PLC), its biological role was unclear. Antigen 1, a new vaccine candidate, has been identified to interact with Bm86, but the region of protein interaction with Bm86 was unknown.

This study had two aims. Firstly, we aimed to identify sequences for Bm86, Antigen 1 and members of the PLC pathway from various life stages and tissues of a South African strain of *R. microplus*. Secondly, we aimed to map the protein interaction domain of Antigen 1 with Bm86.

A total of five novel *R. microplus* transcriptomes were assembled. The transcriptomes for *R. microplus* larvae and nymph and three adult tissues (salivary gland, gut and ovary) were successfully assembled with an overall 94% BUSCO completeness when compared to the arthropod dataset. From these assembled transcriptomes the RNA and protein sequence for Bm86, Antigen 1, each G protein and PLC proteins were identified.

The sequence for Bm86 from each transcript set was identified. Despite numerous missense mutations in the sequences when compared to a reference sequence, most domains previously predicted in Bm86 were corroborated. In addition, all Bm86 sequences from these transcriptome assemblies' group with the American clade when compared in a Maximum Likelihood tree. Of interest, the sequence for Bm86 identified from the larval transcriptome differs significantly from other Bm86 sequences from different life stages and tissues. This may be an indication of the different blood meal component accessed by larva; pointing to a role for Bm86 in response to different ligands in the larval meal. The inclusion of this novel sequence in the current Bm86 vaccine is currently being evaluated in cattle vaccine trials (invention disclosure filed and patent pending).

The sequence for Antigen 1 that has been identified is conserved between life stages and tissues, and the domain regions are highly conserved. Five missense mutations were identified in the N-terminal region of the Antigen 1 protein, which preceded all domains. It is also this variable region that is proposed to interact with Bm86 based on the yeast-two-hybrid results of this study.

The G proteins identified in the gut transcriptome include the $G_{\alpha i/o/s/q}$ as well as a G_{γ} and G_{β} subunit. The G protein that may be involved in the proposed Bm86 signalling pathway depends on the PLC protein employed. As numerous G proteins were identified in the gut and given that they all have the capacity to activate various PLCs, it is not possible to predict which G protein is utilised in the proposed pathway. It is also interesting to note that G proteins are themselves feasible drug targets.

Phospholipase C proteins identified include $PLC_{\beta, \gamma, \delta, \eta}$ and $_{\varepsilon}$. Only PLC_{β}-21C and PLC_{γ} were identified in the gut transcriptome assembly. Either PLC may respond to Bm86 in the gut; however, we proposed that the PLC involved in the hypothesised Bm86-mediated signalling pathway, is PLC_{γ}. While a G protein is often implicated in the activation of a PLC, PLC_{γ} has been seen to be directly activated by proteins with similar domain architectures to Bm86 (i.e. EGF domains) (Kamat and Carpenter, 1997). It is thus possible that a G protein may not be necessary for the activation of the proposed Bm86 signalling pathway if the pathway involves a PLC_{γ}.

It is vital to corroborate these findings with further studies; including (but not limited to) Sanger sequencing to confirm the various sequences found, X-ray crystallography to elucidate the protein structures and interactions, microscopy to confirm cellular localisation of the identified proteins and *in vivo* functional assays to confirm protein functionality in the biology of *R. microplus*.

This is the first study to identify Bm86 and Antigen 1 sequence variation within a tick and between life stages in South Africa, as well as identify a possible region of Antigen 1 interaction with Bm86. It is also the first study to identify putative PLC and G protein transcripts in *R. microplus*. These findings are particularly important in the design of tick control strategies such as vaccine and acaricide development; providing insight into the sequence variation and biological mechanism of two vaccine antigens as well as identifying potential novel drug targets in the biology of *R. microplus* for inclusion in future control strategies.

REFERENCES

Avissar, S., Schreiber, G., Danon, A., Belmaker, R.H., 1988. Lithium inhibits adrenergic and cholinergic increases in GTP binding in rat cortex. Nature 331, 440–442.

