EXPRESSION OF ODORANT RECEPTOR GENES IN SAVANNAH TSETSE FLY *GLOSSINA MORSITANS MORSITANS* WESTWOOD, 1850 (DIPTERA: GLOSSINIDAE)

CYRUS TERER TARE

MASTER OF SCIENCE
(Molecular Biology and Bioinformatics)

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

2018

Expression of Odorant Receptor Genes in Savannah Tsetse Fly *Glossina morsitans morsitans* Westwood, 1850 (Diptera: Glossinidae)

Tare Cyrus Terer

A Thesis submitted for the Degree of Master of Science in Molecular Biology and Bioinformatics in the Jomo Kenyatta University of Agriculture and Technology

2018

DECLARATION

This thesis is my original work and has not been presented elsewhere for the award of degree.

Signature.....… Date.....

Cyrus Terer Tare

This thesis has been submitted for examination with our approval as Supervisors:

Signature……..… Date…….

Dr. Steven Ger Nyanjom

JKUAT, Kenya

Signature……..… Date……...

Prof. Fred Wamunyokoli

JKUAT, Kenya

Signature.....… Date.....

Dr. George Fredrick Obiero

TU-K, Kenya

DEDICATION

To my Aunt Jennifer Tare

ACKNOWLEDGEMENTS

I sincerely thank the Almighty God for granting me the gift of life. He gave me patience, understanding, and great health to pursue my studies.

My appreciation goes to my supervisors Prof. Fred Wamunyokoli, Dr. Steven Ger Nyanjom and Dr. George Obiero for their continued support, guidance, and patience during the entire period of my work. I am also grateful to Dr. Joel Bargul for the assistance he gave me. I extend my gratitude to the technicians in Biochemistry department for their encouragement and support.

I am also grateful and greatly indebted to Maurice Korir and Feisal Yegon, from whom I got a lot of support and guidance right from the beginning of the program. Thank you, brothers.

TABLE OF CONTENTS

DECLARATION.....ii
DEDICATION.....iii
ACKNOWLEDGEMENTS.....iv
TABLE OF CONTENTS.....v
LIST OF TABLES.....viii
LIST OF FIGURES.....ix
LIST OF APPENDICES.....x
LIST OF ABBREVIATIONS AND ACRONYMS.....xi
ABSTRACT.....xiii
CHAPTER ONE.....1
INTRODUCTION.....1
1.1 Background Information.....1
1.2 Problem Statement.....4
1.3 Justification.....5
1.4 Research Questions.....6
1.5 Objectives.....6
1.5.1 Overall Objective.....6
1.5.2 Specific Objective.....7
CHAPTER TWO.....8
LITERATURE REVIEW.....8
2.1 Distribution of tsetse fly in Africa.....8
2.2 Trypanosomiasis Control Strategies.....10
2.3 Vector Control Approaches.....11
2.3.1 Insecticides.....11
2.3.2 Odour-baited traps.....12
2.3.3 Sterile insect technique.....12
2.3.4 Chemoreception in Vector Control.....13
2.4 Chemoreception in Insects.....14
2.4.1 Soluble chemosensory proteins.....17

2.4.2 Odorant degrading enzymes....17
2.4.3 Odorant Binding Proteins and Pheromone Binding Proteins....18
2.4.4 Gustatory Receptors....19
2.4.5 Ionotropic Receptors....20
2.4.6 Odorant Receptors....21
2.5 Real time PCR for Gene Expression....22
2.6 Bioinformatics Analyses....23
2.6.1 Multiple Sequence Alignment....23
2.6.2 Transmembrane Domains....24
2.6.3 Conserved Domains....24
2.6.4 Gene Ontology....25
2.6.5 Protein Structure Analysis....25
CHAPTER THREE....27
MATERIALS AND METHODS....27
3.1 Study Area and Preparation of Insects and Tissue....27
3.1.1 Study Site....27
3.1.2 Insect Rearing....27
3.1.3 Handling of Samples....27
3.2 Determination of Levels of OR Genes in *G. m. morsitans* samples....28
3.2.1 Sample Preparation for total RNA Extraction....28
3.2.2 Extraction of Total RNA....28
3.2.3 Determination of RNA yield and quality....29
3.2.4 Determination of RNA integrity....29
3.2.5 Synthesis of cDNA by Reverse Transcription....30
3.2.6 Amplification of synthesised cDNA....30
3.2.7 Gel Electrophoresis of PCR products....33
3.3 Real Time Polymerase Chain Reaction....34
3.3.1 Analysis of RT-PCR Data....34
3.4 Comparative Bioinformatics Analyses....35
CHAPTER FOUR....37
RESULTS....37

4.1 Expression of OR Genes in *G. m. morsitans* samples.....37
4.1.1 RNA Quality Analysis.....37
4.1.2 Amplification of synthesised cDNA.....38
4.2 Expression of *Glossina morsitans morsitans* OR genes.....43
4.3 Bioinformatics Analyses.....47
4.3.1 Conserved Domains.....47
4.3.2 Protein Structure and Ligands.....47
4.3.2 Transmembrane Domains.....53
4.3.3 Multiple Sequence Alignment.....54
CHAPTER FIVE.....55
DISCUSSION.....55
CHAPTER SIX.....60
CONCLUSIONS AND RECOMMENDATIONS.....60
6.1 Conclusions.....60
6.2 Recommendations.....60
REFERENCES.....61
APPENDICES.....82

LIST OF TABLES

Table 3.1: Primer sequences for amplifying *G. m. morsitans* Ors……..…….31

Table 4.1: Nanodrop readings of single strand mRNA extracted from *G. m. morsitans* samples……..…….37

Table 4.2: *G. m. m* ORs and best protein blast match to *D. melanogaster*, Conserved domain and Gene Ontology Databases…….49

LIST OF FIGURES

Figure 2.1: Tsetse fly distribution in Africa……..……..9

Figure 4.1: Gel images showing the integrity of the mRNA……...……..38

Figure 4.2: Gel images of *G. m. morsitans* OR genes amplified using conventional PCR…….39

Figure 4.3: Odorant Receptor profiles of *G. m. morsitans* samples……...…...44

Figure 4.4: TMDs of GmmOR1 output by TMHMM version 2 algorithm...……53

LIST OF APPENDICES

Appendix i: *Glossina morsitans morsitans* ORs and their best protein blast match to Drosophila melanogaster, mosquitoes and other *Glossina* species……..…...82

Appendix ii: Protein Structure of the odorant receptors and their respective ligands…….....88

Appendix iii: Multiple sequence alignment of *G. m. morsitans* protein sequences with homologs from *D. melanogaster*, mosquitoes and other *Glossina*……..…96

LIST OF ABBREVIATIONS AND ACRONYMS

AAT	African Animal Trypanosomiasis
AL	Antennal lobe
CDD	Conserved Domain Database
cDNA	Complementary Deoxyribonucleic Acid
CSPs	Chemosensory soluble proteins
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GPCRs	G-Protein Coupled Receptors
GRs	Gustatory Receptors
GRNs	Gustatory Receptor Neurons
HAT	Human African Trypanosomiasis
mRNA	Messenger Ribonucleic Acid
IRs	Ionotropic Receptors
OBPs	Odorant Binding Proteins
ODEs	Odorant Degrading Enzymes
ORs	Odorant Receptors
ORNs	Olfactory Receptor Neurons
PBPs	Pheromone Binding Proteins
PCR	Polymerase Chain Reaction
RT-PCR	Real Time PCR
RNA	Ribonucleic Acid
SIT	Sterile insect technique
SNMPs	Sensory Neuron Membrane Proteins

TMD	Transmembrane Domain
WHO	World Health Organisation

ABSTRACT

Tsetse flies are the sole cyclical vectors of African trypanosomes with negative agricultural and economic impacts in sub-Saharan Africa, yet management of their populations has not been adequately addressed. Ecologically relevant molecules like odorant receptors and gustatory receptors were recently identified computationally. However, understanding the ecological behaviour of the tsetse remains limited because of absence to informative expression of the molecular data. This experimental study aimed to establish the expression levels of *Glossina morsitans morsitans* odorant receptor (OR) genes in the antennae and legs at different developmental stages of adults' for sexes. Whole tissue mRNA was extracted from fresh samples of *G. m. morsitans* larvae, pupae, adults' antennae and legs and used to synthesise cDNA. The *Glossina morsitans morsitans* Odorant Receptor genes expression levels were quantified from the cDNA relative amounts. *Glossina morsitans morsitans* protein sequences and *Drosophila melanogaster* orthologs in FlyBase and mosquitoes and other sequenced *Glossina* sequences from VectorBase were aligned for conserved sequences using MUSCLE alignment tool and their transmembrane domains ascertained with TMHMM tool. All the GmmORs had the 7-transmembrane domain, a hallmark feature for ORs, except GmmOR20. In adults, the OR genes were highly expressed in antennae than in the legs. GmmOR3 and GmmOR45 were dimorphically expressed in antennae in females and males respectively, while GmmOR26 and GmmOR20 were differentially expressed at high levels in female and male tsetse legs respectively. High expression OR genes in developmental stages, and particularly GmmOR28 in pupae, may suggest their involvement in development. The exclusive expression of GmmOR20 in larvae is indicative of the role this gene may play especially on feeding habits of the larvae. Higher expression of GmmOR15 in larvae also suggested that the gene product is critical in escape responses similar to confirmed role of its homolog in *D. melanogaster*, DmOR45. These results open avenues to determine functional roles of tsetse ORs by matching the receptors to their ecological odor ligands. This knowledge of ecologically active ORs will aid in designing effective tsetse control targets. Accordingly, the expressed OR genes established in this study should be subjected to functional studies to determine their specificity and, subsequently, their suitability to be used to be targeted in developing pest control measures.

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Tsetse flies, *Glossina* species, are the sole cyclical vectors of trypanosomes that cause Trypanosomiasis in animals (African Animal Trypanosomiasis, AAT) and sleeping sickness in humans (Human African Trypanosomiasis, HAT) (Spickler, 2010). Human African trypanosomiasis takes either the chronic or acute form depending on the parasite involved: *Trypanosoma brucei gambiense* accounts for more than 97% of the chronic reported cases and is found in 25 countries in west and central Africa. *Trypanosoma brucei rhodesiense* is found in 12 countries in eastern and southern Africa, and this acute form represents under 3% of reported cases and causes an acute infection (Barrett *et al.,* 2003). Trypanosomiasis is fast re-emerging (Simmaro *et al.,* 2011) and has considerable impact on public health and agricultural economic development in sub-Saharan Africa affecting a third of sub-Saharan Africa livestock and human livelihoods (Perry, 2015). Wild animals form the main reservoir of all the parasite species (Horan *et al.,* 2010; Pierre *et al.,* 2010). Infection of domesticated livestock usually occurs when domestic animals share pasture with wild herbivores (Black, Seed, and Murphy, 2001). Ecologically active vector species are *Glossina palpalis, G. m. morsitans, G. swynnertoni, G. pallidipes, G. fuscipes, G. austeni, G. longipalpis, G. brevipalpis, G. fusca, G. tabaniformis, G. vanhoofi* and *G. tachinoides* are important *Glossina* (Hoare, 1970; Magez and Radwanska, 2014).

The *Glossina* species are divided into three main categories: Riverine, Savannah and Forest species (Leak, 1999; Gouteux and Jarry, 1998). The Riverine species are

mainly found in humid areas and close to water bodies like lake basins, riverine areas, swamps and mangrove. They prefer reptilian hosts like crocodiles (Gooding and Krafsur, 2005). The Forest species are found in the thick forested areas of Central Africa like Cameroon, Mozambique and some isolated parts of South Africa, and prefer blood-hosts like monkeys (Simo *et al.,* 2008). The Savannah species are distributed in most of the dry expansive African Savannah grasslands and prefer hoofed-mammals like buffaloes. Whereas the forest species have been found to prefer human hosts thereby transmitting sleeping sickness, the savannah species transmit nagana to livestock (Waiswa *et al.,* 2006).

Feeding preferences of different species of tsetse fly bear no relationship to the relative abundance of the vertebrate hosts in a given habitat (Gikonyo *et al.,* 2003). For instance, in Zimbabwe *G. m. morsitans* showed non-preference for the impala (*Aepyceros melampus*) with respect to landing and engorgement compared to ox (*Bos indicus*), bush pig (*Potamochoerus porcus*), and warthog (*Potamochoerus aethiopicus*) (Muturi *et al.,* 2011). Similarly, in Kenya, *G. pallidipes* were shown to alight and engorge on buffalo (*Syncerus caffer*) or ox but not waterbuck (*Kobus defassa*), even when these animals are kept in adjacent enclosures (Bett, Saini, and Hassanali, 2015).

Tsetse flies rely mainly on chemoreception to execute critical behaviours such as host location, predator avoidance, mate pursuit and identification of larviposition and resting sites (Masiga *et al.,* 2014). Chemoreception involves a cascade of events by different proteins. The olfactory proteins consist of odorant binding proteins (OBPs), pheromone binding proteins (PBPs), odorant degrading enzymes (ODEs) and

chemosensory proteins (CSPs) which belong to soluble proteins and are found in the chemosensilla lymph (Leal, 2013). These functions in conjunction with receptor proteins including ORs, gustatory receptors (GRs), ionotropic receptors (IRs) and sensory neuron membrane proteins (SNMPs) (Leal, 2013), to complete the chemosensory cascade. The OBPs and CSPs bind hydrophobic ligands, transport them through the sensilla lymph and activate the ORs located within the Olfactory Receptor Neurons (ORNs) (Leal, 2013).

The exploitation of *Glossina* odor sensing system remains the most promising avenue to controlling trypanosomiases (Omolo *et al.,* 2009). Tsetse fly, as indeed other insects largely depend on its ORs to discriminate chemical compounds in its environment. This ability is dependent on the number and sensitivity of their ORs (Vosshall *et al.,* 2000). The sensitivity of the ORs is directly proportional to their expression levels, and such patterns are critical in deducing the roles of individual ORs through the insects' life stages, and even body regions (Wright and Smith, 2004).

Apart from antennae, insects also use other body parts like maxillary palps (Robertson *et al.,* 2003) for olfaction. The distribution and expression levels of ORs in the various body parts of an insect are role-specific (Du *et al.,* 2018). In *Aedes albopictus,* OR2 (AalOR2) is expressed in the antennae and its orthologs from *Anopheles gambiae* and *Ae. aegypti* with a female-biased expression (Hansson *et al.,* 2012). The expression of AalOR2 during larval stage suggests its important role in the feeding habits, whereas in adult females this gene is important for locating appropriate larviposition sites (Guidobaldi *et al.,* 2014). In silkworm moth *Bombyx*

mori, the sex pheromone receptor (PR), BmorOR1, is a male-antennae specific gene (Nakagawa *et al.,* 2005; Sakurai *et al.,* 2004). And in diamondback moth, *Plutella xyllotella*, three ORs (PxylOR1, PxylOR5 and PxylOR6) and an OBP gene were expressed in male legs (Sun *et al.,* 2013). These genes may have olfactory associated functions in male insect legs or may display biological functionality in a non-olfactory context (Sengul and Tu, 2008).

This study focussed on the expression of *G. morsitans morsitans* OR genes at different developmental stages and in two body tissues (antennae and legs) of both females and males. Considering that OR genes are critical in modulating tsetse olfactory behaviour, understanding their differential expression in both tsetse sexes is key in designing and improving potent tsetse fly attractants and repellents. The study findings inform tissue-based expression of the individual OR genes in *G. m. morsitans.*

1.2 Problem Statement

An estimated 50 million cattle are at risk of nagana infection in sub-Saharan Africa, with an estimated economic loss of up to US$ 4.5 billion in cattle production (Perry, 2015). Further, up to 70,000 people are infected annually, and without early treatment the disease is usually fatal. The disease has been known to disrupt daily activities, cause food insecurity, neglect of homestead, poor academic performance and even result in school drop-outs, and death (Bukachi, Wandibba and Nyamongo, 2017). Given that no effective tsetse control method currently exists, there is an urgent need for effective novel alternative strategies. It is believed that genes involved in olfaction can be targeted in developing innovative approaches to control

of tsetse flies. Therefore, determining expression levels of individual OR gene is critical towards such innovative molecular control approaches. The molecules, especially host odors, used by tsetse fly to locate food sources and pheromones critical for reproduction have been unravelled (Obiero *et al.,* 2014). The sequencing of *Glossina* genome has availed further information to be utilised towards ascertaining how molecules, such as pheromones and odorants, relate with the diverse tsetse fly molecular repertoire. Subsequently, tsetse chemoreceptors were published (Obiero *et al.,* 2014). However, these chemoreceptors were only computationally predicted and annotated but no functional study has been carried out yet. Further, the specific tissues that express ORs in tsetse are not yet profiled. Therefore, this study aimed to establish the differential expression of OR genes in developmental stages of tsetse flies.

1.3 Justification

Despite significant progress made in controlling trypanosomiasis, the drugs currently in use are ineffective owing to the complexity of their administration and toxicity. Most control strategies that target the tsetse populations are unsustainable and/or not friendly to the environment. Traps have are effective but cover small geographical areas and are labour demanding. This necessitates exploration to discover other alternatives to control the tsetse vector. It is thought that ORs offer potential targets in controlling tsetse fly because of the important role it plays in provoking behaviour for finding blood-meal host, resting sites and larviposition sites. The first study utilising receptors to deploy control measures was done in *Anopheles gambiae* where synthetic compound VUAA1 ((*N*-(4-ethylphenyl)-2-((4-ethyl-5-(3-pyridinyl)-4H-1,2,4-triazol-3-yl)thio)acetamide) was used as an agonist for co-receptor ORCO itself

(Jones *et al.,* 2012). This compound has been found to be active in ORCOs from other species as well, including dipterans (Andersson *et al.,* 2010). Odorant receptor (OR) are thought to bind hydrophobic odors delivered by OBPs from the external environment. They play critical role in discrimination and delivery of odors to ORNs which is the first step in signal transduction to higher brain centres. Analysis of the expression levels of the ORs in different body parts and developmental stages of the tsetse fly will reveal those that are highly expressed, hence could be playing crucial role in chemoreception; and can be targeted in designing tools and targets such as traps and deterrents for controlling tsetse fly through olfactory mediated behaviour. It can also be used to carrying out comparative functional studies with other tsetse species to understand role of chemoreception in tsetse flies.

1.4 Research Questions

1. What are the expression levels of OR genes in the antennae, and legs of both male and female *G. m. morsitans*?
2. What are the expression levels of OR genes in the larvae and pupae of *G. m. morsitans*?
3. What are the bioinformatics properties of GmmORs, like protein structure and ligands, and how they compare in other insects?

1.5 Objectives

1.5.1 Overall Objective

To determine the differential expression of OR genes in larvae, pupae, antennae and legs of both males and females *G. m. morsitans*.

1.5.2 Specific Objective

1. To determine differential expression of OR genes in the antennae and legs of males and females *G. m. morsitans*.
2. To determine the expression level of OR genes in *G. m. morsitans* larvae and pupae.
3. To carry out comparative analysis of *G. m. morsitans* ORs with orthologs from selected insects (Diptera: culicidae) using bioinformatics tools.

CHAPTER TWO

LITERATURE REVIEW

2.1 Distribution of tsetse fly in Africa

Tsetse flies are found in 37 countries of the sub-Saharan Africa covering approximately 10 million square kilometres (Figure 2.1) (Franco *et al.,* 2014; Simarro *et al.,* 2012). Glossina species are divided into three groups consisting of Forest (fusca), Riverine (palpalis) and Savannah (morsitans) based on their distribution and morphology (Vreysen *et al.,* 2013). Their geographical distribution is dependent on habitats with suitable land cover and hosts (Moore *et al.,* 2012). However, extensive land use due to increased population pressure has increased bush burning and land clearing for settlements and farming. Such practices have shaped tsetse fly distribution trends across Africa (Ilemobade, 2009).

Compared to savannah and riverine groups, the forest species are the largest and are mostly found in moist forests of West and Central Africa (Steverding, 2008). For instance, *G. brevipalpis* is mostly found in the dense vegetation like in South East region of Kwale County in Kenya (Mbahin *et al.,* 2013). Forest species preferentially feed on animals such as hippopotamus, cattle, bush pigs and buffaloes, and they transmit *Trypanosoma brucei brucei*, *T. suis* or *T. simiae* that causes nagana in livestock (Cecchi *et al.,* 2015).