Bamji-Mirza, M., Yoa, Z., 2011. Phospholipases, AOCS Lipid Library. https://doi.org/10.21748/lipidlibrary.39190

Barrero, R.A., Guerrero, F.D., Black, M., McCooke, J., Chapman, B., Schilkey, F., Pérez de León, A.A., Miller, R.J., Bruns, S., Dobry, J., Mikhaylenko, G., Stormo, K., Bell, C., Tao, Q., Bogden, R., Moolhuijzen, P.M., Hunter, A., Bellgard, M.I., 2017. Gene-enriched draft genome of the cattle tick *Rhipicephalus microplus* : assembly by the hybrid Pacific Biosciences/Illumina approach enabled analysis of the highly repetitive genome. Int. J. Parasitol. https://doi.org/10.1016/j.ijpara.2017.03.007

Bastos, R.G., Ueti, M.W., Knowles, D.P., Scoles, G.A., 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. https://doi.org/10.1186/1756-3305-3-111

Bausek, N., Zeidler, M.P., 2014. G 73B is a downstream effector of JAK/STAT signalling and a regulator of Rho1 in *Drosophila* haematopoiesis. J. Cell Sci. 127, 101–110. https://doi.org/10.1242/jcs.132852

Bloomquist, B.T., Shortridge, R.D., Schneuwly, S., Perdew, M., Montell, C., Steller, H., Rubin, G., Pak, W.L., 1988. Isolation of a putative Phospholipase C gene of *Drosophila*, norpA, and its role in phototransduction. Cell 54, 723–733. https://doi.org/10.1016/S0092-8674(88)80017-5

Bronstein, R., Levkovitz, L., Yosef, N., Yanku, M., Ruppin, E., Sharan, R., Westphal, H., Oliver, B., Segal, D., 2010. Transcriptional regulation by CHIP/LDB complexes. PLoS Genet. 6, 12–15. https://doi.org/10.1371/journal.pgen.1001063

Carty, D.J., Padrell, E., Codina, J., Birnbaumer, L., Hildebrandt, J.D., Iyengar, R., 1990. Distinct guanine nucleotide binding and release properties of the three Gi proteins. J.Biol.Chem. 265, 6268–6273.

Chahdi, A., Daeffler, L., Gies, J.P., Landry, Y., 1998. Drugs interacting with G protein alpha subunits: selectivity and perspectives. Fundam. Clin. Pharmacol. 12, 121–132. https://doi.org/10.4161/epi.22552

Cockcroft, S., 2006. The latest phospholipase C, PLCη, is implicated in neuronal function. Trends Biochem. Sci. 31, 4–7. https://doi.org/10.1016/j.tibs.2005.11.003

Dahdal, D., Reeves, D.C., Ruben, M., Akabas, M.H., Blau, J., 2010. *Drosophila* pacemaker neurons require G protein signaling and GABAergic inputs to generate twenty-four-hour behavioral rhythms. Neuron 68, 964–977. https://doi.org/10.1016/j.neuron.2010.11.017

Egger-Adam, D., Katanaev, V.L., 2010. The trimeric G protein Go inflicts a double impact on axin in the Wnt/Frizzled signalling pathway. Dev. Dyn. 239, 168–183. https://doi.org/10.1002/dvdy.22060

Elia, N., Frechter, S., Gedi, Y., Minke, B., Selinger, Z., 2005. Excess of Gβe over Gqαe *in vivo* prevents dark, spontaneous activity of *Drosophila* photoreceptors. J. Cell Biol. 171, 517–526. https://doi.org/10.1083/jcb.200506082

Freeman, J.M., Davey, R.B., Kappmeyer, L.S., Kammlah, D.M., Olafson, P.U., 2010. Bm86 midgut protein sequence variation in South Texas cattle fever ticks. Parasit. Vectors 3, 101–109. https://doi.org/10.1186/1756-3305-3-101

Freissmuth, M., Waldhoer, M., Bofill-Cardona, E., Nanoff, C., 1999. G protein antagonists signalling mechanism and potential target sites for drug action. TiPS 20, 237–245.

Fuse, N., Yu, F., Hirose, S., 2013. Gprk2 adjusts Fog signalling to organize cell movements in *Drosophila* gastrulation. Development 140, 4246–4255. https://doi.org/10.1242/dev.093625

García-García, J.C., Montero, C., Redondo, M., Vargas, M., Canales, M., Boue, O., Rodríguez, M., Joglar, M., Machado, H., González, I.L., Valdés, M., Méndez, L., 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. https://doi.org/10.1016/S0264-410X(99)00548-4

García-García, J.C., Gonzalez, I.L., Gonzalez, D., Valdes, M., Mendez, L., Lamberti, J., D'Agostino, B., Citroni, D., Fragoso, H., Ortiz, M., Rodriguez Valle, M., de la Fuente, J., 1999. Sequence variations

in the *Boophilus microplus* Bm86 locus and implications for immunoprotection in cattle vaccinated with this antigen. Exp. Appl. Acarol. 23 23, 883–895. https://doi.org/10.1023/A:1006270615158

Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C.A., Finley, R.L., White, K.P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R.A., McKenna, M.P., Chant, J., Rothberg, J.M., 2003. A protein interaction map of *Drosophila melanogaster*. Science (80-.). 302, 1727–1736. https://doi.org/10.1126/science.1090289

Granderath, S., Stollewerk, A, Greig, S., Goodman, C.S., O'Kane, C.J., Klämbt, C., 1999. loco encodes an RGS protein required for *Drosophila* glial differentiation. Development 126, 1781–1791. https://doi.org/10.1002/(SICI)1096-9861(19981207)402:1<32::AID-CNE3>3.0.CO;2-V

Hamada, K., Shimizu, T., Matsui, T., Tsukita, S., Tsukita, S., Hakoshima, T., 2000. Structural basis of the membrane-targeting and unmasking mechanisms of the radixin FERM domain. EMBO J. 19, 4449–4462. https://doi.org/10.1093/emboj/19.17.4449

Hardie, R.C., Martin, F., Cochrane, G.W., Juusola, M., Georgiev, P., Raghu, P., 2002. Molecular basis of amplification in *Drosophila* phototransduction: Roles for G protein, Phospholipase C, and Diacylglycerol Kinase. Neuron 36, 689–701. https://doi.org/10.1016/S0896-6273(02)01048-6

Heinz, D.W., Ryan, M., Smith, M.P., Weaver, L.H., Keana, J.F.W., Hayes Griffith, O., 1996. Crystal structure of phosphatidylinositol-specific Phospholipase C from *Bacillus cereus* in complex with glucosaminyl(α1→6)-D-myo-inositol, an essential fragment of GPI anchors. Biochemistry 35, 9496–9504. https://doi.org/10.1021/bi9606105

Hiramoto, M., Hiromi, Y., 2006. ROBO directs axon crossing of segmental boundaries by suppressing responsiveness to relocalized Netrin. Nat. Neurosci. 9, 58–66. https://doi.org/10.1038/nn1612

Höller, C., Freissmuth, M., Nanoff, C., 1999. G proteins as drug targets. Cell. Mol. Life Sci. 55, 257–270. https://doi.org/10.1007/s000180050288

Ja, W.W., Wiser, O., Austin, R.J., Jan, L.Y., Roberts, R.W., 2006. Turning G proteins on and off using peptide ligands. ACS Chem. Biol. 1, 570–574. https://doi.org/10.1021/cb600345k.Turning

Kaewmongkol, S., Kaewmongkol, G., Inthong, N., Lakkitjaroen, N., Sirinarumitr, T., Berry, C.M., Jonsson, N.N., Stich, R.W., Jittapalapong, S., 2015. Variation among Bm86 sequences in *Rhipicephalus* (*Boophilus*) *microplus* ticks collected from cattle across Thailand. Exp. Appl. Acarol. 66, 247–256. https://doi.org/10.1007/s10493-015-9897-0

Kain, P., Chakraborty, T.S., Sundaram, S., Siddiqi, O., Rodrigues, V., Hasan, G., 2008. Reduced odor responses from antennal neurons of Gq , Phospholipase C , and rdgA mutants in *Drosophila* support a role for a phospholipid intermediate in insect olfactory transduction. J. Neurosci. 28, 4745–4755. https://doi.org/10.1523/JNEUROSCI.5306-07.2008

Kamat, A., Carpenter, G., 1997. Phospholipase C-γ1: Regulation of enzyme function and role in growth factor-dependent signal transduction. Cytokine Growth Factor Rev. 8, 109–117. https://doi.org/10.1016/S1359-6101(97)00003-8

Kariya, K.I., Kim Bui, Y., Gao, X., Sternberg, P.W., Kataoka, T., 2004. Phospholipase Cε regulates ovulation in *Caenorhabditis elegans*. Dev. Biol. 274, 201–210. https://doi.org/10.1016/j.ydbio.2004.06.024

Katanaev, V.L., Ponzielli, R., Sémériva, M., Tomlinson, A., 2005. Trimeric G protein-dependent frizzled signalling in *Drosophila*. Cell 120, 111–122. https://doi.org/10.1016/j.cell.2004.11.014

Katanayeva, N., Kopein, D., Portmann, R., Hess, D., Katanaev, V.L., 2010. Competing activities of heterotrimeric G proteins in *Drosophila* wing maturation. PLoS One 5. https://doi.org/10.1371/journal.pone.0012331

Kaziro, Y., Itoh, H., Kozasa, T., Nakafuku, M., Satoh, T., 1991. Signal-transducing structure and function of GTP-binding proteins. Annu. Rev. Biochem. 60, 349–400.

Kemp, D.H., Koudstaal, D., Roberts, J. a, Kerr, J.D., 1975. Feeding of *Boophilus microplus* larvae on a partially defined medium through thin slices of cattle skin. Parasitology 70, 243–254.

Kim, S., McKay, R.R., Miller, K., Shortridge, R.D., 1995. Multiple subtypes of Phospholipase C are encoded by the norpA gene of *Drosophila melanogaster*. J. Biol. Chem.

https://doi.org/10.1074/jbc.270.24.14376

Kimple, A.J., Bosch, D.E., Giguere, P.M., Siderovski, D.P., 2011. Regulators of G-protein signalling and their Galpha substrates: Promises and challenges in their use as drug discovery targets. Pharmacol. Rev. 63, 728–749. https://doi.org/10.1124/pr.110.003038.discovery

Kornmann, B., Osman, C., Walter, P., 2011. The conserved GTPase Gem1 regulates endoplasmic reticulum-mitochondria connections. Proc. Natl. Acad. Sci. 108, 14151–14156. https://doi.org/10.1073/pnas.1111314108

Latchman, D.S., 2015. Post-transcriptional processes, in: Gene Control. Garland Science, pp. 205–282.

Leach, A., Bruls, M., Grimsby, J. *et al*., 2015. Genetic diversity and protective efficacy of the RTS,S/AS01 malaria vaccine. N. Engl. J. Med. 373, 2025–2037. https://doi.org/10.1056/nejmoa1505819

Lees, K., Bowman, A.S., 2007. Tick neurobiology: Recent advances and the post-genomic era. Invertebr. Neurosci. 7, 183–198. https://doi.org/10.1007/s10158-007-0060-4

Linder, M.E., Ewald, D. a, Miller, R.J., Gilman, a G., 1990. Purification and characterization of Go alpha and three types of Gi alpha after expression in *Escherichia coli*. J. Biol. Chem. 265, 8243–8251.

Louw, E., van der Merwe, N.A., Neitz, A.W.H., Maritz-Olivier, C., 2013. Evolution of the tissue factor pathway inhibitor-like Kunitz domain-containing protein family in *Rhipicephalus microplus*. Int. J. Parasitol. 43, 81–94. https://doi.org/10.1016/j.ijpara.2012.11.006

Maritz-Olivier, C., Stutzer, C., Jongejan, F., Neitz, A.W.H., Gaspar, A.R.M., 2007. Tick anti-hemostatics: targets for future vaccines and therapeutics. Trends Parasitol. 23, 397–407. https://doi.org/10.1016/j.pt.2007.07.005

Martínez-Arzate, S.G., Sánchez-Bermúdez, J.C., Sotelo-Gómez, S., Diaz-Albiter, H.M., Hegazy-Hassan, W., Tenorio-Borroto, E., Barbosa-Pliego, A., Vázquez-Chagoyán, J.C., 2019. Genetic diversity of Bm86 sequences in *Rhipicephalus* (*Boophilus*) *microplus* ticks from Mexico: Analysis of haplotype distribution patterns. BMC Genet. 20, 1–12. https://doi.org/10.1186/s12863-019-0754-8

Moorhouse, D.E., Tatchell, R.J., 1966. The feeding processes of the cattle tick *Boophilus microplus* (Canestrini): a study in host-parasite relations. Parasitology 56, 623–631.

Nakamura, Y., Hamada, Y., Fujiwara, T., Enomoto, H., Hiroe, T., Tanaka, S., Nose, M., Nakahara, M., Yoshida, N., Takenawa, T., Fukami, K., 2005. Phospholipase C-delta1 and -delta3 are essential in the trophoblast for placental development. Mol. Cell. Biol. 25, 10979–88. https://doi.org/10.1128/MCB.25.24.10979-10988.2005

Neer, E.J., 1995. Heterotrimeric G proteins: organizers of transmembrane signals. Cell 80, 249–257. https://doi.org/10.1016/0092-8674(95)90407-7

Neves, S.R., Ram, P.T., Iyengar, R., 2002. G protein pathways. Science (80-.). 296, 1636–1639. https://doi.org/10.1126/science.1071550