The savannah flies occupy the woody savannah habitats of West, Central and East Africa. They are medium-sized and mainly feed on ungulates and other large animals (Steverding, 2008). However, blood-meal analyses from several savannah species

revealed that they get their blood meals from different wild animals (Simo *et al.,* 2008).

On the other hand, the riverine flies are the smallest in body size (Steverding, 2008). They prefer less densely vegetated, woody areas around forests and riverine vegetation (Cecchi *et al.,* 2008) to more arid areas as preferred by *G. tachinoides* (Bouyer *et al.,* 2010). This group of Glossina mainly feed on reptilian hosts and humans causing nagana and sleeping sickness, respectively (Simo *et al.,* 2008).

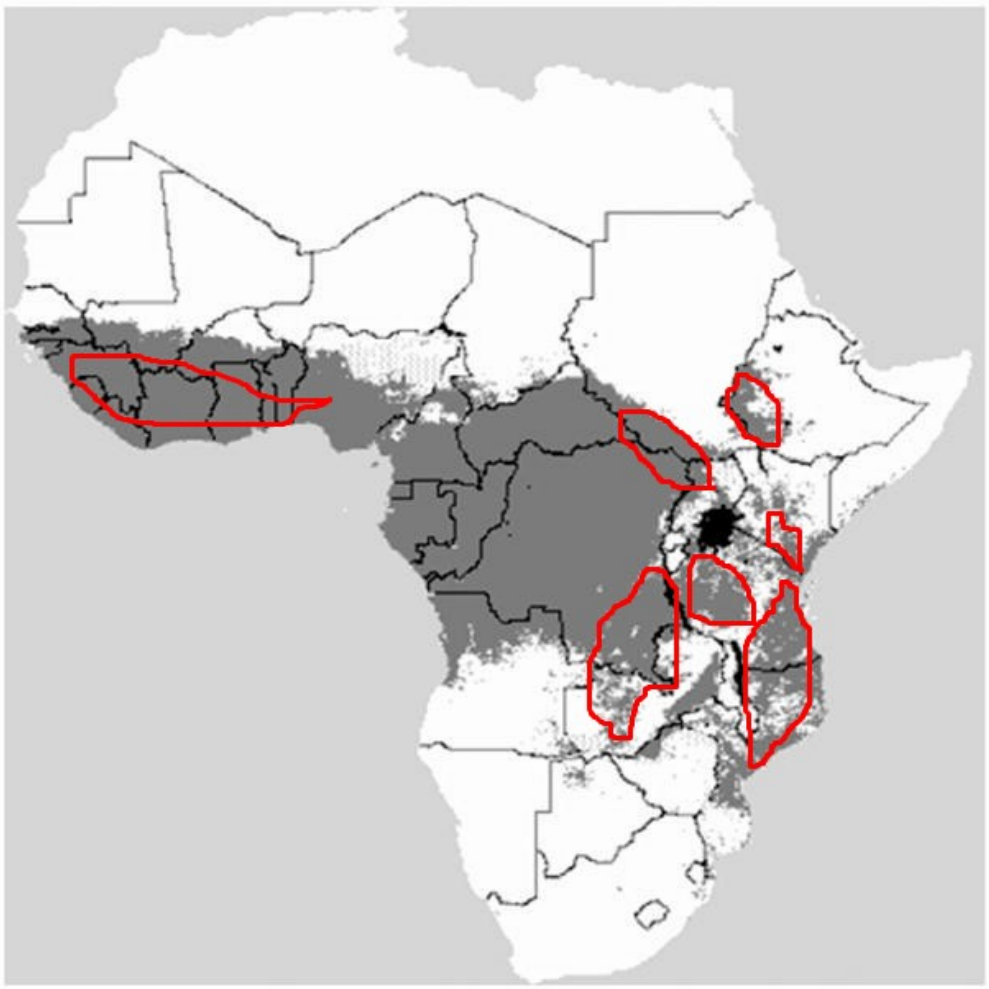


Figure 2.1. Tsetse fly distribution in Africa: Area shaded in grey indicates countries infested by tsetse flies and regions circled in red are *G. m. morsitans* infested areas. Adapted and modified from (Simarro *et al.,* 2012).

2.2 Trypanosomiasis Control Strategies

In both animals and humans, trypanosomiasis is controlled chemotherapeutically (medical/or veterinary) by treating the sick hosts and entomological approach where the populations of tsetse flies are targeted. The chemotherapeutic drugs are the main line of management of the disease (WHO, 2013). Drugs such as pentamidine has been used effectively against *T. b. gambiense* and suramin against *T. b. rhodesiense* in the first stage while Malarsopol, Eflornithine and Nifurtimox are effective against the parasites in their second stages. Even so, the drugs are expensive, developed resistance, have low efficacy and difficult to administer especially during the third stage of the disease because of their cytotoxic nature and usually result in death, and thereby ineffective against the parasites (Perry, 2015). The parasites' ability to change its surface glycoprotein has made it difficult to develop vaccines (Black and Mansfield, 2016). This has shifted disease management to the control of the vector which offers potential avenues to controlling the disease (Simarro *et al.,* 2008).

Tsetse fly control and eradication strategies that have been used in the past include; spraying the tsetse infested areas with insecticides (Vreysen *et al.,* 2000), use of targets impregnated with insecticides (Ryan, 2002), odour baited traps (Esterhuizen *et al.,* 2011) and sterile insect technique (SIT) (Allsopp, 2001). Trypanosomiasis treatment and tsetse fly control has resulted to annual economic losses of over US$ 4 billion in cattle production alone, and a further USD 4.7 billion in losses in Agricultural Domestic Product per year (Perry, 2015). The evasive nature of trypanosome from the hosts' immune system has hampered the development process of an effective vaccine. Therefore, current long term control of trypanosomiasis target the tsetse vector and genes related to olfaction are potential targets in

developing better and improved methods of tsetse control based on traps, baits and olfactory mediated behaviour (Aksoy *et al.,* 2017).

2.3 Vector Control Approaches

Insect vector remains at the centre of control of trypanosomiasis in an integrated approach. Different approaches that have been used to target the tsetse vector such as sterile insect technique (Vreysen *et al.,* 2007), aerial and ground spraying (Cavalloro, 1987; Hilgenkamp, 2006; Clark *et al.,* 2013) clearing of bushes and use of baits, traps and targets developed on the basis of chemoreception (Takken and Knols, 2010). To date, chemoreception is the most promising intervention in controlling tsetse vectors; it plays crucial role in finding suitable mates, habitats and also for host specificity.

2.3.1 Insecticides

Currently, insecticides have been extensively used as the efficient vector control intervention which involves sequential aerosol spraying technique, or selective spraying application of insecticides to animals on which tsetse feed. For instance, ground and aerial spraying of tsetse fly infested areas have been carried out in Nigeria, Namibia, Botswana and Zimbabwe and ongoing in Zambia and Angola (Simarro *et al.,* 2008). The aerosol technique can effectively clear large areas of tsetse flies in a relatively short time, but it is expensive and management intensive (Vreysen *et al.,* 2013). Insecticides used have varied suitabilities depending on their effectiveness and the drugs' impact on the environment. Dichlorodiphenyltrichloroethane (*DDT*) is one of the earliest insecticide that was used in Uganda, while dieldrin was used in Nigeria and Cameroon (Djouaka *et al.,*

2016). Other insecticides that have been commonly used include endosulfan, Malathion and parathion (Vreysen *et al.,* 2007).

2.3.2 Odour-baited traps

Mechanical devices such as traps are also used to kill or weaken tsetse flies through insecticides (Rozendaal, 1997). The traps attract tsetse flies by taking advantage of their primary host-seeking behaviours, visual and olfactory stimulation. The developments of potent attractants in the past as well as the production of second-generation synthetic insecticides has made this form of control technique highly successful (Ryan, 2002; Vreysen *et al.,* 2007).

Trapping of the tsetse flies using odour-baited traps decreases the population of flies and hence reduces human-fly contact. Vector control using insecticide treated traps has the advantage of being managed by the communities for sustainability, especially when materials for the traps could be sourced locally with minimum cost (Esterhuizen *et al.,* 2011). Traps create a visual stimulus to which tsetse flies respond by flying into them (Cardé and Willis, 2008). The efficiency of the traps have greatly been improved by attractants in animals' urine such as acetone, butanol, and 1-acetone-3-a; and this has popularized the use of cow urine baited traps to control *Glossina morsitans morsitans* in Ethiopia, and to control *G. swynnertoni* in the Serengeti and Maasai Mara National parks (Harraca, Syed, and Guerin, 2009).

2.3.3 Sterile insect technique

Sterile insect technique (SIT) involves sterilization of male flies through radiation to make them infertile. The irradiated males are then released into wild populations to compete with natural males. Females are inseminated once in their lifetime, therefore

when mated with sterile males they become unable to produce offspring (Vreysen *et al.,* 2006). In 1994, SIT use in Zanzibar led to the successful eradication of *Glossina austeni* from the island of Unguja (Vreysen *et al.,* 2000). Though costly to sterilize large number of male tsetse flies, SIT has been found to slow down reproduction rate (Vreysen, 2001).

2.3.4 Chemoreception in Vector Control

Chemoreceptors have been targeted in vector control. Recently, natural odorants that interfere with responses to sex pheromones, at the molecular, neurophysiological, and behavioral levels have been unravelled (Renou, 2014; Reisenman, Lei and Guerenstein, 2016). Fine tuning to compounds that affect responses to sex pheromone receptors have been found to lower environmental risks as compared to modulators of the CO_2 receptor complex. This is because pheromone receptors are more specific (Jones *et al.,* 2011).

Of particular interest is the (*E*)-4,8-dimethyl-1,3,7-nonatriene (DMNT), a plant volatile that interferes with the detection of both sex pheromone and host-derived attractants in *Spodoptera littoralis* (Lepidoptera), and thereby reducing attraction to both sex pheromones and host attractants (Knight and Light, 2012). Because DMNT also attracts natural enemies of herbivores it might be a good target for sustainable agricultural pest control (Stenberg *et al.,* 2015).

The first of such studies was performed against receptors from the malaria mosquito, *Anopheles gambiae*, leading to the identification of the synthetic compound VUAA1 (*N*-(4-ethylphenyl)-2-((4-ethyl-5-(3-pyridinyl)-4H-1,2,4-triazol-3-yl)thio)acetamide) as an agonist for the co-receptor Orco itself (Jones *et al.,* 2011).

The use of insect repellents has been a major approach to controlling the spread of insect-borne disease. *N*, *N*-diethyl-*m*-toluamide (DEET) is the active ingredient in most current insect repellents, acting as an airborne spatial repellent and as a contact irritant (DeGennaro, 2015; Leal; 2014). Additional repellents have been developed, such as p-menthane-3,8-diol (PMD), IR3535 and picaridin. Orco antagonists (such as OX1w) are able to block activation by Orco agonists when Orco is part of a heteromeric OR (Kepchia *et al.,* 2017).

Most recent efforts to target ORs for the development of new repellent compounds have involved identification of odorant specificity subunits that recognize behaviorally important odorants (Carey *et al.,* 2010; Wang *et al.,* 2010) and in a few cases, subsequent large-scale ligand screening (Jones *et al.,* 2011; Rinker *et al.,* 2012). However, odorant specificity subunit families are quite divergent across species and there is variation in the odorants and odorant specificity subunits that are important for various species-specific behaviors (Carey and Carlson, 2011; Ramdya and Benton, 2010).

2.4 Chemoreception in Insects

Chemoreception is mediated by membrane proteins including gustatory receptors (GRs), ionotropic receptors (IRs), and odorant receptors (ORs) expressed on gustatory and sensory neurons dendrites. In addition, chemoperception work in conjunction with soluble proteins such as odorant binding proteins (OBPs), pheromone binding proteins (PBPs), odorant digestive enzymes (ODEs) and chemosensory proteins (CSPs) that are expressed in the sensillia lymph that surround the neuron dendrites (Breed and Moore, 2010; Leal, 2013).

Characterization of the *D. melanogaster* ORs indicate that olfactory signalling is distinct in insects (Kaupp, 2010). Drosophila ORs have the protein C terminus at the cell surface and the N terminus within the cell. This is the reversed membrane topology of classical GPCRs in most species (Benton *et al.,* 2006). The Drosophila OR family is encoded by 60 genes. Most members are "tuning" ORs that define the odorant specificity of the olfactory neuron in which they are expressed (Ryan, 2002; Deng *et al.,* 2011).

Chemoreception in insects is divided into olfaction and gustation. Olfaction involves the ability of organisms to detect and discriminate volatile compounds from the surrounding through the numerous pores that fabricate the olfactory tissues (Boron and Boulpaep, 2009). This sensory ability is important for such behaviours as finding mates, locate food sources and detect enemies (Wright and Smith, 2004).

Generally, insect ORs are known to have seven transmembrane domains (Clyne *et al.,* 1999; Mombaerts, 1999) and are expressed in the ORN within hair-like structures known as sensillia (Jacquin-Joly and Merlin, 2004; Vosshall *et al.,* 1999). Antennae is the main olfactory organ in tsetse flies (Liu, 2010) as in other insects too (Ma *et al.,* 2014; Leal, 2013; Pelosi *et al.,* 2006). In *D. melanogaster*, ORs have been reported to be expressed in ORN in the antennae and maxillary palps (Robertson *et al.,* 2003). The power of the olfactory system to discriminate compounds largely depends on the molecular diversity of their OR repertoire (Vosshall *et al.,* 2000).

Almost all Olfactory Receptor Neurons (ORNs) express a canonical co-receptor OR generally named ORCO, which is highly conserved in many divergent insect species, and confers correct function to the other receptors (Sato *et al.,* 2008). This

conservation in diverse insect order and genera suggests that ORCO undergo different selection pressure from other members of the OR gene family (Larsson *et al.,* 2004).

The ORCO also function in the stability of ORs, forming heterodimeric complexes with the tuning ORs (Elmore *et al.,* 2003). It is also essential for transporting and depositing ORs in the specialized sensory cilia on ORN dendrites, thus stabilizing the specific functional ORs in the ciliary membrane (Benton *et al.,* 2006). The binding of the odorant to the membrane receptor activates a series of signalling cascades in the ORN causing action potentials to be generated.

Lots of insects chemoperception has been gleaned from *Drosophila melanogaster* as a genetic tool box that principally focus on the olfactory system (Dobritsa *et al.,* 2003); whereby ORs are expressed following a conserved pattern. Every ORN typically express only one family type of receptor (often up to three conventional ORs), together with the ubiquitous ORCO. In addition, ORNs act in a manner that determines a specific molecular response profile along with the ORCO (Dobritsa *et al.,* 2003).

It has been shown that ORN can also act in a manner that determines a specific molecular response profile along with the ORCO (Dobritsa, *et al.,* 2003). ORNs expressing the same type of OR, converge to a single glomerulus in the antennal lobe (AL), which represents the first olfactory neurophil in the insect brain. These OR family of genes are highly divergent, even among insects of the same species (Vogt *et al.,* 2002).

The ORs determines the sensitivity and specificity of the ORN, which in turn governs innate and learned olfactory behaviours, such as attraction to food and pheromones and avoidance of repellents (Leal, 2013).

2.4.1 Soluble chemosensory proteins

Chemosensory proteins (CSPs) are soluble proteins secreted out of synthesising cells and play diverse role. Based on where they are found, CSPs are involved in CO_2 detection, chemical signal transmission, regenerating legs and chemoperception (Leal, 2013). Numerous CSPs are expressed in the antennae (Sanchez-Gracia *et al.,* 2009), while others are found in the legs (Ozaki *et al.,* 2008) brain, proboscis, and even wings (Zhang *et al.,* 2013). Overall, CSPs in insects play crucial roles such as in transportation of pheromones and larval development, for instance in *Apis mellifera* (Dani *et al.,* 2010), and in cellular regeneration of legs in cockroach *Blatta germanica* (Niu *et al.,* 2016). Generally, CSPs are highly expressed in the lymph of chemosensilla and have great binding activity towards odorants and pheromones (Qiao *et al.,* 2013). In *G. m. morsitans*, five CSPs have been identified and established to be related to female host seeking behaviour (IGGI, 2014; Liu *et al.,* 2012).

2.4.2 Odorant degrading enzymes

Like CSPs, Odorant Degrading Enzymes are present in the sensillum lymph of antennae. They are involved in odor degradation and clearance of the olfactory signal (Leal, 2013). Glutathione-transferase or Cytochrome P450 are intracellular ODEs that have been reported to degrade odors entering the support cells of the sensilla (Younus *et al.,* 2014). In *Phyllopertha diversa*, cytochrome P450 has been found be

responsible with the degradation of pheromones localized in the plasma membrane of the dendrite (Leal, 2013).

2.4.3 Odorant Binding Proteins and Pheromone Binding Proteins

Odorant binding Proteins (OBPs) and Pheromone Binding Proteins (PBPs) are small, soluble proteins with a general composition of between 130-150 amino acids residues and are highly diverse group of olfactory proteins even in insects of the same species (Leal, 2013). Odorant binding proteins bind general odors while PBPs bind pheromones (De Bruyne and Baker, 2008). The OBPs and PBPs bind and transport volatile odors across hydrophilic sensillum to ORs located within ORNs (Zhou, 2010). The insect OBPs have α-helical structure that differs from that of mammalian OBPs which have a β-barrel fold and a carboxy terminal α-helix (Doty, 2015).

Genome sequencing of several insect species has revealed varying number of OBPs. Culex mosquito has 48 OBP genes (Leal *et al.,* 2013), *An. gambiae* has 60 (Qiao *et al.,* 2011), Drosophila has 35 OBP genes (Anholt and Williams, 2010) while *Glossina morsitans morsitans* has 20 OBPs genes (Liu *et al.,* 2010). OBPs have been found to be important in the chemosensory system at the point of odor detection. The OBPs have been reported to solubilize and concentrate odorants in the sensillum lymph, thereby enhancing the signal reaching the ORs (Leal, 2013). Female *An. gambiae* are attracted by volatile compounds released from human skin (Verhulst *et al.,* 2011), where OBPs have been established to be critical in identifying and discriminating hosts by temperature.

Insects have a unique ability to discriminate their hosts, and OBPs are critical for establishing such specificity (Bell and Cardé, 2013). For instance, the savanna species fed on ox, buffalo, kudu and human unlike other Glossina species that have different feeding preference. This phenomenon has been linked to the expression of numerous OBPs in the different antennal sensilla could mean different odorant binding specificities (Torr and Solano, 2010). In *Anopheles stephensi* two OBPs, OBP1 and OBP7, have been found highly expressed in the female antenna and suggested possible roles in female olfactory response and blood feeding behaviours (Sengul and Tu, 2010). However, some OBPs have been found expressed in other body parts other than the main olfactory organ, antenna, and could therefore be involved in non-olfactory functions such as detecting safe landing places (Ogwang *et al.,* 2014).

2.4.4 Gustatory Receptors

Gustatory receptors (GRs) belong to the super-family of G-Protein Coupled Receptors (GPCRs). The *G. m. morsitans* and *D. melanogaster* GRs are clustered into four groups: the CO_2 (Jones, *et al.,* 2007), sugar (Slone, *et al.,* 2007), GR43a-like (Sato, *et al.,* 2011), and bitter classes (Lee, *et al.,* 2009). These receptors share no sequence similarity with vertebrate GRs and their topology is inverted compared to classic GPCRs (Zhang, *et al.,* 2011; Xu, *et al.,* 2012). Chemically-mediated behaviours include detection and feeding, avoidance of harmful and toxic substances and the recognition of mating partners and mating sites are modulated by gustation. Taste is an extremely important sensory modality in the behaviour of insects (Aksoy *et al.,* 2014). Contact chemosensation has been confirmed to play critical roles in discrimination of sweet and bitter taste and in the detection of non-volatile

pheromones. Also, the possibility of GRs participating in metabolism, and other sensory modalities such as the sensing of temperature and water are highly likely (Montell, 2010).

In *Anopheles gambiae,* 76 GRs have been reported and another 79 in *Aedes aegypti* (Pelletier *et al.,* 2010). *D. melanogaster* has 60 GRs (Lee, 2009) whereas *G. m. morsitans* has only 14 (Obiero *et al.,* 2014). Generally, a number of GRs encoded by the *Drosophila, Anopheles* and *Aedes* genomes do not have recognizable homologs in the other insects suggesting significant differences in ligand specificities or sensitivities (Lee, 2009). Many GRs are co-expressed in the same GRNs. However, there have been surprises in *Drosophila* GR repertoire where at least four members of the GR family — including CO_2 sensitive GR21a and GR63a — are expressed in olfactory receptor neurons (Dunipace *et al.,* 2001; Scott *et al.,* 2001).

Gustatory receptors (GRs) are present in the antennae and maxillary palps, and they detect volatile and soluble chemical signals. The signals are then transmitted to the central nervous system and translated to phenotypes (Lee *et al.,* 2009).

2.4.5 Ionotropic Receptors

Drosophila has also provided a genetic model that principally focus on the olfactory system (Dobritsa *et al.,* 2003). Numerous antennal neurons express chemosensory receptor genes known as ionotropic receptors, IRs. The IRs are not related to ORs in insects, but have evolved from ionotropic glutamate receptors (iGluRs), a conserved family of synaptic ligand-gated ion channels (Rytz *et al.,* 2013). The expression analysis of the IR genes closely correspond the spatial distribution of the

physiological coeloconic classes. This provides evidence that the receptors define the odor response profiles of these neurons (Boulpaep and Boron, 2005).

Studies on the olfactory responses has revealed that IR neurons are generally more narrowly-tuned and less sensitive than the other receptor expressing neurons. This is perhaps why IR neurons are slower to respond but adapt less quickly (Galizia and Lledo, 2013). Like the OR neurones, the IR neurones specialise in detecting distinct classes of odors; the strongest IR ligands are recognized only weakly or not at all by ORs, and the strongest OR ligands (predominantly esters, alcohols, and ketones) not stimulating any IR neurons. However, a key property of the IRs is with their specific targeting to the ciliated dendrites of ORNs (Rytz *et al.,* 2013).

2.4.6 Odorant Receptors

Odorant receptors are expressed in ORNs and are important in hormonal regulation, neurotransmission and photoreception (Alcedo, Maier and Ch'ng, 2010). ORs are composed of a protein complex that has a transmembrane core important in evoking odour recognition in the sensory neurons (Sato *et al.,* 2008), a crucial step for external signal transduction (Gudermann *et al.,* 1995).

The repertoires and characterization of 60 OR genes have been reported in the model fruit fly *D. melanogaster* (Vosshall *et al.,* 2000), and also in mosquitoes *Anopheles gambiae* (79 ORs) and *Aedes aegypti* (131 ORs) (Pelletier *et al.,* 2010). The Southern house mosquito, *Culex quinquefasciatus*, has the largest repertoire of ORs (181 ORs) of all dipteran species whose genomes have been hitherto sequenced (Arensburger *et al.,* 2010). A further significant insight into the mosquito sense of smell was recently obtained by the functional characterization of 50 ORs in *An.*

gambiae (Scialò *et al.,* 2012). Each OR is expressed within a spatially restricted sub-types of ORNs. One exceptional receptor, ORCO, is expressed in most ORNs of both the adult and larval stages (Clyne *et al.,* 1999; Vosshall *et al.,* 2000; Kreher *et al.,* 2005).

Glossina morsitans morsitans has 46 ORs, smaller number compared to other Diptera even though some specific receptors appear in multiple copies (Obiero *et al.,* 2014). Like in other insects (Bohbot *et al.,* 2007), the *G. m morsitans* ORCO has been found to be conserved amongst other OR genes (Obiero *et al.,* 2014). This conservation has been linked to the critical roles that ORCO play in modulating responses of the other receptors (Sato *et al.,* 2008).

Glossina morsitans morsitans has numerous CO_2 sensitive chemoreceptors capable of utilisation as the vector control tools that exploit such unique adaptations. However, the information on the expression of these ORs in tsetse have not been reported; this will form the aim of this study — to determine the expression profiles of the ORs in *G. m. morsitans.*

2.5 Real time PCR for Gene Expression

Real-time polymerase chain reaction (RT-PCR) is often used to determine the expression of target genes. However, conducting an RT-PCR analysis is challenging. This is due to difficulties in establishing the amount and good quality of the starting material, the quality and quantity of cDNA, and enzymatic efficiencies during the reaction. An appropriate reference gene is used to stabilize and thus normalize the reactions in RT-PCR analysis that arise due to experimental variability. Despite these challenges, the accuracy, sensitivity, and reproducibility of the RT-PCR process

offers huge potential for high throughput application (Derveaux, Vandesompele, and Hellemans, 2010).

2.6 Bioinformatics Analyses

2.6.1 Multiple Sequence Alignment

Multiple Sequence Alignment (MSA) is the alignment of three or more biological sequences (protein or nucleic acid) of similar length. The output of such alignment enables inference of homology and the evolutionary relationships possible between the sequences studied (Pearson, 2013).

There are a number of MSA tools available. Clustal Omega tool (Sievers and Higgins, 2014) uses seeded guide trees and HMM profile-profile techniques to generate alignments. Kalign (Deorowicz, Debudaj-Grabysz, and Gudyś, 2014) is very fast MSA tool that concentrates on local regions and is suitable for large alignments. MAFFT (Katoh and Standley, 2013) uses Fast Fourier Transforms that is suitable for medium-large alignments and MUSCLE (Edgar, 2004) is accurate MSA tool, and particularly good with proteins for medium alignments. On the other hand, T-Coffee (Di Tommaso *et al.,* 2011) has consistency and attempts to mitigate the pitfalls of progressive alignment methods. It is however suitable for small alignments.

Multiple sequence alignments reveal sequence homology and phylogenetic relationships that can be used to assess the sequences' shared evolutionary origins (Pearson, 2013). Multiple sequence alignment is often used to assess sequence conservation of protein domains, structures, and even individual amino acids or nucleotides. And can be helpful in detecting historical and familial relations between

sequences of proteins or amino acids and determining certain structures or locations on sequences (Marti-Renom, Madhusudhan and Sali, 2004).

2.6.2 Transmembrane Domains

Predicting tansmembrane domains (TMDs) utilises Hidden Markov Model that can be represented in the simplest dynamic Bayesian network and assumes the system under investigation to be in a hidden state. Hidden refers to the state sequence through which the model passes, not to the parameters of the model (Zhu and Wang, 2015).

Hidden Markov Models (HMMs) are an extremely versatile statistical representation that can model protein sequences in many ways, depending on what features of the protein are represented by the Markov states (Bystroff and Krogh, 2008). For protein structure prediction, states have been chosen to represent either homologous sequence positions, local or secondary structure types, or transmembrane locality. The resulting models can be used to predict common ancestry, secondary or local structure, or membrane topology by applying one of the two standard algorithms for comparing a sequence to a model (Zaki and Bystroff, 2007).

2.6.3 Conserved Domains

Domains are distinct functional and/or structural units of a protein. Function and structure often coincide rather often in an independently folding unit of a polypeptide chain (Andreeva, 2011). Domains are often the recurring sequence or structure units, which may exist in various sequences. In molecular evolution such domains may have been utilized as building blocks, and may have been recombined in different arrangements to modulate protein function (Marchler-Bauer and Bryant, 2004).

The aim of conserved domain curation is to provide insights into how patterns of residue conservation and divergence in a family relate to functional properties, and how such entwine by providing detailed information that help understand the sequence, structure, and function relationships (Alexandersson *et al.,* 2014). Conserved domain databases have, therefore, provided elaborate information that supplement and enrich the traditional multiple sequence alignments that form the foundation of domain models. Such information include 3-dimensional structures and conserved core motifs, conserved features, and even phylogenetic organization (Chakrabarti *et al.,* 2006).

2.6.4 Gene Ontology

Gene Ontology is critical for interpretating large-scale molecular biology omics experiments. Omics experiments measure gene products (RNA and proteins), variation in the DNA sequence of genes, or small molecules metabolized by proteins. Subsequently, relating them all to gene function. GO attempts to reconcile representation of gene and gene product features across all species (Alexandersson *et al.,* 2014).

2.6.5 Protein Structure Analysis

Proteins fold into one or more specific spatial conformations driven by a number of non-covalent interactions such as hydrogen bonding, ionic interactions, Van der Waals forces, and hydrophobic packing (Sung, 2015). Such folding is critical for determing the role of protein in an organism. The determination of protein's three-dimensional structure is paramount in understanding the functions of proteins at a molecular level (Roy, Kucukural and Zhang, 2010). Such structural biology studies

employs techniques such as X-ray crystallography, NMR spectroscopy, and dual polarisation interferometry (Feng, Pan and Zhang, 2011).

The complexities of protein structure make the elucidation of a complete protein structure extremely difficult even with the most advanced analytical equipment (Biasini *et al.,* 2014). Nuclear magnetic resonance (NMR) or X-ray crystallography provide high-resolution analysis of the three-dimensional structure of a protein analysis in a more complete manner (Faraggi *et al.,* 2018). To determine the three-dimensional structure of a protein by X-ray diffraction, a large, well-ordered single crystal is required. X-ray diffraction allows measurement of the short distances between atoms and yields a three-dimensional electron density map, which can be used to build a model of the protein structure (Miao *et al.,* 2015). The use of NMR to determine the three-dimensional structure of a protein has some advantages over X-ray diffraction in that it can be carried out in solution and thus the protein is free of the constraints of the crystal lattice (Fan and Lane, 2016).

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study Area and Preparation of Insects and Tissue

3.1.1 Study Site

This study was carried out at Jomo Kenyatta University of Agriculture and Technology (JKUAT), Biochemistry Department Laboratory. The *G. m. morsitans* flies (newly emerged larvae, 1 day old pupae and 21 day-old adults) used in this study were sourced from Biotechnology Research Institute of Kenya Agricultural Livestock Research Organization (BRI-KALRO) in Muguga, Kenya.

3.1.2 Insect Rearing

The flies were reared under standard laboratory conditions (temperature 25 ± 1°C; relative humidity 75 ± 10%; fed on sterilized pig blood after every 24 hours using an artificial membrane) in the insectary at BRI-KALRO (Moloo, 1971).

3.1.3 Handling of Samples

The adult flies were collected in the morning and transported to JKUAT Biochemistry Department laboratory for utilisation in RNA work using cages labelled as either males or females. Once the pregnant adult females gave birth, the newly born larvae were placed in a sterile plastic pyrex tube and fully immersed in a steel flask of liquid nitrogen. Some larvae were left to develop in to pupae after about four hours of parturition and were wrapped in sterilised piece of aluminium foil before transportation.

3.2 Determination of Levels of OR Genes in *G. m. morsitans* samples

Expression analysis of odorant receptor genes in *G. m. morsitans* samples; both male and female antennae and leg tissues, larvae and pupae began with the extraction of mRNA, quantification, cDNA synthesis and amplification by conventional PCR and subsequently expression analysis by gel electrophoresis.

3.2.1 Sample Preparation for total RNA Extraction

Total RNA was extracted from a pool of 10 larvae and 10 pupae and from 100 pairs antennae and 100 legs dissected from adult male and female *G. m. morsitans* flies. Prior to total RNA extraction, the flies were frozen at –20°C for 10 minutes before antennae were dissected under a microscope and legs were removed from adults and placed alongside larvae and pupae in separate 350 µL of PBS (pH 7.4) buffer for RNA extractions following TRIzol reagent protocol (Rio *et al.,* 2010).

3.2.2 Extraction of Total RNA

The samples were homogenized to a fine paste using RNase free micro pestles and the contents mixed thoroughly by vortexing for 1 minute. The lysate was then centrifuged at 12,000 rpm at 4^0C for 5 minutes and 250 µL of the supernatant transferred into a clean RNase free micro centrifuge tubes containing 750 µL of TRIzol LS (Invitrogen, Carlsbad, CA, USA) reagent (Moloo, 1971). The mixture was then capped and left to stand at room temperature for 10 minutes to allow for complete dissociation of nucleoprotein complexes before adding 200 µL of chloroform. It was again capped thoroughly and left to incubate at room temperature for further 10 minutes. The samples were then centrifuged at 12,000 rpm at 4^0C for another 10 minutes. The colourless upper aqueous phase, containing the RNA, was pipetted out carefully into a clean RNase free micro centrifuge tube. RNA was

precipitated out by adding 1 µL of Glycogen and 500 µL of Isopropanol followed by an overnight incubation at -20^{0}C (Sambrook and Russel, 2001). After the overnight incubation, the samples were then centrifuged at 12,000 rpm for 10 minutes at 4^{0}C. The liquid was discarded with caution such that the pellet was not dislodged and lost. Afterwards, the RNA pellet was washed by gently inverting the tube containing the pellet in 50µL of 75% ethanol. The tubes were centrifuged at 12,000rpm at 4^{0}C for 10 minutes and the supernatant discarded. A further 30 seconds centrifugation of the samples was done to remove excess liquid. Subsequently, the pellet obtained was re-suspended in 12 µL of nuclease free water and then allowed to stand at room temperature for 5 minutes before gently tapping the sides of tube to dissolve it completely. The sample was then treated in 2 µL 10X reaction buffer and 2 µL Dnase, to remove genomic DNA followed by incubation at 37^{0}C for 37 minutes. 1 µL of 50mM EDTA was then added followed by incubation at 65^{0}C for 10 minutes (Sambrook and Russel, 2001).

3.2.3 Determination of RNA yield and quality

The extracted RNA was quantified using Nanodrop 2000 spectrophotometer reader (Biospec-mini, Shimadzu Corporation, Tokyo). The machine was blanked with 2 µL nuclease free water before 2 µL of RNA sample loaded. After quantification, the data generated was recorded for further analysis.

3.2.4 Determination of RNA integrity

A 1.2% formaldehyde agarose gel was prepared to ascertain the integrity of the extracted RNA. This was done by mixing 2ml of 10X MOPS buffer; 3-(N-morpholino) propanesulfonic acids (209.3 gmol^{-1}), sodium acetate (136.1 gmol^{-1}), EDTA (372.24 gmol^{-1}), formaldehyde (37%) in 1 litre of Diethyl pyrocarbonate

(DEPC)-treated H_2O, 18ml of nuclease free water and 1.8ml of 37% formaldehyde in a conical flask containing 1.2g of agarose. The mixture was heated in a microwave for 30 seconds and left to cool to about 40ºC before 0.5 µL of Ethidium Bromide (EtBr) (0.5µg/ml) added and allowed to polymerize in a gel slab with combs in place. 5 µL of RNA sample was then mixed with 2 µL of the dye and loaded into the wells alongside a DNA standard and left to run at 100V 80mA 50W for 10 minutes. The images captured were stored for later analyses.

3.2.5 Synthesis of cDNA by Reverse Transcription

Reverse transcription from the respective RNA (20ng/µL) samples (larvae, pupae, antennae and legs) was carried out using cDNA Synthesis Kit (Fischer, ThermoScientific, USA) following the manufacturer's specifications, with a final reaction volume of 20 µL. Into a clean RNase free PCR tube, 4 µL RNA (20ng/µL) template, 1 µL oligo $(dT)_{18}$ and 7 µL of nuclease free water were added, respectively. These components were gently mixed and briefly centrifuged before being incubated at 65°C for 5 minutes. After incubation, the tube was chilled in ice and spanned gently. Then, the following components were added in their respective order: 4 µL ×5 Reaction buffer, 1 µL Ribolock RNase Inhibitor, 2 µL 10mM dNTP mix, and 1 µL Rivert Aid M-MULV-RT. These were then mixed gently before the mixture was incubated at 42°C for 60 minutes and the reaction terminated by heating at 70°C for 5 minutes and finally chilled on ice.

3.2.6 Amplification of synthesised cDNA

Specific primers for *Glossina morsitans morsitans* ORs retrieved from VectorBase (Giraldo-Calderón *et al.,* 2015) were designed using Primer3 v4.0 algorithm (Untergasser *et al.,* 2012; Castresana, 2000). The primers, with annealing

temperatures ranging between 55.0 ºC to 58.0ºC, and length of between 120 and 200bp, were synthesized by Inqaba Biotech, South Africa (Table 3.1).

Table 3.1: Quantitative Real Time Primers designed for *Glossina morsitans morsitans* Odorant Receptors (ORs) Genes.

	Primers Sequences 5' to 3'	
Gmm OR	**Forward Sequence**	**Reverse Sequence**
GAPDH	TAAAATGGGTGGATGGTGAGAGTC	CTACGATGAAATTAAGGCAAAAGT
OR1	CTGTTCGCTCTATCTCATGC	GAAATAGAGCTCCGGAATCG
OR2	TCCCGCTTGGTATGGTTTAG	AGAGACCCACAGAACGCACT
OR3	TCACCGGAGCGTATCCTAAT	ATATGGCCGATCAAGACGAC
OR4	GGCGATTGTTGGGTATCG	AAACTGGTCGCTATGCTGGT
OR5	TGGCTATTGGTATGGAGGTC	GATACCGCCCGCTAACATAG
OR6	CAGGCGGCCATCAATATC	CGGTAATTGCACAGTGACG
OR7	CTGCAACCATCATCACAACC	GTCCGCAGATGATACGATGA
OR8	ATTCGAAGCGACCACTTACC	CCAGCTCGTTGGTGAAATAG
OR9	GCGCACTTCAGAGGAGATG	CGACCGGCATATACAAGGTT
OR10	TACAGCTCTCGCAGTTCTGG	AGCGAAAGAACGCAGTAAGC
OR11	TCTATGGAGGCATGTCCTTG	GGCCAAAACCCAGATACTCA
OR12	ACGATGTAACCGATGTGCTG	CAGATTAGCTGCTCGTTTGC
OR13	TACACGGGTCTGACCACATT	ACAGGTCCCGTTGTAATTCG
OR14	GCGTTGTACGTATGGGATTC	TTACCGCCGTTTTCTTGG
OR15	GGTAACGCGACGTTCATTC	CCCTGAGGTATCGGCAAAT
OR16	ACTCAATGGCCAGTCTGACA	CACCGAGATCGTCTTGGATA
OR17	CTTAGTCGCAACCATGTGGA	TGATGAGTGCCGTGTAAAGG
OR18	GCTTGGAAGCTGAAGCAAGT	GATCATCATTCGGACCAAAA
OR19	CCAAGCTTAGCCCAATGC	AAGGAGACATTCCGCTCTGA
OR20	GGAGACCAAAACTCGCAGAT	CCACGGAACGCAACTTTAG
OR21	GCCCGATAGAACGTTGGTT	ACATCCAATAAGCCGAGGAG
OR22	CAAACGTGCCGCTTTTCT	GGTATGAATGTGGCTCTTCG
OR23	TCTAACGGCGGGATCTTCTA	CCATACGTGCTCCCTTACCT

OR24	CTGGGCATAGAACCTTACGC	GTGGCCTCCAGGAAATTGT
OR25	TTACGATGGCAAATTGTTGG	TTCCGATTCGGTCAATATCC
OR26	CTCGTTGGCTGTACGCATTA	GCCAATCATCACAAGTACCG
OR27	CTGCATATGGCACGGAAA	GTATGCTGCCGTCATACCAT
OR28	AGAGCTCAAAGGCCAGCTC	CGTTGTCAGTGGCTACAGGA
OR29	TGACGAGCATCCTTGTGAAC	ATAATGGTGCCAGGCTGAAC
OR30	GTTCAGCCTGGCACCATTAT	GCAGCAAAGGCTTGAGTTG
OR31	TCAACACAAGCACCGCTTAC	CCATCAGCACCAACAAACC
OR32	GACGTAGGCTGGAATGCAC	TGCCGCAAGTGTAGGTAATG
OR33	AAAAGAATTGGCGACGCTAT	AATTGAGTCGATACCCTCCT
OR34	CGAAGAGCCACATTCATACC	TGGATGGTCAAAGCTGTACG
OR35	GTTGCCGGTCTGTGTAATGA	GGAATGCCTCGAACTTATCG
OR36	TACCGTTATGCTGTGCGTTC	TACGGTAAGCAGCATCAGGA
OR37	CGCAGAATCCTTTGGGTAGT	ATTTTCGCAGGAGTCTGAGG
OR38	CCACCGCTGTTAACTCATCA	GAAGATTCCGCAAGCAAGAC
OR39	GCCGATCGTTTAAGTGAAGC	TCGCATACGTCGCTGTAAGT
OR40	GTGGCGCTGTGTTTACGATA	CTCACCACTCTTGCTCGTTG
OR41	AGGATAGAGGCGCACGATAC	CCCGTTTGCCTGTTATTAGC
OR42	CACGATACGAATACGCCTTG	CCCGGTATCAACGCAGTTA
OR43	GGCACACAATACGAATACGC	GGATCAATTCCCGGTATCAG
OR44	CGATCCCAATACCCATACG	CGCTTTCATTGGTCTCCTC
OR45	CAGTGATCGTATGGCCGTTA	GCTGCAAAGTTTCCATAGGC
OR46	CGCGAATGTAGCGATGATAG	GTGAAGGAAATAGCCCATCG

The integrity of *G. m. morsitans* antennae, leg, pupae and larvae cDNA was validated through conventional PCR amplification carried out in a 96-well Applied Biosystem (ABI Gene AMP 9700) PCR system along with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) specific primers (Fwd 5'-TAAAATGGGTGGATGGTGAGAGTC-3' and Rev 5'-CTACGATGAAATTAAGGCAAAAGT-3') as internal control (Attardo *et al.,* 2006).