Nijhof, A.M., Balk, J.A., Postigo, M., Jongejan, F., 2009. Selection of reference genes for quantitative RT-PCR studies in *Rhipicephalus* (*Boophilus*) *microplus* and *Rhipicephalus appendiculatus* ticks and determination of the expression profile of Bm86. BMC Mol. Biol. 10, 112–124. https://doi.org/10.1186/1471-2199-10-112

Nikolaidou, K.K., Barrett, K., 2004. A Rho GTPase signalling pathway is used reiteratively in epithelial folding and potentially selects the outcome of Rho activation. Curr. Biol. 14, 1822–1826. https://doi.org/10.1016/j.cub.2004.09.080

Perner, J., Kropáčková, S., Kopáček, P., Ribeiro, J.M.C., 2018. Sialome diversity of ticks revealed by RNAseq of single tick salivary glands. PLoS Negl. Trop. Dis. 12, 1–17. https://doi.org/10.1371/journal.pntd.0006410

Quan, F., Wolfgang, W.J., Forte, M., 1993. A *Drosophila* G-protein alpha subunit, Gf alpha, expressed in a spatially and temporally restricted pattern during *Drosophila* development. Proc. Natl. Acad. Sci. U. S. A. 90, 4236–40.

Rand, K.N., Moore, T., Sriskantha, A., Spring, K., Tellam, R., Willadsen, P., Cobon, G.S., 1989. Cloning and expression of a protective antigen from the cattle tick *Boophilus microplus*. Proc. Natl. Acad. Sci. U. S. A. 86, 9657–61. https://doi.org/10.1073/pnas.86.24.9657

Rasenick, M.M., Watanabe, M., Lazarevic, M.B., Hatta, S., Hamm, H.E., 1994. Synthetic peptides as probes for G protein function: Carboxyl-terminal Gαs peptides mimic Gs and evoke high-affinity

agonist binding to β-adrenergic receptors. J. Biol. Chem. 269, 21519–21525.

Roberts, J.A., 1971. Behaviour of larvae of cattle tick, *Boophilus microplus* (canestrini), on cattle of differing degrees of resistance. J. Parasitol. 57, 651-. https://doi.org/10.2307/3277933

Saunders, C.M., Larman, M.G., Parrington, J., Cox, L.J., Royse, J., Blayney, L.M., Swann, K., Lai, F.A., 2002. PLCζ: a sperm-specific trigger of Ca^{2+} oscillations in eggs and embryo development. Development 129, 3533–3544.

Schetters, T., Bishop, R., Crampton, M., Kopáček, P., Lew-Tabor, A., Maritz-Olivier, C., Miller, R., Mosqueda, J., Patarroyo, J., Rodriguez-Valle, M., Scoles, G.A., de la Fuente, J., 2016. Cattle tick vaccine researchers join forces in CATVAC. Parasit. Vectors 9, 105. https://doi.org/10.1186/s13071-016-1386-8

Seifert, G.W., Springell, P.H., Tatchell, R.J., 1968. Radioactive studies on the feeding of larvae, nymphs, and adults of the cattle tick, *Boophilus microplus* (Canestrini). Parasitology 58, 415–430. https://doi.org/10.1017/S0031182000069444

Shibatohge, M., Kariya, K.I., Liao, Y., Hu, C.D., Watari, Y., Goshima, M., Shima, F., Kataoka, T., 1998. Identification of PLC210, a *Caenorhabditis elegans* Phospholipase C, as a putative effector of Ras. J. Biol. Chem. 273, 6218–6222. https://doi.org/10.1074/jbc.273.11.6218

Shieh, B.H., Zhu, M.Y., Lee, J.K., Kelly, I.M., Bahiraei, F., 1997. Association of INAD with NORPA is essential for controlled activation and deactivation of *Drosophila* phototransduction *in vivo*. Proc. Natl. Acad. Sci. U. S. A. 94, 12682–7. https://doi.org/VL - 94

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sossai, S., Peconick, A.P., Sales, P.A., Marcelino, F.C., Vargas, M.I., Neves, E.S., Patarroyo, J.H., 2005. Polymorphism of the Bm86 gene in South American strains of the cattle tick *Boophilus microplus*. Exp. Appl. Acarol. 37, 199–214. https://doi.org/10.1007/s10493-005-3262-7

Suh, P.-G., Park, J.-I., Manzoli, L., Cocco, L., Peak, J.C., Katan, M., Fukami, K., Kataoka, T., Ryu, S.Y.& S.H., 2008. Multiple roles of phosphoinositide-specific Phospholipase C isozymes. BMB Rep. 41, 415–434.