The cDNA synthesised from the larvae, pupae, adult's antennae and legs were used as templates for the PCR amplification in a total volume of 20 µL containing: 6 µL nuclease free water, 10 µL ×2 Dream Taq master mix, and 1 µL each of the forward and reverse primers and 2 µL (20ng/µL) of cDNA sample. The negative control was prepared from the cDNA samples that had been made with all the reagents and amplification requirements met except the inclusion of Reverse Transcriptase enzyme (Sambrook and Russel, 2001). The PCR conditions were: an initial denaturation temperature of 95ºC for 5 minutes, and 35 cycles of subsequent denaturation temperature at 94ºC for 30seconds, annealing temperature of 55ºC for 45 seconds, extension temperature of 72ºC for 1 minute, and one last cycle of final extension at temperatures 72ºC for 7 minutes. The amplified RT-PCR products were analysed on a 2% agarose gel.

3.2.7 Gel Electrophoresis of PCR products

The generated PCR products from cDNA amplification were analysed using 2% agarose gel electrophoresis. The gel was prepared with 0.4g agarose in 20ml of buffer, heated and after cooling to about 50ºC, 0.5 µL Ethidium bromide (EtBr) (0.5µg/ml) was added. About 10 µL aliquots of the PCR products were subjected to gel electrophoresis in 2% agarose NA (Pharmacia Biotech AB, Uppsala, Sweden) at 100Volts 80mA and 50W in ×1 TBE buffer, for 10 minutes. DNA standard of 100bps (Biolabs), was loaded in the first well of the gel slab to act as a molecular size marker. The resulting gel images were viewed on a UV trans-illuminator and photo images taken (Dorak, 2007). The images generated were saved and analysed before conducting RT-PCR.

3.3 Real Time Polymerase Chain Reaction

The OR genes that successfully amplified during primer validation were further quantified using RT-PCR. Quantification was done using 2×Maxima SYBR Green/ROX master mix (Fermentas, Thermoscientific, UK) following the manufacturer's instructions with a Stratagene Mx3000P qPCR system (Agilent Technologies Ltd, Cheshire, UK). Each reaction amounted to 12.5µL and consisted of the following final concentrations; 6.25µL 1×Maxima SYBR Green/ROX master mix, 0.5µL of 2ng/ µL of cDNA template, 0.5µL each of 0.3µM forward and reverse primers. The reaction mixture was topped up using 4.25µL nuclease free water, mixed gently without creating bubbles, centrifuged briefly and placed in the thermocycler. The following thermocycling conditions were used; one cycle of initial denaturation at 95°C for 10 minutes, 40 cycles of; 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 45 seconds. Data acquisition, and exportation on to excel file, was automatically performed by the thermocycler during the extension step. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used for initial normalization of the template cDNA. Tests were done in triplicates.

3.3.1 Analysis of RT-PCR Data

The expression level of OR genes in the pupae, larvae, antennae and leg tissues of the male and female *G. m. morsitans* was calculated by delta delta Ct (ddCt) method in which the discrepancy between the Ct for the ORs and GAPDH was first calculated to normalize the variation in the amount of cDNA in each reaction (Livak and Schmittgen, 2001; Makovets, 2013). Descriptive statistics (means and plots of averages) were done using a statistical tool — R commander (Braun and Murdoch,

2007). Prior to ANOVA (Kim, 2014), data were log transformed using the formula log10 (ddCt + 0.001) and multiple comparison was performed using Tukey's test (Silverman, 2018) where $p<0.05$ was considered significant.

3.4 Comparative Bioinformatics Analyses

The published *G. m. morsitans* OR protein sequences were used to search and retrieve orthologs in mosquitoes (*Anopheles gambiae, Aedes aegypti* and *Culex quinquefasciatus)*, other *Glossina* species (*G. pallidipes, G. fuscipes fuscipes, G. brevipalpis,* and *G. austeni*) in VectorBase (Giraldo-Calderón *et al.,* 2015) and *Drosophila* in FlyBase (Gramates *et al.,* 2017). These orthologs were identified using tBLASTx algorithm with cut-off e-value of <1.0e-05 with low complexity filter. The *Glossina* OR protein sequences were screened in the NCBI's conserved domain database (CDD) (Marchler-Bauer *et al.,* 2017) for conserved domains, at e-value <1.5e-05. The *Glossina* OR amino acid sequences together with identified orthologs in *Drosophila* and mosquitoes were aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) tool (Edgar, 2004). Aligned sequences were viewed using Jalview (Waterhouse, 2009) tool and manually edited to remove gaps. The *G. m. morsitans* protein sequences were screened for the seven transmembrane domains (TMDs) using TMHMM (V4.0) (Krogh *et al.,* 2001). Putative functions of the *G. m. morsitans* ORs were inferred using gene ontology database (Ashburner *et al.,* 2000; Consortium, 2004). Protein three-dimensional structure prediction and determination of binding sites was performed using RaptorX online tool under default parameters (Morten *et al.,* 2012).

CHAPTER FOUR

RESULTS

4.1 Expression of OR Genes in *G. m. morsitans* samples

4.1.1 RNA Quality Analysis

The RNA extracted from *G. m. morsitans* larvae, pupae, adults' antennae and legs, was quantified using Nanodrop 2000 spectrophotometer reader (Biospec-mini, Shimadzu Corporation, Tokyo), all the RNA showed $A_{260/280}$ nm absorbance ratio of between 1.753 and 2.148 with the highest recorded in larvae and lowest in male legs (Table 4.1). The highest concentration was for pupae (2979.2 ng/mL) and lowest concentration was for male legs (650.65 ng/mL) (Table 4.1).

Table 4.1: Nanodrop readings for mRNA extracted from *G. m. morsitans* samples

Sample Name	Concentration (ng/mL)	Absorbance (260nm/280nm)
Larvae	2665.7	2.148
Pupae	2979.2	2.062
Male Antennae	1221.5	1.862
Male Legs	650.65	1.753
Female Antennae	1457.6	1.777
Female Legs	911.71	1.767

RNA concentrations were given in (ng/mL) and quality (absorbance at wavelength 260nm/280nm) in the six *Glossina morsitans morsitans* study samples. The highest concentration was recorded in pupae at 2979 ng/mL and absorbance in larvae at 2.1.

The RNA integrity was analysed by formaldehyde agarose gel electrophoresis. Visual inspection of the gel images was used to confirm the presence of bright, clear bands corresponding to the 28S and 18S ribosomal subunits, providing qualitative assessment of the rRNA integrity. The larvae, pupae, adults' antennae and leg samples all had the 28S and 18S RNA bands, confirming that the RNA was intact and of good quality (Figure 4.1).

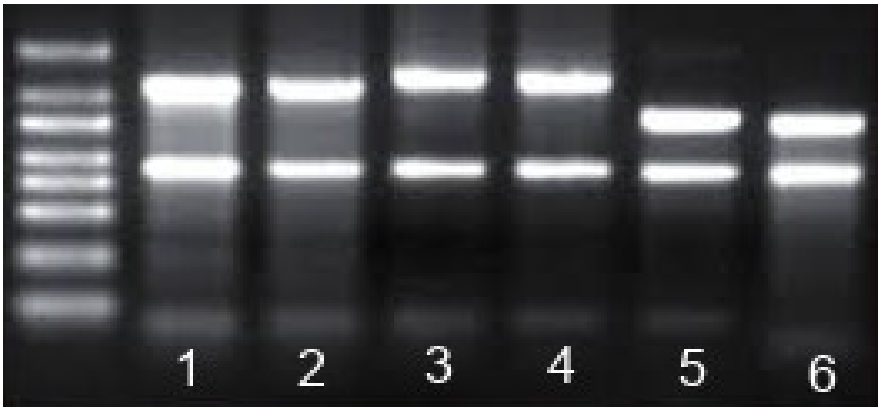


Figure 4.1: Gel image showing the integrity of the total RNA extracted from *G. m. morsitans*. The 18S and 28S ribosomal bands present indicated good integrity of the RNA samples. The first lane represent molecular marker and **1:** Male leg, **2:** Male antennae, **3:** Female leg, **4:** Female Antennae, **5**: Larvae, **6:** Pupae.

4.1.2 Amplification of synthesised cDNA

Amplification of the cDNA reverse transcribed from total RNA extracted from *G. m. morsitans* samples (adults' antennae and legs, larvae and pupae) was carried out with each of the 46 OR genes using conventional PCR. Forty-five OR genes were amplified in larva, forty in pupae, forty-four in female antennae, forty-three in female

legs, forty-five in male legs while forty were detected in the male antennae. The positive control GAPDH was amplified in all the test samples while the negative control did not show any amplification (Figure 4.2a to 4.2x).

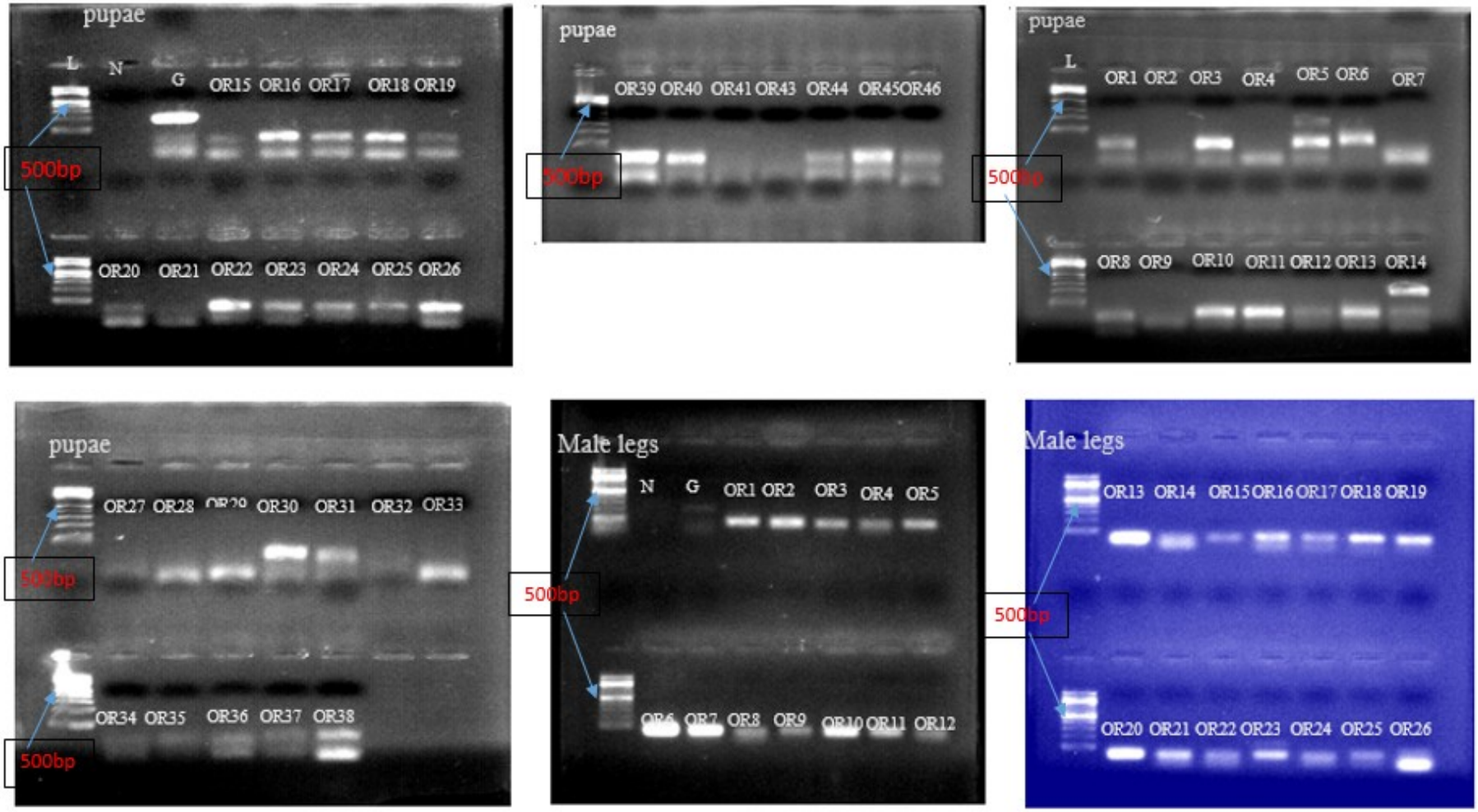

pupae
L
N
G
OR15 OR16 OR17 OR18 OR19
500bp
OR20 OR21 OR22 OR23 OR24 OR25 OR26
pupae
OR39 OR40 OR41 OR43 OR44 OR45OR46
500bp
pupae
L
OR1 OR2 OR3 OR4 OR5 OR6 OR7
500bp
OR8 OR9 OR10 OR11 OR12 OR13 OR14
pupae
OR27 OR28 OR30 OR31 OR32 OR33
500bp
OR34 OR35 OR36 OR37 OR38
500bp
Male legs
N
G
OR1 OR2 OR3 OR4 OR5
500bp
OR6 OR7 OR8 OR9 OR10 OR11 OR12
Male legs
OR13 OR14 OR15 OR16 OR17 OR18 OR19
500bp
OR20 OR21 OR22 OR23 OR24 OR25 OR26

Figure 4.2(A): Conventional PCR amplification of OR genes in Pupae. Forty odorant receptor genes were amplified.

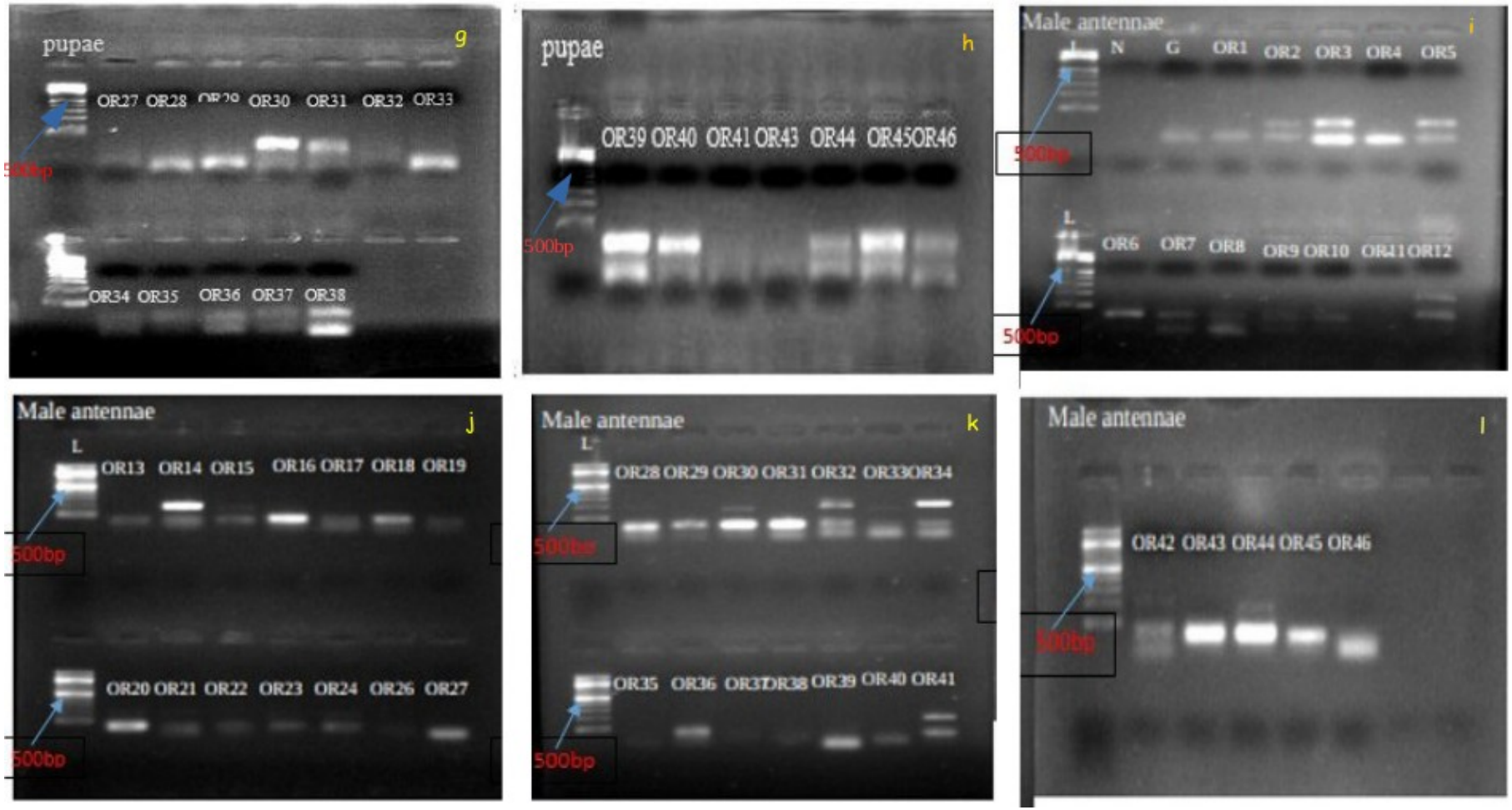

pupae
g
OR27 OR28 OR30 OR31 OR32 OR33
500bp
OR34 OR35 OR36 OR37 OR38
pupae
h
OR39 OR40 OR41 OR43 OR44 OR45OR46
500bp
Male antennae
i
L
N G OR1 OR2 OR3 OR4 OR5
500bp
L
OR6 OR7 OR8 OR9 OR10 OR11 OR12
500bp
Male antennae
j
L
OR13 OR14 OR15 OR16 OR17 OR18 OR19
500bp
OR20 OR21 OR22 OR23 OR24 OR26 OR27
500bp
Male antennae
k
L
OR28 OR29 OR30 OR31 OR32 OR33OR34
500bp
OR35 OR36 OR37 OR38 OR39 OR40 OR41
500bp
Male antennae
l
OR42 OR43 OR44 OR45 OR46
500bp

Figure 4.2(B): Forty ORs in pupae (e-h) and forty were detected in the male antennae (i-l). **L:** 100bp marker, **N:** Negative control; **G:** GAPDH

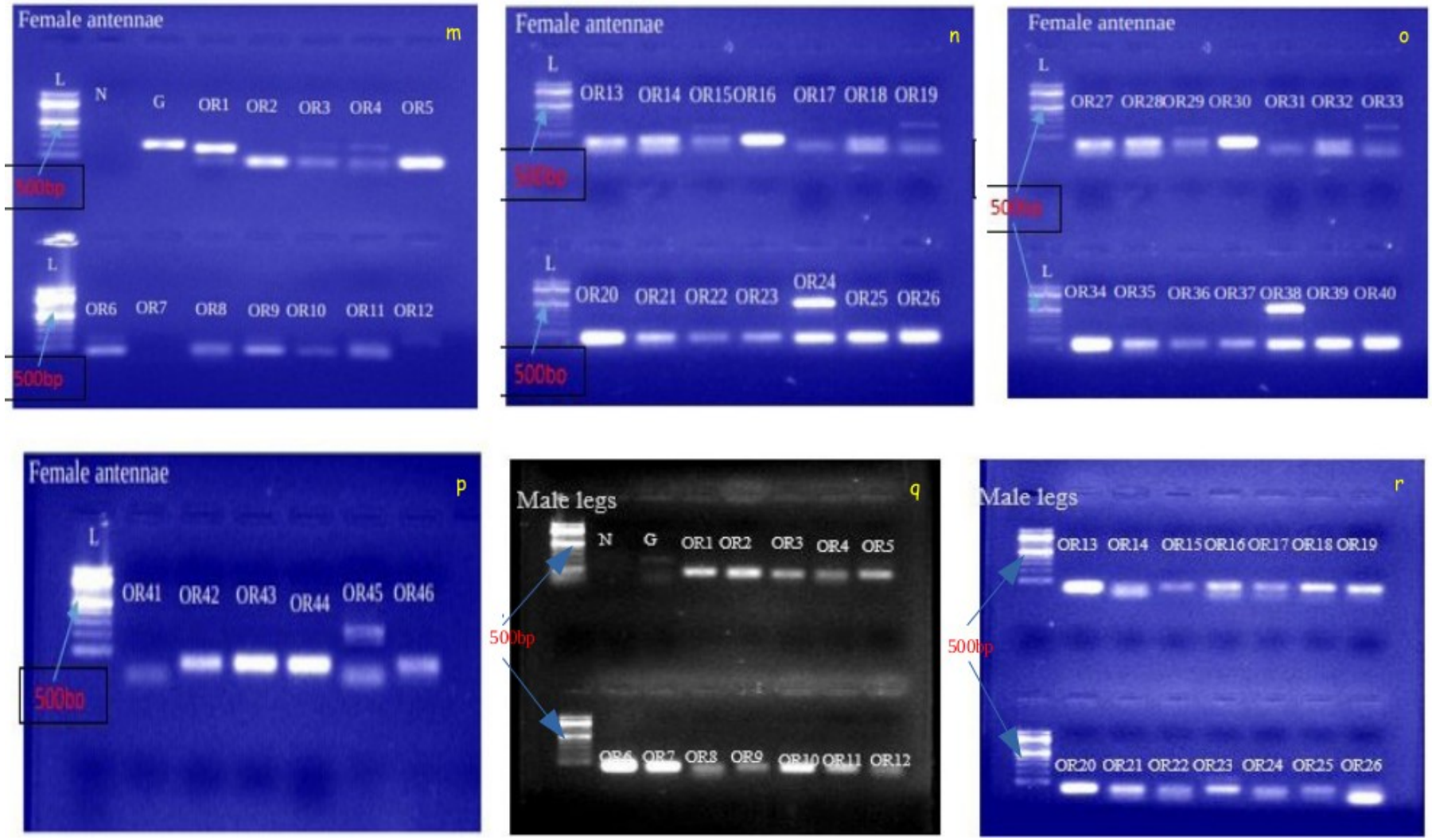

Female antennae
m
L
N
G
OR1
OR2
OR3
OR4
OR5
500bp
L
OR6
OR7
OR8
OR9
OR10
OR11
OR12
500bp
Female antennae
n
L
OR13
OR14
OR15
OR16
OR17
OR18
OR19
500bp
L
OR20
OR21
OR22
OR23
OR24
OR25
OR26
500bp
Female antennae
o
L
OR27
OR28
OR29
OR30
OR31
OR32
OR33
L
OR34
OR35
OR36
OR37
OR38
OR39
OR40
Female antennae
p
L
OR41
OR42
OR43
OR44
OR45
OR46
500bp
Male legs
q
N
G
OR1
OR2
OR3
OR4
OR5
500bp
OR6
OR7
OR8
OR9
OR10
OR11
OR12
Male legs
r
OR13
OR14
OR15
OR16
OR17
OR18
OR19
500bp
OR20
OR21
OR22
OR23
OR24
OR25
OR26

Figure 4.2(C): Forty-four OR genes amplified in female antennae (m-p). **L:** 100bp marker, **N:** Negative control; **G:** GAPDH

Figure 4.2(D): Forty-five in male legs (q-t), and forty-three in female legs (u-x).

Figure 4.2 (A-D): Agarose gel (2%) electrophoresis of the amplified *G. m. morsitans* OR genes in all the six samples.
Gel images of *G. m. morsitans* OR genes that were amplified using conventional PCR in the samples. Some gels had multiple bands, for instance OR3 and OR34 in male antennae and OR4, OR19, OR45 in female antennae. This was due to the formation of primer dimers as a result of low quantities of cDNA during amplification. **L:** 100bp molecular marker (New England Biolabs), **N:** Negative control; a

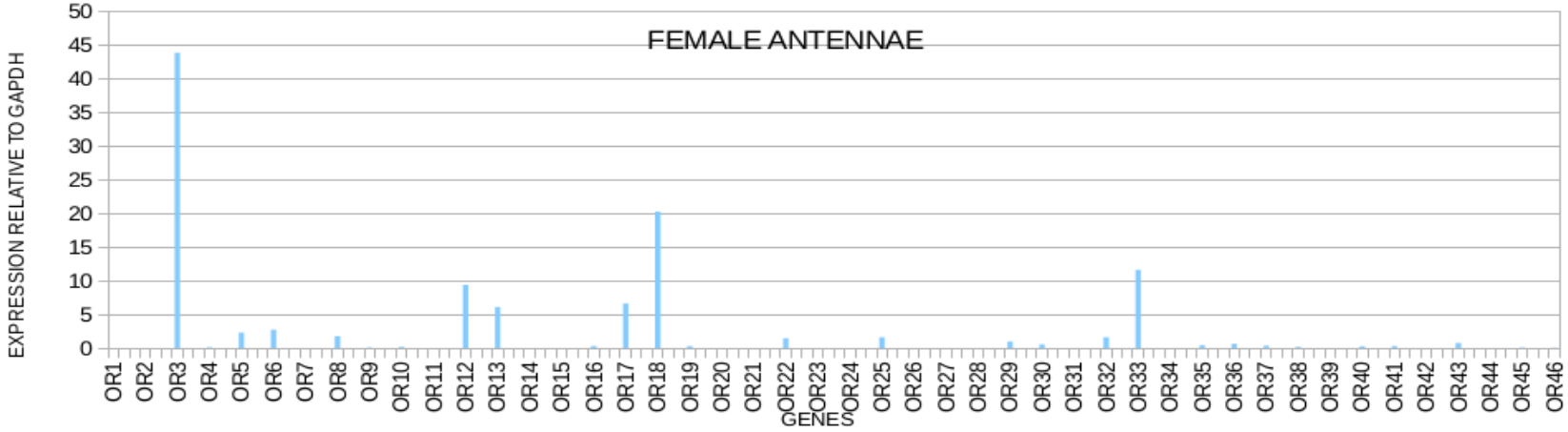
reaction tube containing all the PCR ingredients but lacking the polymerase enzyme, **G:** GAPDH. The positive control (GAPDH) was amplified in all the samples.

4.2 Expression of *Glossina morsitans morsitans* OR genes

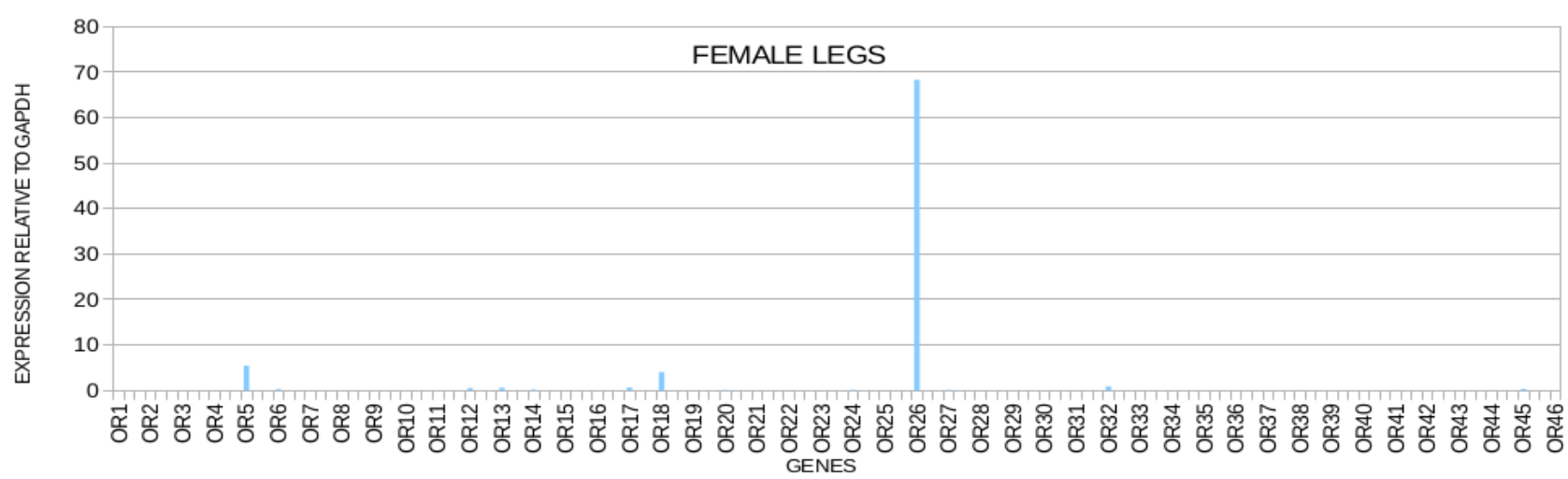
The relative amounts of transcripts generated for each OR gene estimated by qPCR technique are presented in Figure 4.3. Overall, the *G. m. morsitans* OR genes were enriched in the antennae compared to the leg tissue (Figure 4.3). The OR gene expression in female and male *G. m. morsitans* antennae were almost the same amount. The female *G. m. morsitans* revealed 13 ORs expressed in the antennae of which 6 ORs were expressed at high levels while 7 ORs were expressed at low levels. Similarly, male *G. m. morsitans* antennae revealed 12 ORs expressed of which 9 ORs were highly expressed and 3 ORs expressed at low levels. *G. m. morsitans* OR3 and OR45 were exceptionally expressed at high levels in the female and male antennae respectively. Four OR genes were detected in female *G. m. morsitans* leg with OR26 expressed at relatively high level while the male *G. m. morsitans* leg had 9 OR genes detected with OR20 highly transcribed. In larvae six ORs were expressed with OR20 at high levels and in pupae OR27 and OR28 were expressed at relatively high levels (Figure 4.3A to 4.3F).

Odorant receptor profiles of *G. m. morsitans* sample

A.



B.



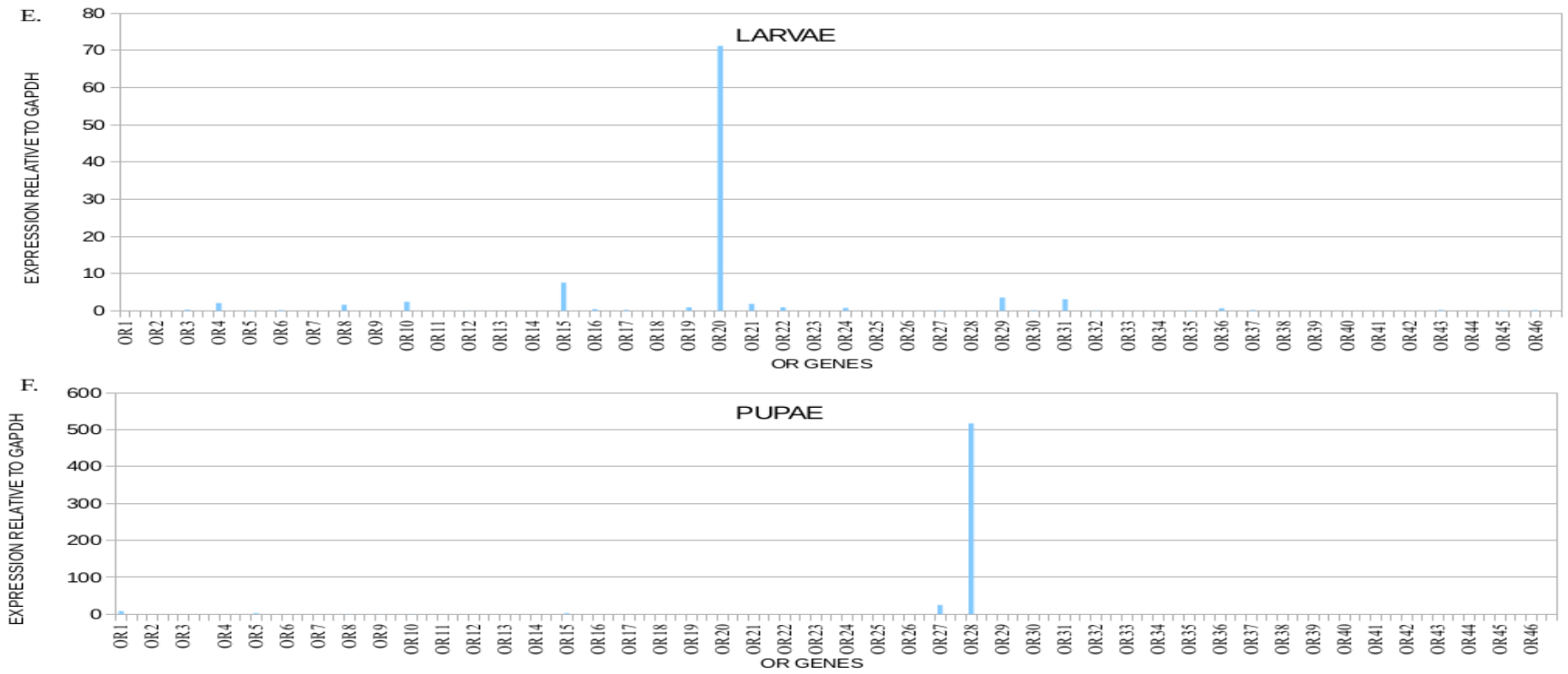


Figure 4.3: Real time polymerase chain reaction (RT-PCR) analysis of *G. m. morsitans* odorant receptor (GmmOR) genes expression levels in female antennae (A), leg tissue (B), male leg (C), antennae tissue (D), larvae (E) and pupae (F). The Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used to normalize the template cDNA. The expression level of Gmm OR genes in female and male *G. m. morsitans* antennae and leg, larvae and pupae was calculated by delta delta Ct (ddCt) method in which the discrepancy

between the Ct for the GmmOR gene and GAPDH was first calculated to normalize the variation in the amount of cDNA in each reaction (Livak and Schmittgen, 2001).

4.3 Bioinformatics Analyses

Blast search against *D. melanogaster,* mosquitoes and sequenced *Glossina* genomes identified orthologs with peptide percentage identities ranging from 22.3% to 98.9% (Table 4.3 and Appendix 1). GmmOR1 had the highest identity to Dmel Orco, confirming the relatively high amino acid identity between Orco from different insects. Similarly, GmmOR1 had relatively high amino acid identity to Orco orthologs in *Aedes aegypti*, *Anopheles gambiae*, *Culex quinquefasciatus*, *G. austeni, G. brevipalpis, G. fuscipes fuscipes* and *G. pallidipes* (Appendix 1). Generally, the OR sequence identity was low among *D. melanogaster* and mosquitoes (*Ae. aegypti*, *An. gambiae* and *Cx. quinquefasciatus*) orthologs and high among *G. austeni, G. brevipalpis, G. f. fuscipes* and *G. pallidipes* orthologs. The low protein sequence similarities were notable, for instance, in GmmOR41, and GmmOR23 whereas high protein sequence identities were recorded in GmmOR1 (97.5%), GmmOR17 (93.3%), and GmmOR40 (96.5%).

4.3.1 Conserved Domains

The NCBI's conserved domain (CDD) analysis of the OR genes yielded domains with specific and general functional characteristics. Except for GmmOR23 being an Alpha-mann_mid — a hydrolase and zinc ion binding entity — that is uncommon with odorant receptors, all the other ORs belonged to the same family of 7tm_6 (Table 4.2).

4.3.2 Protein Structure and Ligands

The RaptorX predicted structures for 45 ORs (Appendix 2). Several ligands were identified, too. Most notable were the ligands for the significantly expressed genes that included Protoporphyrin IX containing Fe, Fe (III) ion and Fe (II) ion,

Magnesium ion, Sulphate ion, Adenosine-5’-monophosphate, Zinc ion, Guanosine-5'-monophosphate, Cytidine-5'-monophosphate, Uridine-5'-monophosphate, Potassium ion, Cadmium ion, 4-(2-Hydroxyethyl)-1-Piperazine ethanesulfonic acid. Nonetheless, GmmOR20 and GmmOR28 had no predicted binding pockets hence lacked ligands to be associated with.

Table 4.2: Odorant receptors (ORs) identified from *Glossina morsitans morsitans* genome and best blast match to *Drosophila melanogaster*, Conserved domain and Gene Ontology Databases

Gmm OR	Gene ID Accession No	Best *D. melanogaster* match (Identities, E-value)	*Gmm* TMDs	Best Rpsblast to CDD DB (E-value)	Match to GO DB (E-value)	GO No.
OR1	GMOY005610	Dmel\Orco (77.5%, 0.0)	5	7tm_6 domain-containing protein (1.19e-51)	odorant binding (0.0)	GO:0005549
OR2	GMOY005796	Dmel\Or2a (47.7%, 1.2e-105)	7	7tm_6 domain-containing protein (6.87e-88)	signal transducer activity (0.0)	GO:0004871
OR3	GMOY004772	Dmel\Or2a (40.4%, 2.0e-49)	2	7tm_6 domain-containing protein (6.87e-88)	odorant binding (0.0)	GO:0005549
OR4	GMOY012195	Dmel\Or59a (38.6%, 3.9e-82)	7	7tm_6 domain-containing protein (4.46e-78)	odorant binding (0.0)	GO:0005549
OR5	GMOY012018	Dmel\Or33b (38.4%, 1.2e-49)	5	7tm_6 domain-containing protein (1.35e-88)	odorant binding (0.0)	GO:0005549
OR6	GMOY009475	Dmel\Or7a (37.9%, 1.1e-65)	5	7tm_6 domain-containing protein (3.64e-86)	odorant binding (0.0)	GO:0005549
OR7	Not Available	Dmel\Or42b (32.3%, 1.13e-56)	6	7tm_6 domain-containing protein (7.90e-55)	odorant binding (2.1e-94)	GO:0005549
OR8	GMOY012193	Dmel\Or7a (34.5%, 8.1e-66)	4	7tm_6 domain-containing protein (3.64e-86)	odorant binding (0.0)	GO:0005549
OR9	GMOY012276	Dmel\Or42a (39.7%, 5.4e-91)	6	7tm_6 domain-containing protein (3.46e-78)	odorant binding (0.0)	GO:0005549
OR10	GMOY012214	Dmel\Or46a (33.8%, 4.3e-56)	6	7tm_6 domain-containing protein (8.95e-71)	odorant binding (0.0)	GO:0005549
OR11	GMOY010761	Dmel\Or46a (31%, 1.3e-44)	6	7tm_6 domain-containing protein (1.08e-74)	odorant binding (0.0)	GO:0005549
OR12	GMOY009271	Dmel\Or94a (39%, 8.2e-69)	3	7tm Odorant receptor (2.82e-76)	odorant binding (0.0)	GO:0005549