Trentelman, J.J.A., Kleuskens, J.A.G.M., Van De Crommert, J., Schetters, T.P.M., 2017. A new method for *in vitro* feeding of *Rhipicephalus australis* (formerly *Rhipicephalus microplus*) larvae: A valuable tool for tick vaccine development. Parasites and Vectors 10, 1–9. https://doi.org/10.1186/s13071-017-2081-0

Ullmann, A.J., Lima, C.M.R., Guerrero, F.D., Piesman, J., Iv, W.C.B., 2005. Genome size and organization in the black-legged tick , *Ixodes scapularis* and the Southern cattle tick , *Boophilus microplus*. Insect Mol. Biol. 14, 217–222. https://doi.org/10.1111/j.1365-2583.2005.00551.x

Vázquez-Manrique, R.P., Nagy, A.I., Legg, J.C., Bales, O.A.M., Ly, S., Baylis, H.A., 2008. Phospholipase C-ε regulates epidermal morphogenesis in *Caenorhabditis elegans*. PLoS Genet. 4, e100043. https://doi.org/10.1371/journal.pgen.1000043

Voogd, T.E., Vansterkenburg, E., Wilting, J., Janssen, L., 1993. Recent research in the biological activity of Suramin. Pharmacol. Rev. 45, 177–199. https://doi.org/0031-6997/93/4502-0177$03.00/0

Wang, T., Dowal, L., El-maghrabi, M.R., Rebecchi, M., Scarlata, S., 2000. The pleckstrin homology domain of Phospholipase C-beta2 links the binding of Gbeta/gamma to activation of the catalytic core. J. Biol. Chem. 275, 7466–7469.

Williams, R.L., Katan, M., 1996. Structural views of phosphoinositide-specific Phospholipase C: Signalling the way ahead. Structure 4, 1387–1394. https://doi.org/10.1016/S0969-2126(96)00146-3

Wolfgang, W.J., Clay, C., Parker, J., Delgado, R., Labarca, P., Kidokoro, Y., Forte, M., 2004. Signalling through Gsα is required for the growth and function of neuromuscular synapses in *Drosophila*. Dev. Biol. 268, 295–311. https://doi.org/10.1016/j.ydbio.2004.01.007

Yu, F., Wang, H., Qian, H., Kaushik, R., Bownes, M., Yang, X., Chia, W., 2005. Locomotion defects, together with Pins, regulates Heterotrimeric G-protein signalling during *Drosophila* neuroblast asymmetric divisions. Genes Dev. 19, 1341–1353. https://doi.org/10.1101/gad.1295505

Appendix 1

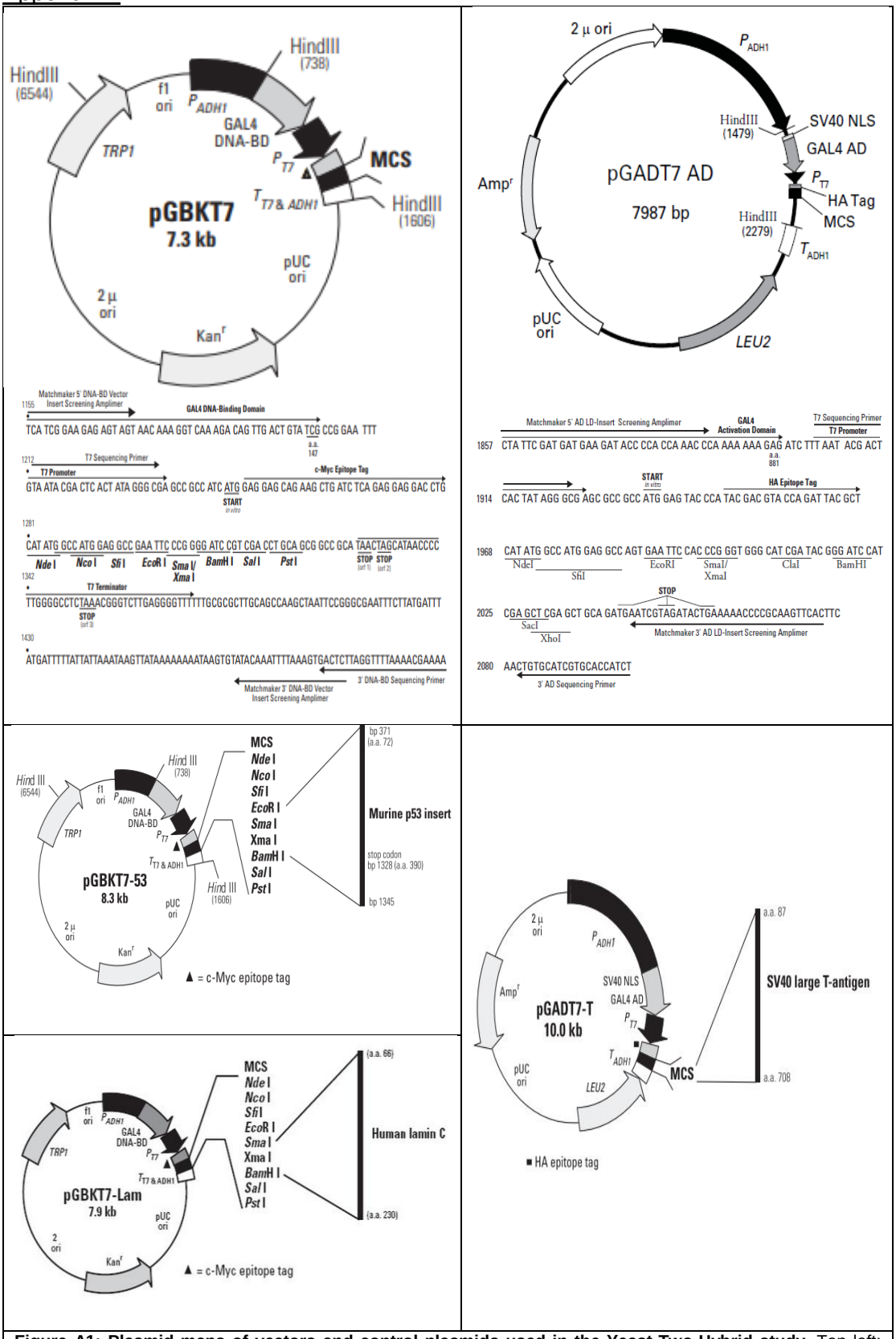
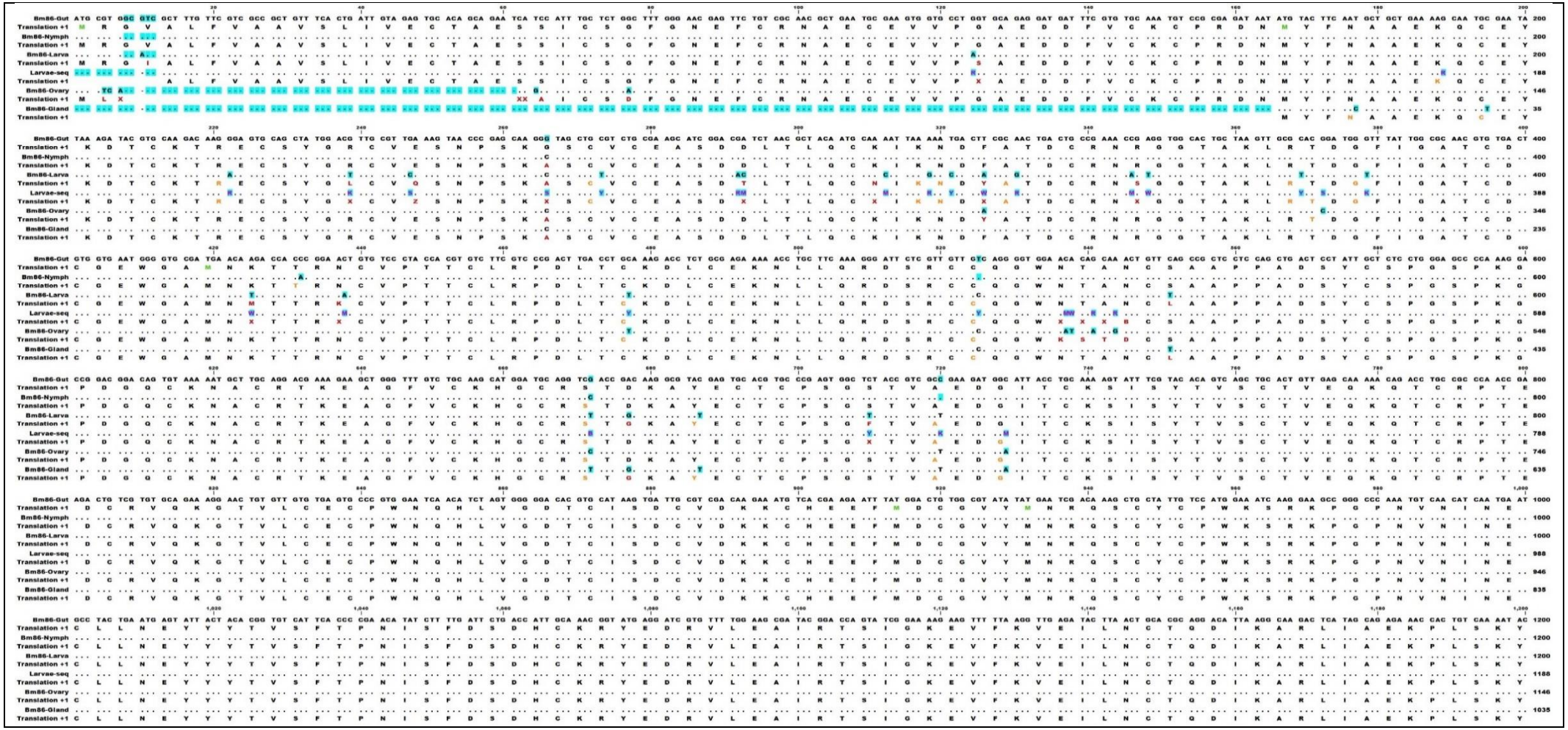