OR13	GMOY003312	Dmel\Or82a (45.2%, 2.3e-76)	6	7tm_6 super family (2.53e-42)	odorant binding	(0.0)	GO:0005549
OR14	GMOY001365	Dmel\Or45a (27.9%, 1.7e-29)	6	7tm_6 domain-containing protein (1.53e-62)	odorant binding	(0.0)	GO:0005549
OR15	GMOY012287	Dmel\Or45a (25.3%, 9.4e-33)	7	7tm_6 domain-containing protein (1.53e-62)	odorant binding	(0.0)	GO:0005549
OR16	GMOY012240	Dmel\Or45a (25.6%, 3.7e-32)	6	7tm_6 domain-containing protein (1.53e-62)	odorant binding	(0.0)	GO:0005549
OR17	GMOY005386	Dmel\Or85d (33.6%, 2.4e-17)	8	7tm_6 domain-containing protein (5.73e-83)	odorant binding	(0.0)	GO:0005549
OR18	GMOY012323	Dmel\Or63a (27.9%, 2.0e-51)	7	7tm_6 domain-containing protein (1.35e-64)	odorant binding	(0.0)	GO:0005549
OR19	GMOY012322	Dmel\Or63a (28.1%, 5.6e-49)	7	7tm_6 domain-containing protein (1.35e-64)	odorant binding	(0.0)	GO:0005549
OR20	GMOY012277	Dmel\Or85d (23.6%, 2.7e-08)	0	7tm_6 domain-containing protein (5.73e-83)	odorant binding	(0.0)	GO:0005549
OR21	GMOY011399	Dmel\Or85e (47.5%, 3.4e-123)	6	7tm_6 super family (1.01e-11)	odorant binding	(0.0)	GO:0005549
OR22	GMOY012256	Dmel\Or88a (36.6%, 5.8e-24)	3	7tm_6 domain-containing protein (5.22e-75)	odorant binding	(0.0)	GO:0005549
OR23	GMOY011902	Dmel\LManI (46%, 0.0)	5	Alpha-mann_mid (2.80e-30)	alpha-mannosidase activity (0.0)		GO:0004559
OR24	GMOY010839	Dmel\Or85b (48.2%, 2.5e-100)	5	7tm_6 domain-containing protein (2.52e-80)	odorant binding	(0.0)	GO:0005549
OR25	GMOY012357	Dmel\Or92a (28.4%, 1.3e-27)	5	7tm_6 domain-containing protein (3.08e-50)	odorant binding	(0.0)	GO:0005549
OR26	GMOY012255	Dmel\Or85c (27.7%, 2.5e-25)	6	7tm_6 domain-containing protein (1.60e-80)	odorant binding	(0.0)	GO:0005549
OR27	GMOY008038	Dmel\Or67c (28.3%, 6.3e-25)	3	7tm_6 super family (6.80e-	odorant binding	(0.0)	GO:0005549

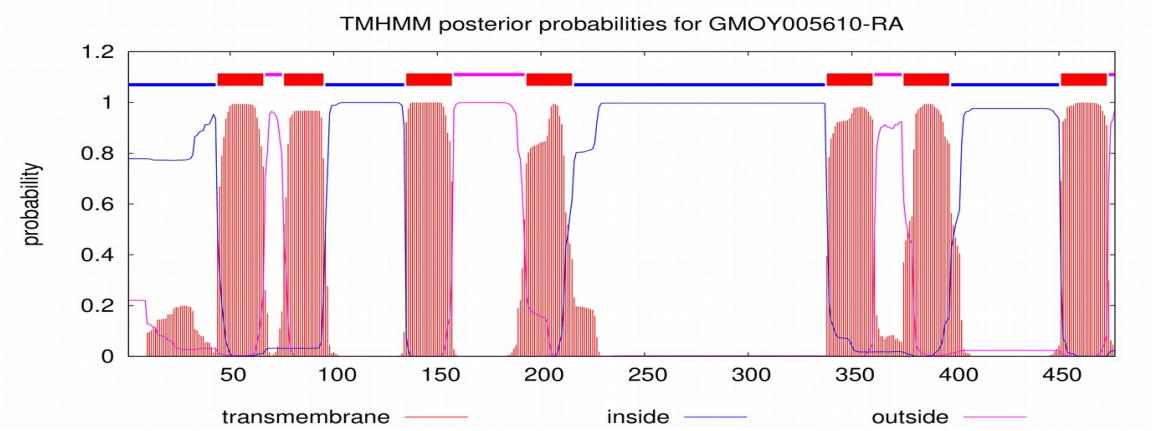
				50)			
OR28	GMOY012199	Dmel\Or85f (26.2%, 0.0)	1	7tm_6 domain-containing protein (3.26e-71)	odorant binding	(0.0)	GO:0005549
OR29	GMOY012283	Dmel\Or67a (20.4%, 4.1e-15)	5	7tm_6 super family (3.48e-39)	odorant binding	(0.0)	GO:0005549
OR30	GMOY012282	Dmel\Or92a (22%, 2.2e-07)	6	7tm_6 domain-containing protein (3.08e-50)	odorant binding	(0.0)	GO:0005549
OR31	GMOY012239	Dmel\Or24a (44.1%, 45.e-104)	7	7tm_6 domain-containing protein (2.19e-75)	odorant binding	(0.0)	GO:0005549
OR32	GMOY005084	Dmel\Or13a (57.9%, 1.4e-143)	4	7tm_6 super family (3.06e-39)	odorant binding	(0.0)	GO:0005549
OR33	GMOY005479	Dmel\Or49b (54.9%, 6.9e-45)	4	7tm_6 domain-containing protein (3.68e-71)	odorant binding	(0.0)	GO:0005549
OR34	GMOY011902	Dmel\Or49a (23%, 2.2e-09)	2	7tm_6 super family (7.23e-41)	odorant binding	(0.0)	GO:0005549
OR35	GMOY012253	Dmel\Or43a (47.5%, 3.9e-96)	7	7tm_6 domain-containing protein (8.62e-77)	odorant binding	(0.0)	GO:0005549
OR36	GMOY012254	Dmel\Or43a (30.2%, 2.4e-32)	5	7tm_6 domain-containing protein (8.62e-77)	odorant binding	(0.0)	GO:0005549
OR37	GMOY012218	Dmel\Or74a (39%, 7.1e-93)	4	7tm_6 super family (2.14e-10)	odorant binding	(0.0)	GO:0005549
OR38	GMOY012208	Dmel\Or47b (32.9%, 6.5e-47)	6	7tm_6 domain-containing protein (6.99e-66)	odorant binding	(0.0)	GO:0005549
OR39	GMOY004392	Dmel\Or88a (27.7%, 3.9e-44)	6	7tm_6 super family (2.23e-37)	odorant binding	(2.5e-131)	GO:0005549
OR40	GMOY012356	Dmel\Or56a (47.2%, 2.4e-76)	4	7tm_6 domain-containing protein (9.26e-50)	odorant binding	(0.0)	GO:0005549
OR41	GMOY006480	Dmel\Or67d (35.3%, 8.9e-62)	6	7tm_6 super family (1.12e-10)	odorant binding	(0.0)	GO:0005549

OR42	GMOY006479	Dmel\Or67d (36.6%, 3.6e-68)	6	7tm_6 super family	(1.12e-10)	odorant binding	(0.0) GO:0005549
OR43	Not Available	Dmel\Or67d (36.6%, 3.6e-68)	6	7tm_6 super family	(1.12e-10)	odorant binding	(0.0) GO:0005549
OR44	GMOY006265	Dmel\Or67d (36.2%, 2.2e-74)	6	7tm_6 super family	(1.12e-10)	odorant binding	(0.0) GO:0005549
OR45	GMOY007896	Dmel\Or67d (38.5%, 9.3e-42)	3	7tm_6 super family	(1.12e-10)	odorant binding	(0.0) GO:0005549
OR46	GMOY003305	Dmel\Or67d (37%, 5.2e-67)	5	7tm_6 super family	(1.12e-10)	odorant binding	(0.0) GO:0005549

4.3.2 Transmembrane Domains

Transmembrane domains were constituted predominantly by the non-polar amino acids. Conserved domain prediction revealed that 45 *G. morsitans morsitans* ORs had transmembrane domains and were involved in odorant binding. The GmmOR23 was the only exception as it didn't have transmembrane domain and was predicted to be involved in alpha mannosidase activity. The start and end of the amino acid sequences denoting the seven transmembrane domains that characterize insect ORs (Figure 4.4), and determined using TMHMM algorithm, were transferred on to the multiple sequence alignments and indicated accurately through colouring (Appendix 3). These TMD analyses revealed helices ranging between 17 and 22 amino acids. Except for serine and threonine residues being common in the TMDs, their composition and number varied greatly from one GmmOR cluster to the other with the highest having thirteen, for instance, *G. fuscipes* in GmmOR5 cluster. However, a few mosquitoes orthologs in GmmOR13, GmmOR11, GmmOR3, GmmOR33 and GmmOR27 clusters, lacked TMDs.

Figure 4.4: Trasmembrane domain depiction of GmmOR1 output by TMHMM



version 2 algorithms. The odorant recptor is predicted to have seven TMDs.

4.3.3 Multiple Sequence Alignment

The *G. m. morsitans* ORs and its orthologs differ in their gene sequences as demonstrated by peptide sequence alignment. These alignments illustrate the highly significant peptide sequence homologies between the *G. m. morsitans* and several other insect OR peptide sequences deposited in the repositories. Multiple sequence alignment revealed conserved and identical residues in the GmmOR clusters of amino acid sequences. The analyses revealed that GmmOR1, GmmOR2, GmmOR4, GmmOR8, GmmOR9, GmmOR10, GmmOR15, GmmOR35, GmmOR18, GmmOR17, GmmOR12, GmmOR24, GmmOR6, GmmOR32, GmmOR44, GmmOR25, GmmOR42, GmmOR41, GmmOR14, and GmmOR21 clusters indicated high probabilities of conservation (Appendix 3).

CHAPTER FIVE

DISCUSSION

The 46 *G. m. morsitans* OR genes analysed in this study had been reported earlier (Macharia *et al.,* 2016; Obiero *et al.,* 2014). *Glossina m. morsitans* was reported to contain fewer ORs relative to closely related dipterans that could be as a result of the restricted diet of tsetse flies (Obiero *et al.,* 2014). This finding was corroborated in a study of other tsetse species (Macharia *et al.,* 2016). The low number of *G. m. morsitans* OR genes compared to other insects (Benton *et al.,* 2006; Clyne *et al.,* 1999; Hill *et al.,* 2002; Robertson and Wanner, 2006; Scott *et al.,* 2014) was thought to be as a result of strict blood-feeding lifestyle of tsetse flies (IGGI, 2014). However, the relative expression of these set of ORs are yet to be determined. This study analysed the differential and relative expression of the full set of *G. m. morsitans* ORs in adult fly antennae and legs both sexes and also in larvae and pupae. This provides a great perspective into the possible role of these ORs at different stages of life of the fly.

Expression profiling revealed that the *G. m. morsitans* ORs genes were highly expressed in the female and male antennae relative to the legs tissues. This supported earlier findings that points to the antennae being the main olfactory organ in tsetse flies (Liu *et al.,* 2010) and as already reported in other insects (Leal, 2013; Ma *et al.,* 2014; Pelosi *et al.,* 2006). In *D. melanogaster*, ORs were reported to be expressed in ORN in the antennae and maxillary palps (Robertson *et al.,* 2003), while in *An.*

gambiae and ants, expression of chemosensory receptors revealed odor-receptor coding signatures that were either tissue or sex-specific (Pitts *et al.,* 2011; Zhou *et al.,* 2012). The varying expression patterns of the OR genes suggests that they may be involved in different, important roles in tsetse olfaction at different developmental stages and body regions.

In pupae and larvae, the over-expressed genes, OR28 and OR20 respectively, may suggest their involvement in development, for instance, of larval specific chemical sensing capability crucial for feeding. Expression of OR genes in larvae is instrumental in understanding the molecular mechanisms involved in pheromone detection (Winnebeck, Millar, and Warman, 2010). Notably, the exclusive expression of GmmOR28 in pupae and GmmOR20 in larvae are suggestive of such functions. The latter, GmmOR20, is homologous to OR85b in *D. melanogaster* that, depending on odour type, has been found to mediate avoidance or acceptance behaviour (Nichols and Luetje, 2010; Hallem and Carlson, 2006). Also, in addition to calcium permeability, this gene encodes wide range of odor stimuli that varies a lot. These behavioural changes are in response to chemicals such as butyl acetate, amyl acetate, and 2-heptanone (Poivet *et al.,* 2013). The expression of GmmOR15 in larvae, homologous to OR45 in *D. melanogaster,* play crucial role in escape response in the larvae (Bellmann *et al.,* 2010; Obiero *et al.,* 2014).

The transcript levels of OR genes reported in both female and male *G. m. morsitans* antennae were approximately equal and could indicate the conserved function of OR genes in *Glossina* as both adult females and males are strict blood-feeders and vectors of protozoan trypanosomes. The GmmOR3 gene that was highly expressed in female *G. m. morsitans* antennae could probably be involved in female tsetse specific

behaviours such as larviposition. While, the GmmOR45 gene that was highly expressed in male *G. m. morsitans* antennae could be involved in mediating male behaviours like mate detection and courtship. Its orthologs are known to perform specific roles as reported in female *Ae. Aegypti* (Sengul and Tu, 2010) *An. gambiae* (Biessmann *et al.,* 2005) and parasitic wasp, *Microplitis mediator* (Ma *et al.,* 2014).

Intriguingly, the distribution of OR genes in *G. m. morsitans* legs with GmmOR26 highly expressed in female leg and GmmOR20 in male legs may pinpoint that the legs could also be involved in functions related to olfaction. The GmmOR15 was reported to be predominately expressed based on RNA-seq data in RPKM (Obiero *et al.,* 2014) and this was observed in the pupae. Its homolog in *D. melanogaster* (DmelOr45a) is associated with larvae response (Bellmann *et al.,* 2010). And in male legs, GmmOR37 and a homolog to OR74a in *D. melanogaster*, has been found to be involved in the behavioural responses to octanol, anisole, and 2-heptanone (Deng *et al.,* 2011), suggestively points to the role legs play in insect olfaction.

It is established that odorant receptors in insects are highly diverse and less conserved even between paralogs, and most case no direct one-to-one homology exist except for between the highly conserved co-receptor ORCO (Miller and Tu, 2008). For instance in this study, GmmOR41, GmmOR20, and GmmOR23, had exceptionally low sequence similarity relative to other paralogs, pointing to probable rapid evolution of the OR genes in the genome, under yet unclear mechanisms (Obiero *et al.,* 2014). However, the sequence divergence observed in different insects ORs could be linked to the insects species-specific expansion of distinct OR genes. *D. melanogaster* genome had 18 ORs not found in mosquitoes while mosquito

genome contained 27 OR genes that were absent in fruit fly (Pitts, Fox, and Zwiebel, 2004). It is important to note also that some receptor sequence did exhibit high sequence similarity in ORs clusters — for instance GmmOR1, GmmOR17, and GmmOR40. Probably, given their primary structural and functional domain similarity, these could be amongst the ancestral genes, attributable to sequence conservation because of the roles these genes play in the organism (Capra and Singh, 2007).

The CDD searches revealed that the *G. m. morsitans* OR clusters had varying number of TMDs on the peptide sequence. The MSA information on this OR revealed a high degree of amino acid residue conservation across the sequences in the cluster. The high peptide sequence conservation level corresponds to the existence of the multi-super families in the cluster. This is because presence of conserved domains in proteins is a function of sequence conservation (Sato *et al.,* 2008). Surprisingly, the identity between GmmOR23 and its orthologs from other insects are among the lowest. This indicates that these genes probably evolved from different origins or are under divergence. If at all it is a correctly annotated OR gene, then functional study is required to validate its role.

In addition to GmmOR20 lacking TMDs, its ortholog in *Aedes aegypti* was irretrievable during data mining in the repositories. The absence of TMDs suggested that these sequences may be involved in non-olfactory functions hence not an odorant receptor. It is, however, unclear if the compositional differences between TMDs confer varying physical properties (Marchler-Bauer *et al.,* 2009). Protein structure and ligand analysis supported earlier research that multiple ligands could bind to a single OR and in turn activate OSN (Ma and Shepherd, 2000). Such ligands

activating the ORs provide possible avenues through vector control measures could be delivered to the insect chemoreceptor system. However, binding of ligands to ORs do not necessarily lead to activation. Other ligands bind and create a competitively antogonizing environment (Araneda et al., 2004; Araneda, Kini and Firestein, 2000). Whereas ligand analysis revealed that GmmOR20 had no predicted ligands, its utilisation for control efforts would be challenging given such possibility. Our study confirmed that previously annotated *G. m. morsitans* OR genes (Obiero *et al.,* 2014) contain the structurally conserved transmembrane domains and are involved in odorant binding.

Although putative tsetse OR genes have been identified using bioinformatics approaches (Macharia *et al.,* 2016; Obiero *et al.,* 2014), no functional studies have reported the binding and response of these ORs to odorants. Neither is there a report in which specific hair-cells the ORs are expressed to separate which type of ORs are expressed on specialist sensillae and which ones on generalist sensillae, as previously pointed out (Voskamp *et al.,* 1999). Our report presents crucial insight into the tissue-differential expression of ORs, which are a major component of the tsetse chemoreceptor system.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The study revealed that the ORs were highly expressed in the antennae, confirming that the antennae is the main olfactory organ. The expressed ORs in the females could indicate the critical roles, such as seeking larviposition sites, performed by ORs in females than the males. More ORs were expressed in the larvae supporting its need for more genes to effect various functions unlike pupae that is immobile.

6.2 Recommendations

1. It is appropriate to perform functional analysis on the highly expressed ORs, for instance GmmOR3 in female antenna, OR26 in female legs, OR45 in male antennae, OR20 male legs, OR20 in larvae, and OR28 pupae, through expression in *Xenopus laevis* oocytes to confirm their specificity.

2. The expressed ORs, for instance GmmOR20 in larvae, GmmOR28 in pupae, should be considered for further functional studies to establish their response to known odors that attract tsetse.

REFERENCES

Aksoy, S., Attardo, G., Berriman, M., Christoffels, A., Lehane, M., Masiga, D., and Toure, Y. (2014). Human African Trypanosomiasis Research Gets a Boost : Unravelling the Tsetse Genome, *8*(4), 6–8.

Aksoy, S., Buscher, P., Lehane, M., Solano, P., and Van Den Abbeele, J. (2017). Human African trypanosomiasis control: Achievements and challenges. *PLoS neglected tropical diseases*, *11*(4), 1-6.

Alcedo, J., Maier, W., & Ch'ng, Q. (2010). Sensory influence on homeostasis and lifespan: molecules and circuits. In *Protein Metabolism and Homeostasis in Aging*. Boston, MA: Springer.

Alexandersson, E., Jacobson, D., Vivier, M. A., Weckwerth, W., & Andreasson, E. (2014). Field-omics—understanding large-scale molecular data from field crops. *Frontiers in plant science*, *5*, 286.

Allsopp, R. (2001). Options for vector control against trypanosomiasis in Africa. *Trends in parasitology*, *17*(1), 15-19.

Andersson, M. N., Larsson, M. C., Blaženec, M., Jakuš, R., Zhang, Q.-H., and Schlyter, F. (2010). Peripheral modulation of pheromone response by inhibitory host compound in a beetle. *J. Exp. Biol. 213*, 3332–3339.

Andreeva, A. (2011). *Classification of proteins: available structural space for molecular modeling. In Homology Modeling.* Florida: Humana Press.

Anholt, R. R. H., and Williams, T. I. (2010). The soluble proteome of the Drosophila antenna. *Chemical Senses*, *35*(1), 21–30.

Araneda RC, Peterlin Z, Zhang X, Chesler A, Firestein S. (2004). A pharmacological profile of the aldehyde receptor repertoire in rat olfactory epithelium. J. Physiol. *555*: 743-756.

Araneda RC, Kini AD, and Firestein S. (2000). The molecular receptive range of an odorant receptor. Nat. Neurosci. *3*: 1248-1255.

Arensburger, P., Megy, K., Waterhouse, R. M., Abrudan, J., Amedeo, P., Antelo, B., … Camara, F. (2010). Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science*, *330*(6000), 86–88.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nat Genet*, *25*(1), 25–29.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., & Krishna, S. (2003). The trypanosomiases. *The Lancet*, *362*(9394), 1469-1480.

Bell, W. J., and Cardé, R. T. (2013). *Chemical ecology of insects*. London: Springer.

Bellmann, D., Richardt, A., Freyberger, R., Nuwal, N., Schwärzel, M., Fiala, A., and Störtkuhl, K. F. (2010). Optogenetically induced olfactory stimulation in Drosophila larvae reveals the neuronal basis of odor-aversion behavior. *Frontiers in Behavioral Neuroscience*, *4*, 27.