Figure A1: Plasmid maps of vectors and control plasmids used in the Yeast-Two-Hybrid study. Top left: pGBKT7 Binding Domain plasmid and Multiple cloning site sequence. Top right: pGADT7 activation domain plasmid and multiple cloning site sequence. Bottom left: pGBKT7 Control plasmids pGBKT7-53 (encodes a p53 gene clone for a positive interaction with SV40 large T-antigen) and pGBKT7-LAM (encodes a Human lamin C gene for a negative interaction with SV40 large T-antigen) for transformation use with Bottom right pGADT7-T (encoding a gene for the SV40 large T-antigen cloned into the BamHI sites of the pGADT7 MCS by Li and Fields (1993)).

Appendix 2


Bm86-Gut
Translation +1
Bm86-Nymph
Bm86-Larva
Larvae-seq
Bm86-Ovary
Bm86-Gland

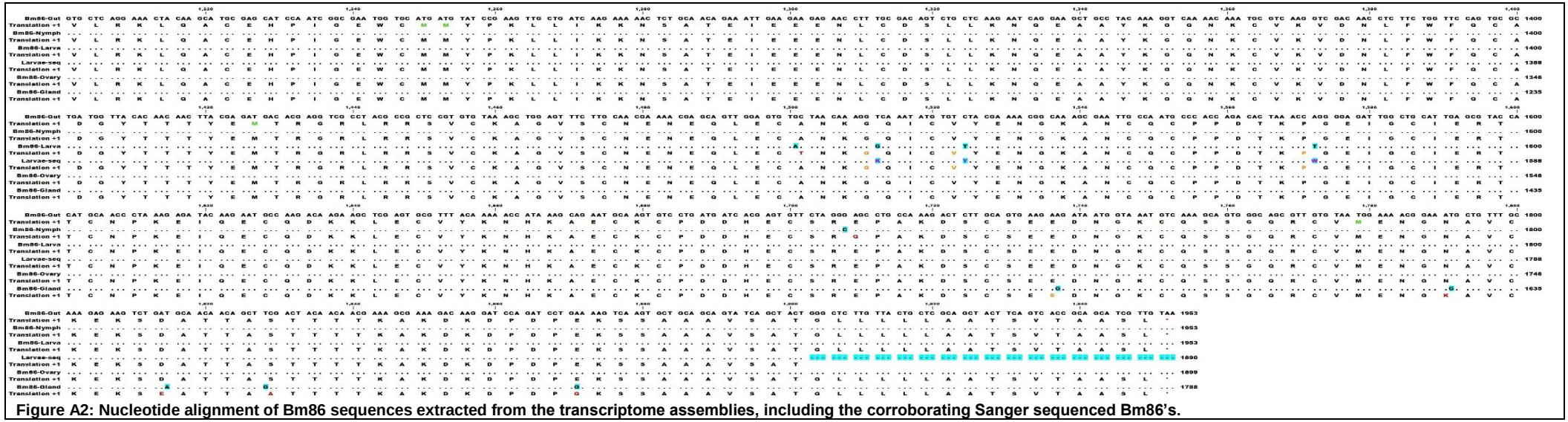


Figure A2: Nucleotide alignment of Bm86 sequences extracted from the transcriptome assemblies, including the corroborating Sanger sequenced Bm86's.