Benton, R., Sachse, S., Michnick, S. W., and Vosshall, L. B. (2006). Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biology*, *4*(2), 240–257.

Bett, M. K., Saini, R. K., and Hassanali, A. (2015). Repellency of tsetse-refractory waterbuck (*Kobus defassa*) body odour to Glossina pallidipes (*Diptera: Glossinidae*): Assessment of relative contribution of different classes and individual constituents. *Acta Tropica*, *146*, 17–24.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., ... & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research*, *42*(W1), W252-W258.

Biessmann, H., Nguyen, Q. K., Le, D., and Walter, M. F. (2005). Microarray-based survey of a subset of putative olfactory genes in the mosquito Anopheles gambiae. *Insect Molecular Biology*, *14*(6), 575–589.

Black, S. J., and Mansfield, J. M. (2016). Prospects for vaccination against pathogenic African trypanosomes. *Parasite immunology*, *38*(12), 735-743.

Black, S. J., Richard Seed, J., and Murphy, N. B. (2001). Innate and acquired resistance to African trypanosomiasis. *Journal of Parasitology*, *87*(1), 1–9.

Bohbot, J., Pitts, R. J., Kwon, H. W., Rützler, M., Robertson, H. M., & Zwiebel, L. J. (2007). Molecular characterization of the Aedes aegypti odorant receptor gene family. *Insect molecular biology*, *16*(5), 525-537.

Boron, W. F., and Boulpaep, E. L. (2009). *Medical physiology: a cellular and molecular approach. GEN, Elsevier*. Philadelphia, PA: Saunders.

Boulpaep, E. L., and Boron, W. F. (2005). *Medical physiology: a cellular and molecular approach.* St. Louis, Mo: Elsevier Saunders.

Bouyer, J., Seck, M. T., Sall, B., Ndiaye, E. Y., Guerrini, L., and Vreysen, M. J. B. (2010). Stratified entomological sampling in preparation for an area-wide integrated pest management program: the example of Glossina palpalis gambiensis (Diptera: Glossinidae) in the Niayes of Senegal. *Journal of Medical Entomology*, *47*(4), 543–552.

Braun, W. John and Murdoch, Duncan J. (2007). *A First Course in Statistical Programming with R*. Cambridge: Cambridge University Press.

Breed, M. D., and Moore, J. (2010). *Encyclopaedia of animal behaviour.* London: *Acad. Press*.

Bukachi, S. A., Wandibba, S., & Nyamongo, I. K. (2017). The socio-economic burden of human African trypanosomiasis and the coping strategies of households in the South Western Kenya foci. *PLoS Neglected Tropical Diseases*, *11*(10), e0006002.

Bystroff, C., & Krogh, A. (2008). *Hidden Markov Models for prediction of protein features. In Protein Structure Prediction*. New York: Humana Press.

Callahan, F. E., Vogt, R. G., Tucker, M. L., Dickens, J. C., and Mattoo, A. K. (2000). High level expression of “male specific” pheromone binding proteins

(PBPs) in the antennae of female noctuiid moths. *Insect Biochemistry and Molecular Biology*, *30*(6), 507–514.

Capra, J. A., and Singh, M. (2007). Predicting functionally important residues from sequence conservation. *Bioinformatics*, *23*(15), 1875–1882.

Cardé, R. T., and Willis, M. A. (2008). Navigational strategies used by insects to find distant, wind-borne sources of odor. *Journal of Chemical Ecology*, *34*(7), 854–866.

Carey AF and Carlson JR. (2011). Insect olfaction from model systems to disease control. Proc Natl Acad Sci USA. 108: 12987–12995.

Carey AF, Wang G, Su CY, Zwiebel LJ, Carlson JR (2010). Odorant reception in the malaria mosquito Anopheles gambiae. Nature. *464*: 66–71.

Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, *17*(4), 540–552.

Cavalloro, R. (1987). *Integrated Tse-tse Fly Control* (Vol. 10530). Florida: *CRC Press*.

Cecchi, G., Mattioli, R. C., Slingenbergh, J., and De La Rocque, S. (2008). Land cover and tsetse fly distributions in sub-Saharan Africa. *Medical and Veterinary Entomology*, *22*(4), 364–373.

Cecchi, G., Paone, M., Herrero, R. A., Vreysen, M. J. B., and Mattioli, R. C. (2015). Developing a continental atlas of the distribution and trypanosomal infection of tsetse flies (Glossina species). *Parasites and Vectors*, *8*(1), 284.

Chakrabarti, S., Lanczycki, C. J., Panchenko, A. R., Przytycka, T. M., Thiessen, P. A., & Bryant, S. H. (2006). Refining multiple sequence alignments with conserved core regions. *Nucleic acids research*, *34*(9), 2598-2606.

Clark, N., Frost, A., Maudlin, I., and Ward, A. (2013). *Technology development assistance for agriculture: putting research into use in low income countries.* London: Routledge.

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. (1999). A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in Drosophila. *Neuron*, *22*(2), 327–338.

Consortium, G. O. (2004). The Gene Ontology (GO) database and informatics resource, *32*, 258–261.

Dani, F. R., Iovinella, I., Felicioli, A., Niccolini, A., Calvello, M. A., Carucci, M. G., … Moneti, G. (2010). Mapping the expression of soluble olfactory proteins in the honeybee. *Journal of Proteome Research*, *9*(4), 1822–1833.

De Bruyne, M., and Baker, T. C. (2008). Odor detection in insects: volatile codes. *Journal of Chemical Ecology*, 34(7), 882–897.

DeGennaro M. (2015). The mysterious multi-modal repellency of DEET. Fly (Austin). *9*: 45–51.

Deng, Y., Zhang, W., Farhat, K., Oberland, S., Gisselmann, G., and Neuhaus, E. M. (2011). The stimulatory Gαs protein is involved in olfactory signal transduction in Drosophila. *PLoS One*, *6*(4), e18605.

Deorowicz, S., Debudaj-Grabysz, A., & Gudyś, A. (2014). Kalign-LCS—A more accurate and faster variant of kalign2 algorithm for the multiple sequence alignment problem. In *Man-Machine Interactions 3* (pp. 495-502). Springer, Cham.

Derveaux, S., Vandesompele, J., and Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. *Methods*, *50*(4), 227–230.

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., ... & Notredame, C. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic acids research*, *39*(suppl_2), W13-W17.

Djouaka, R. J., Atoyebi, S. M., Tchigossou, G. M., Riveron, J. M., Irving, H., Akoton, R., ... & Wondji, C. S. (2016). Evidence of a multiple insecticide resistance in the malaria vector Anopheles funestus in South West Nigeria. *Malaria journal*, *15*(1), 565.

Dobritsa, A. A., van Naters, W. van der G., Warr, C. G., Steinbrecht, R. A., and Carlson, J. R. (2003). Integrating the molecular and cellular basis of odor coding in the *Drosophila* antenna. *Neuron*, *37*(5), 827–841.

Dorak, M. T. (2007). *Real-time PCR*. New York: Taylor and Francis, 270 Madison Avenue.

Doty, R. L. (2015). *Handbook of olfaction and gustation,* 3rd Ed. New York: John Wiley and Sons.

Du, L., Zhao, X., Liang, X., Gao, X., Liu, Y., & Wang, G. (2018). Identification of candidate chemosensory genes in Mythimna separata by transcriptomic analysis. *BMC genomics*, *19*(1), 518.

Dunipace, L., Meister, S., McNealy, C., and Amrein, H. (2001). Spatially restricted expression of candidate taste receptors in the Drosophila gustatory system. *Current Biology*, *11*(11), 822–835.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, *32*(5), 1792-1797

Elmore, T., Ignell, R., Carlson, J. R., and Smith, D. P. (2003). Targeted mutation of a *Drosophila* odor receptor defines receptor requirement in a novel class of sensillum. *The Journal of Neuroscience*, *23*(30), 9906–9912.

Esterhuizen, J., Rayaisse, J. B., Tirados, I., Mpiana, S., Solano, P., Vale, G. A., … Torr, S. J. (2011). Improving the cost-effectiveness of visual devices for the control of riverine tsetse flies, the major vectors of human African trypanosomiasis. *PLoS Negl Trop Dis*, *5*(8), e1257.

Fan, T. W. M., & Lane, A. N. (2016). Applications of NMR spectroscopy to systems biochemistry. *Progress in nuclear magnetic resonance spectroscopy*, *92*, 18-53.

Faraggi, E., Dunker, A. K., Sussman, J. L., & Kloczkowski, A. (2018). Comparing NMR and X-ray protein structure: Lindemann-like parameters and NMR disorder. *Journal of Biomolecular Structure and Dynamics*, *36*(9), 2331-2341.

Feng, W., Pan, L., & Zhang, M. (2011). Combination of NMR spectroscopy and X-ray crystallography offers unique advantages for elucidation of the structural basis of protein complex assembly. *Science China Life Sciences*, *54*(2), 101-111.

Franco, J. R., Simarro, P. P., Diarra, A., and Jannin, J. G. (2014). Epidemiology of human African trypanosomiasis. *Clin Epidemiol*, *6*, 257–275.

Galizia, C. G., and Lledo, P. M. (2013). *Olfaction In Neurosciences - From molecule to behavior: a university textbook*. Berlin Heidelberg: Springer.

Gikonyo, N. K., Hassanali, A., Njagi, P. G. N., and Saini, R. K. (2003). Responses of *Glossina morsitans morsitans* to blends of electroantennographically active compounds in the odors of its preferred (buffalo and ox) and non-preferred (waterbuck) hosts. *Journal of Chemical Ecology*, *29*(10), 2331–2345.

Giraldo-Calderón GI, Emrich SJ, MacCallum RM, Maslen G, Dialynas E, Topalis P, Ho N, Gesing S, VectorBase Consortium, Madey G, Collins FH, Lawson D. (2015). VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Research *43*:D707-13.

Gooding, R. H., and Krafsur, E. S. (2005). Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. *Annual Review of Entomology*, *50*(8), 101–23.

Gouteux, J.-P., and Jarry, M. (1998). Tsetse flies, biodiversity and the control of sleeping sickness. Structure of a *Glossina* guild in southwest Cote d'Ivoire. *Acta Oecologica*, *19*(5), 453–471.

Gramates LS, Marygold SJ, dos Santos G, Urbano J-M, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, Falls K, Goodman JL, Hu Y, Ponting L, Schroeder AJ, Strelets VB, Thurmond J, Zhou P and the FlyBase Consortium. (2017). FlyBase at 25: looking to the future. Nucleic Acids Res. *45*(D1):D663-D671

Gudermann, T., Nürnberg, B., and Schultz, G. (1995). Receptors and G proteins as primary components of transmembrane signal transduction. *Journal of molecular medicine*, *73*(2), 51-63.

Guidobaldi, F., May-Concha, I. J., and Guerenstein, P. G. (2014). Morphology and physiology of the olfactory system of blood-feeding insects. *Journal of Physiology-Paris*, *108*(2), 96–111.

Hallem, E. A., & Carlson, J. R. (2006). Coding of odors by a receptor repertoire. *Cell*, *125*(1), 143-160.

Hansen, I. A., Rodriguez, S. D., Drake, L. L., Price, D. P., Blakely, B. N., Hammond, J. I., … Romero, A. (2014). The Odorant Receptor Co-Receptor from the Bed Bug, *Cimex lectularius L .*, *9*(11).

Hansson, B. S., Giordano, E., Polito, C. L., Digilio, F. A., and Scialo, F. (2012). Molecular and Functional Characterization of the Odorant Receptor2 (OR2) in the Tiger Mosquito *Aedes albopictus*, *7*(5), 1–11.

Harraca, V., Syed, Z., and Guerin, P. M. (2009). Olfactory and behavioural responses of tsetse flies, Glossina spp., to rumen metabolites. *Journal of Comparative Physiology A*, *195(*9), 815–824.

Hilgenkamp, K. (2006). *Environmental health: Ecological perspectives,* 1st Ed. Boston, Massachusetts: *J*ones and Bartlett Learning.

Hill, C. A., Fox, A. N., Pitts, R. J., Kent, L. B., Tan, P. L., Chrystal, M. A., … Zwiebel, L. J. (2002). G protein-coupled receptors in Anopheles gambiae. *Science*, *298*(5591), 176–178.

Hoare, C. A. (1970). Systematic description of the mammalian trypanosomes of Africa. *Systematic description of the mammalian trypanosomes of Africa.* London: Alien and Unwin.

Horan, R. D., Fenichel, E. P., Wolf, C. A., and Gramig, B. M. (2010). Managing infectious animal disease systems. *Annu. Rev. Resour. Econ.*, *2*(1), 101–124.

I. G. G. I (2014). Genome sequence of the tsetse fly (Glossina morsitans): vector of African trypanosomiasis. *Science*, *344*(6182), 380–386.

Ilemobade, A. A. (2009). Tsetse and trypanosomosis in Africa: the challenges, the opportunities. *Onderstepoort Journal of Veterinary Research*, *76*(1), 35–40.

Jacquin-Joly, E., and Merlin, C. (2004). Insect olfactory receptors: contributions of molecular biology to chemical ecology. *Journal of Chemical Ecology*, *30*(12), 2359–2397.

Jones, P. L., Pask, G. M., Rinker, D. C., and Zwiebel, L. J. (2011). Functional agonism of insect odorant receptor ion channels. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8821–8825.

Jones, P. L., Pask, G. M., Romaine, I. M., Taylor, R. W., Reid, P. R., Waterson, A. G., et al. (2012). Allosteric antagonism of insect odorant receptor ion channels. *PLoS ONE 7*(2), 1-7.

Jones, W. D., Cayirlioglu, P., Kadow, I. G., and Vosshall, L. B. (2007). Two odorant receptors together mediate carbon dioxide detection in *Drosophila*. *Nature, 445*(7123), 86-90.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772-780.

Kaupp, U. B. (2010). Olfactory signalling in vertebrates and insects: differences and commonalities. *Nature Reviews Neuroscience*, *11*(3), 188.

Kepchia, D., Moliver, S., Chohan, K., Phillips, C., & Luetje, C. W. (2017). Inhibition of insect olfactory behavior by an airborne antagonist of the insect odorant receptor co-receptor subunit. *PloS one*, *12*(5), e0177454.

Kim, H. Y. (2014). Analysis of variance (ANOVA) comparing means of more than two groups. *Restorative dentistry & endodontics*, *39*(1), 74-77.

Knight, A. L., & Light, D. M. (2012). Monitoring codling moth (Lepidoptera: Tortricidae) in sex pheromone-treated orchards with (E)-4, 8-dimethyl-1, 3, 7-

nonatriene or pear ester in combination with codlemone and acetic acid. *Environmental entomology*, *41*(2), 407-414.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. *Journal of Molecular Biology*, *305*(3), 567–580.

Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H., and Vosshall, L. B. (2004). Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron*, *43*(5), 703–714.

Leak, S. G. A. (1999). *Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis*. Wallingford, UK: CABI publishing.

Leal, W. S. (2013). Odorant Reception in Insects : Roles of Receptors, Binding Proteins, and Degrading Enzymes. *58*(1), 373-391.

Leal, W. S. (2014). The enigmatic reception of DEET—the gold standard of insect repellents. Curr Opin Insect Sci. *6*: 93–98.

Lee, Y., Moon, S. J., and Montell, C. (2009). Multiple gustatory receptors required for the caffeine response in *Drosophila*. *Proc Natl Acad Sci, 106*(11), 4495-4500.

Liu, R., He, X., Lehane, S., Lehane, M., Berriman, M., Field, L. M., and Zhou, J. (2012). Expression of chemosensory proteins in the tsetse fly *Glossina morsitans morsitans* is related to female host-seeking behaviour, 21(November 2011), 41–48.

Liu, R., Lehane, S., He, X., Lehane, M., and Berriman, C. H. M. (2010). Characterisations of odorant-binding proteins in the tsetse fly Glossina morsitans morsitans. *67*(6). pp 919–929.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)*, 25(4), 402–8.

Ma, L., Gu, S.-H., Liu, Z.-W., Wang, S.-N., Guo, Y.-Y., Zhou, J.-J., and Zhang, Y.-J. (2014). Molecular characterization and expression profiles of olfactory receptor genes in the parasitic wasp, Microplitis mediator (Hymenoptera: Braconidae). *Journal of Insect Physiology*, *60*, 118–126.

Ma M and Shepherd GM (2000). Functional mosaic organization of mouse olfactory receptor neurons. Proc Natl Acad Sci USA. 97: 12869-12874

Macharia, R., Mireji, P., Murungi, E., Murilla, G., Christoffels, A., Aksoy, S., and Masiga, D. (2016). Genome-Wide comparative analysis of chemosensory gene families in five tsetse fly species. *PLoS Negl Trop Dis*, *10*(2), e0004421.

Magez, S., and Radwanska, M. (2014). *Trypanosomes and trypanosomiasis*. Verlag: Springer.

Makovets, S. (2013). Basic DNA electrophoresis in molecular cloning: a comprehensive guide for beginners. *DNA Electrophoresis: Methods and Protocols*, *3*(1) 11–43.

Marchler-Bauer A et al. (2017). *"CDD/SPARCLE: functional classification of proteins via subfamily domain architectures."*, Nucleic Acids Res.*45* (D)200-3.

Marchler-Bauer, A., & Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic acids research*, *32*(suppl_2), W327-W331.

Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., Deweese-Scott, C., Fong, J. H., … Bryant, S. H. (2009). CDD: Specific functional annotation with the Conserved Domain Database. *Nucleic Acids Research*, *37*(SUPPL. 1), 205–210.

Marti-Renom, M. A., Madhusudhan, M. S., & Sali, A. (2004). Alignment of protein sequences by their profiles. *Protein Science*, *13*(4), 1071-1087.

Masiga, D., Obiero, G., Macharia, R., Mireji, P., and Christoffels, A. (2014). Chemosensory receptors in tsetse flies provide link between chemical and behavioural ecology. *Trends in parasitology*, *30*(9), 426-428.

Mbahin, N., Affognon, H., Andoke, J., Tiberius, M., Mbuvi, D., Otieno, J., … Saini, R. K. (2013). Parasitological prevalence of bovine trypanosomosis in Kubo division of Kwale county of coastal: baseline survey. *American Journal of Animal and Veterinary Sciences*, 8(1), 28.

Miao, J., Ishikawa, T., Robinson, I. K., & Murnane, M. M. (2015). Beyond crystallography: Diffractive imaging using coherent x-ray light sources. *Science*, *348*(6234), 530-535.

Miller, R., & Tu, Z. (2008). Odorant Receptor C-Terminal Motifs in Divergent Insect Species. *Journal of Insect Science*, *8*, 53.

Moloo, S. K. (1971). An artificial feeding technique for Glossina. *Parasitology*, *63*(3), 507–512.

Mombaerts, P. (1999). Seven-transmembrane proteins as odorant and chemosensory receptors. *Science*, *286*(5440), 707–711.

Montell, C. (2010). NIH Public Access, *19*(4), 345–353.

Moore, S., Shrestha, S., Tomlinson, K. W., and Vuong, H. (2012). Predicting the effect of climate change on African trypanosomiasis: integrating epidemiology with parasite and vector biology. *Journal of the Royal Society Interface*, *9*(70), 817–830.

Morten Källberg, Haipeng Wang, Sheng Wang, Jian Peng, Zhiyong Wang, Hui Lu & Jinbo Xu (2012). Template-based protein structure modeling using the RaptorX web server. *Nature Protocols 7*, 1511–1522.

Muturi, C. N., Ouma, J. O., Malele, I. I., Ngure, R. M., Rutto, J. J., Klaus, M., … Masiga, D. K. (2011). Tracking the Feeding Patterns of Tsetse Flies (*Glossina Genus*) by Analysis of Bloodmeals Using Mitochondrial Cytochromes Genes, *6*(2), 1–6.

Nakagawa, T., Sakurai, T., Nishioka, T., and Touhara, K. (2005). Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science*, *307*(5715), 1638–1642.

Nichols, A. S., & Luetje, C. W. (2010). Transmembrane segment 3 of Drosophila melanogaster odorant receptor subunit 85b contributes to ligand-receptor interactions. *Journal of Biological Chemistry*, jbc-M109.

Niu, D.-J., Liu, Y., Dong, X.-T., and Dong, S.-L. (2016). Transcriptome based identification and tissue expression profiles of chemosensory genes in Blattella germanica (Blattaria: Blattidae). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *18*, 30–43.

Obiero, G. F. O., Mireji, P. O., Nyanjom, S. R. G., Christoffels, A., Robertson, H. M., and Masiga, D. K. (2014). Odorant and Gustatory Receptors in the Tsetse Fly *Glossina morsitans morsitans*. *PLoS Neglected Tropical Diseases*, *8*(4), 1–8.

Ogwang, M O, Ngaira, J., Wamunyokoli, F. W, Nyanjom, S. (2014). Identification of *Glossina morsitans morsitans* odorant binding proteins genes in *Glossina fuscipes fuscipes*: A Preliminary study JAGST, 1*6*(2), 72–85.

Omolo, M. O., Hassanali, A., Mpiana, S., Esterhuizen, J., Lindh, J., Lehane, M. J., ... & Tirados, I. (2009). Prospects for developing odour baits to control Glossina fuscipes spp., the major vector of human African trypanosomiasis. *Plos Neglected Tropical Diseases*, *3*(5), e435.

Ozaki, K., Utoguchi, A., Yamada, A., and Yoshikawa, H. (2008). Identification and genomic structure of chemosensory proteins (CSP) and odorant binding proteins (OBP) genes expressed in foreleg tarsi of the swallowtail butterfly Papilio xuthus. *Insect Biochemistry and Molecular Biology*, *38*(11), 969–976.

Pearson, W. R. (2013). An introduction to sequence similarity ("homology") searching. *Current protocols in bioinformatics*, *42*(1), 3-1.

Pelletier, J., Hughes, D. T., Luetje, C. W., and Leal, W. S. (2010). An odorant receptor from the southern house mosquito *Culex pipiens quinquefasciatus* sensitive to oviposition attractants. *PloS One*, *5*(4), e10090.

Pelosi, P., Zhou, J.-J., Ban, L. P., and Calvello, M. (2006). Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences*, *63*(14), 1658–1676.

Perry, B. D. (2015). Towards a healthier planet: Veterinary epidemiology research at the International Laboratory for Research on Animal Diseases (ILRAD) and the International Livestock Research Institute (ILRI), 1987–2014

Pierre Gerber, Harold Mooney, Jeroen Dijkman, Shirley Tarewali, and C. de H. (2010). Livestock in a Changing Landscape: Experiences and Regional Perspectives. *Island Press*, *2*, 210.

Pitts, R. J., Fox, A. N., and Zwiebel, L. J. (2004). A highly conserved candidate chemoreceptor expressed in both olfactory and gustatory tissues in the malaria vector Anopheles gambiae. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(14), 5058–5063.

Pitts, R. J., Rinker, D. C., Jones, P. L., Rokas, A., and Zwiebel, L. J. (2011). Transcriptome profiling of chemosensory appendages in the malaria vector Anopheles gambiae reveals tissue-and sex-specific signatures of odor coding. *BMC Genomics*, *12*(1), 271.

Poivet, E., Gallot, A., Montagné, N., Glaser, N., Legeai, F., and Jacquin-Joly, E. (2013). A Comparison of the Olfactory Gene Repertoires of Adults and Larvae in the Noctuid Moth *Spodoptera littoralis*. *PLoS ONE*, *8*(4).

Qiao, H.-L., Deng, P.-Y., Li, D.-D., Chen, M., Jiao, Z.-J., Liu, Z.-C., … Kan, Y.-C. (2013). Expression analysis and binding experiments of chemosensory proteins indicate multiple roles in Bombyx mori. *Journal of Insect Physiology*, *59*(7), 667–675.

Qiao, H., He, X., Schymura, D., Ban, L., Field, L., Dani, F. R., … Iatrou, K. (2011). Cooperative interactions between odorant-binding proteins of Anopheles gambiae. *Cellular and Molecular Life Sciences*, *68*(10), 1799–1813.

Ramdya P, and Benton R. (2010). Evolving olfactory systems on the fly. Trends Genet. *26*(7):307–16.

Reisenman, C. E., Lei, H., & Guerenstein, P. G. (2016). Neuroethology of olfactory-guided behavior and its potential application in the control of harmful insects. *Frontiers in physiology*, *7*, 271.

Renou, M. (2014). Pheromones and general odor perception in insects. In *Neurobiology of chemical communication*. Boca Raton, FL: CRC Press.

Rinker DC, Jones PL, Pitts RJ, Rutzler M, Camp G, Sun LJ, et al., (2012). Novel high-throughput screens of Anopheles gambiae odorant receptors reveal candidate behavior-modifying chemicals for mosquitoes. Physiol Entomol. *37*: 33–41.

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor Protocols*, *2010*(6), pdb-prot5439.

Robertson, H. M., and Wanner, K. W. (2006). The chemoreceptor superfamily in the honey bee, Apis mellifera: expansion of the odorant, but not gustatory, receptor family. *Genome Research*, *16*(11), 1395–1403.

Robertson, H. M., Warr, C. G., and Carlson, J. R. (2003). Molecular evolution of the insect chemoreceptor gene superfamily in Drosophila melanogaster. *Proceedings of the National Academy of Sciences*, *100*(suppl 2), 14537–14542.

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols*, *5*(4), 725.

Rozendaal, J. A. (1997). *Vector control: methods for use by individuals and communities*. New York: World Health Organization.

Ryan, M. (2002). *Insect chemoreception: fundamental and applied*. London: Kluwer Academic Publishers.

Rytz, R., Croset, V., and Benton, R. (2013). Ionotropic Receptors (IRs): Odorant ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochemistry and Molecular Biology, 8*(13), 1-10.

Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., … Nishioka, T. (2004). Identification and functional characterization of a sex pheromone receptor in the silkmoth Bombyx mori. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(47), 16653–16658.

Sambrook, J., and Russel, D. W. (2001). Molecular Cloning; A laboratory manual 3rd ed. Chapter 8. GEN, vitro amplification of DNA, by polymerase chain reaction. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sanchez-Gracia, A., Vieira, F. G., and Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity*, *103*(3), 208–216.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., and Touhara, K. (2008). Insect olfactory receptors are heteromeric ligand-gated ion channels, *452*(4), 1-7.

Sato, K., Tanaka, K., and Touhara, K. (2011). Sugar-regulated cation channel formed by an insect gustatory receptor. *Proc Natl Acad Sci, 108*(28), 11680-11685.

Scialò, F., Hansson, B. S., Giordano, E., Polito, C. L., and Digilio, F. A. (2012). Molecular and functional characterization of the odorant receptor2 (OR2) in the tiger mosquito *Aedes albopictus*. *PLoS ONE*, *7*(5), 1–11.

Scott, J. G., Warren, W. C., Beukeboom, L. W., Bopp, D., Clark, A. G., Giers, S. D., … Leichter, C. A. (2014). Genome of the house fly, Musca domestica L., a global vector of diseases with adaptations to a septic environment. *Genome Biology*, *15*(10), 466.

Scott, K., Brady, R., Cravchik, A., Morozov, P., Rzhetsky, A., Zuker, C., and Axel, R. (2001). A chemosensory gene family encoding candidate gustatory and olfactory receptors in Drosophila. *Cell*, *104*(5), 661–673.

Sengul, M. S., and Tu, Z. (2008). Characterization and expression of the odorant-binding protein 7 gene in *Anopheles stephensi* and comparative analysis among five mosquito species. *Insect Molecular Biology*, *17*(6), 631–645.

Sengul, M. S., and Tu, Z. (2010). Expression analysis and knockdown of two antennal odorant-binding protein genes in Aedes aegypti. *Journal of Insect Science*, *10*(171), 1–18.

Sengul, M. S., and Tu, Z. (2010). Identification and characterization of odorant-binding protein 1 gene from the Asian malaria mosquito, Anopheles stephensi. *Insect Molecular Biology*, *19*(1), 49–60.

Sievers, F., & Higgins, D. G. (2014). Clustal Omega, accurate alignment of very large numbers of sequences. In *Multiple sequence alignment methods.* Totowa, NJ: Humana Press,

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

Simarro, P. P., Cecchi, G., Franco, J. R., Paone, M., Diarra, A., Ruiz-Postigo, J. A., … Jannin, J. G. (2012). Estimating and mapping the population at risk of sleeping sickness. *PLoS Negl Trop Dis*, *6*(10), 1859.

Simarro, P. P., Diarra, A., Ruiz Postigo, J. A., Franco, J. R., & Jannin, J. G. (2011). The Human African Trypanosomiasis Control and Surveillance Programme of the World Health Organization 2000–2009: The Way Forward. *PLoS Neglected Tropical Diseases*, *5*(2), e1007.

Simarro, P. P., Jannin, J., and Cattand, P. (2008). Eliminating human African trypanosomiasis: where do we stand and what comes next? *PLoS Med*, *5*(2), e55.

Simo, G., Njiokou, F., Mbida Mbida, J. A., Njitchouang, G. R., Herder, S., Asonganyi, T., and Cuny, G. (2008). Tsetse fly host preference from sleeping sickness foci in Cameroon: Epidemiological implications. *Infection, Genetics and Evolution*, *8*(1), 34–39.

Slone, J., Daniels, J., and Amrein, H. (2007). Sugar receptors in *Drosophila*. *Current Biol, 17*(20), 1809-1816.

Spickler, A. R. (2010). *Emerging and exotic diseases of animals*. New York: Center for Food Security and Public Health.

Stenberg, J. A., Heil, M., Åhman, I., & Björkman, C. (2015). Optimizing crops for biocontrol of pests and disease. *Trends in plant science*, *20*(11), 698-712.

Steverding, D. (2008). The history of African trypanosomiasis. *Parasites and Vectors*, *1*(1), 3.

Sun, M., Liu, Y., and Wang, G. (2013). Expression patterns and binding properties of three pheromone binding proteins in the diamondback moth, *Plutella xyllotella*. *Journal of Insect Physiology*, *59*(1), 46–55.

Sung, S. S. (2015). Peptide folding driven by Van der W aals interactions. *Protein Science*, *24*(9), 1383-1388.

Takken, W., and Knols, B. G. J. (2010). *Olfaction in vector-host interactions*. Berlin, Wageningen Academic Pub.

Torr, S. J., and Solano, P. (2010). 12. Olfaction in Glossina–host interactions: a tale of two tsetse. *Olfaction in Vector-Host Interactions*, *265*(1), 1-6.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, *40*(15), 115-115.

Verhulst, N. O., Qiu, Y. T., Beijleveld, H., Maliepaard, C., Knights, D., Schulz, S., … Haasnoot, G. W. (2011). Composition of human skin microbiota affects attractiveness to malaria mosquitoes. *PloS One*, *6*(12), e28991.

Vogt, R. G., Rogers, M. E., Franco, M. D., and Sun, M. (2002). A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (*Lepidoptera*) and the organization of OBP genes in *Drosophila melanogaster* (*Diptera*). *Journal of Experimental biology, 205*, 719-744.

Voskamp, K. E., Everaarts, E., and Den Otter, C. J. (1999). Olfactory responses to attractants and repellents in tsetse. *Medical and Veterinary Entomology*, *13*(4), 386–392.

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. (1999). A spatial map of olfactory receptor expression in the Drosophila antenna. *Cell*, *96*(5), 725–736.

Vosshall, L. B., Wong, a M., and Axel, R. (2000). An olfactory sensory map in the fly brain. *Cell*, *102*(2), 147–59.

Vreysen, M. J. B. (2001). Principles of area-wide integrated tsetse fly control using the sterile insect technique. *Médecine Tropicale*, *61*(4–5), 397–410.

Vreysen, M. J. B., Hendrichs, J., and Enkerlin, W. R. (2006). The sterile insect technique as a component of sustainable area-wide integrated pest management of selected horticultural insect pests. *Journal of Fruit and Ornamental Plant Research*, *14,* 107.

Vreysen, M. J. B., Robinson, A. S., Hendrichs, J., and Kenmore, P. (2007). Area-wide integrated pest management (AW-IPM): principles, practice and prospects. In *Area-Wide Control of Insect Pests.* New York. CHAP, Springer.

Vreysen, M. J. B., Saleh, K. M., Ali, M. Y., Abdulla, A. M., Zhu, Z.-R., Juma, K. G., … Feldmann, H. U. (2000). Glossina austeni (Diptera: Glossinidae) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. *Journal of Economic Entomology*, *93*(1), 123–135.

Vreysen, M. J. B., Seck, M. T., Sall, B., and Bouyer, J. (2013). Tsetse flies: Their biology and control using area-wide integrated pest management approaches. *Journal of Invertebrate Pathology*, 112, S15–S25.

Waiswa, C., Picozzi, K., Katunguka-Rwakishaya, E., Olaho-Mukani, W., Musoke, R. A., and Welburn, S. C. (2006). Glossina fuscipes fuscipes in the trypanosomiasis endemic areas of South Eastern Uganda: apparent density, trypanosome infection rates and host feeding preferences. *Acta Tropica*, *99*(1), 23–29.

Wang G, Carey AF, Carlson JR, Zwiebel LJ (2010). Molecular basis of odor coding in the malaria vector mosquito Anopheles gambiae. Proc Natl Acad Sci USA. *107*: 4418–4423.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191.

WHO. (2013). Control and surveillance of human African trypanosomiasis. *World Health Organization Technical Report Series*, *1*(984), 1–237.

Winnebeck, E. C., Millar, C. D., and Warman, G. R. (2010). Why does insect RNA look degraded? *Journal of Insect Science (Online)*, *10*(159), 159.

Wright, G. A., and Smith, B. H. (2004). Different thresholds for detection and discrimination of odors in the honey bee (*Apis mellifera*). *Chemical Senses*, *29*(2), 127–135.

Xu, W., Zhang, H. J., and Anderson, A. (2012). A sugar gustatory receptor identified from the foregut of cotton bollworm *H. armigera*. *J Chem Ecol, 38*(12), 1513-1520.

Younus, F., Chertemps, T., Pearce, S. L., Pandey, G., Bozzolan, F., Coppin, C. W., … Oakeshott, J. G. (2014). Identification of candidate odorant degrading gene/ enzyme systems in the antennal transcriptome of Drosophila melanogaster. *Insect Biochemistry and Molecular Biology*, *53*, 30–43.

Zaki, M., & Bystroff, C. (Eds.). (2007). *Protein structure prediction*. London: Springer Science & Business Media.

Zhang, H. J., Anderson, A. R., Trowell, S. C., Luo, A. R., Xiang, Z. H., and Xia, Q. Y. (2011). Topological and functional characterization of an insect gustatory receptor. *PLoS One, 6*(8), 24111.

Zhang, Y.-N., Jin, J.-Y., Jin, R., Xia, Y.-H., Zhou, J.-J., Deng, J.-Y., and Dong, S.-L. (2013). Differential expression patterns in chemosensory and non-chemosensory tissues of putative chemosensory genes identified by transcriptome analysis of insect pest the purple stem borer Sesamia inferens (Walker). *PloS One*, *8*(7), e69715.

Zhou, J.-J. (2010). Chapter ten-Odorant-binding proteins in insects. *Vitamins and Hormones*, *83*, 241–272.

Zhou, X., Slone, J. D., Rokas, A., Berger, S. L., Liebig, J., Ray, A., … Zwiebel, L. J. (2012). Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent ant species reveals sex-specific signatures of odor coding. *PLoS Genet*, *8*(8), 1-7.

Zhu, S., & Wang, Y. (2015). Hidden Markov induced Dynamic Bayesian Network for recovering time evolving gene regulatory networks. *Scientific reports*, *5*, 17841.

APPENDICES

Appendix i: *Glossina m. morsitans* ORs and their best protein blast match to *D. melanogaster*, mosquitoes and other *Glossina* species

Table: Odorant receptors (ORs) identified from *Glossina morsitans morsitans* genome and best blast match to Mosquitoes (*Anopheles gambiae, Aedes aegypti, Culex quinquefasciatus*) and *Glossina* (*G. austeni, G. brevipalpis, G. fuscipes*, and *G. pallidipes*) orthologs

***Gmm*OR**	**Best *aaeg* Match (Id, *E*-value)**	**Best *aga*m Match (Id, *E*-value)**	**Best *cqu*e Match (Id, *E*-value)**	**Best *gaut* Match (Id, *E*-value)**	**Best *gbri* Match (Id, *E*-value)**	**Best *gfui* Match (Id, *E*-value)**	**Best *gpai* Match (Id, *E*-value)**
OR1	Orco (Or7) (70.1%, 2e-16)	Orco (Or7) (72.7%, 0.0)	Orco (68.5%, 8e-169)	Orco (85.9%, 0.0)	Orco (Or1) (80.1%, 2e-130)	Orco (97.5%, 0.0)	Orco (95.7, 0.0)
OR2	OR37 (23.3%, 2e-20)	OR34 (23.5%, 5e-21)	OR13 (22.4%, 5e-16)	GAUT045920-RA (93.9%, 0.0)	GBRI035583-RA (84.0%, 0.0)	GFUI028755-RA (91.0%, 0.0)	GPAI004010-RA (95.4%, 0.0)
OR3	lethal giant larva homologue (27.0%, 0.043)	odorant receptor 1 (24.1%, 4e-11)	odorant receptor 1 (28.7%, 7e-09)	GAUT050371-RA (91.5%, 3e-178)	GBRI018062-RA (74.0%, 5e-142)	odorant receptor (85.9%, 5e-121)	odorant receptor (93.4%, 0.0)
OR4	odorant receptor 34 (22%, 2e-16)	odorant receptor 1 (23.5%, 1e-17)	odorant receptor 13 (24.8%, 1e-21)	Odorant receptor 23a (98.2%, 0.0)	Odorant receptor 23a (83.6%, 0.0)	GFUI042981-RA (91.7%, 0.0)	GPAI039747-RA (97.7%, 0.0)
OR5	odorant receptor 34 (24.8%, 2e-10)	odorant receptor 1 (24.2%, 2e-11)	odorant receptor 12 (26.4%, 3e-10)	odorant receptor (71.4%, 3e-178)	GBRI036342-RA (77.4%, 1e-140)	odorant receptor (88.2%, 3e-151)	GPAI034198-RA (93.2%, 6e-164)
OR6	odorant receptor 37 (23.6%, 2e-06)	odorant receptor 34 (22.3%, 1e-04)	odorant receptor 13 (27.0%, 4e-05)	odorant receptor (94.5%, 0.0)	odorant receptor (83.4%, 0.0)	odorant receptor (92.7%, 0.0)	GPAI031315-RA (95.3%, 0.0)
OR7	odorant receptor	odorant receptor 57	odorant receptor	Uncharacterised	Odorant receptor	Uncharacterised	Uncharacteris

	(Or5) (37.3%, 4e-10)	(Or57) (34.9%, 4e-09)	(Or99) (24.3, 1e-07)	protein (89.2%, 0.0)	(69.2%, 0.0)	protein (92.4%, 7.2e-155)	ed protein (97.1%, 3.5e-160)
OR8	odorant receptor 5 (26.5%, 1e-11)	odorant receptor 2 (23.4%, 3e-11)	odorant receptor 13 (24.6%, 2e-11)	GAUT050213-RA (95%, 0.0)	GBRI044639-RA (41.9%, 2e-103)	GFUI003499-RA (47.2%, 2e-97)	GPAI031326-RA (96.5%, 1e-97)
OR9	odorant receptor 8 (23.7%, 1e-10)	odorant receptor 34 (22.7%, 1e-10)	odorant receptor (23.6%, 3e-09)	Odorant receptor 42a (96.2%, 0.0)	Odorant receptor 42b (85.7%, 4e-177)	Odorant receptor 42a (94.6%, 0.0)	GPAI029610-RA (93.8%, 1e-154)
OR10	odorant receptor 34 (23.7%, 3e-20)	odorant receptor 37 (22.5%, 2e-17)	Odorant receptor 94a, putative (27.0%, 5e-20)	GAUT011101-RA (93.4%, 0.0)	odorant receptor (41.0%, 2e-95)	GFUI037305-RA (87.0%, 0.0)	GPAI009882-RA (90.4%,0.0)
OR11	odorant receptor 34 (23.2%, 8e-11)	odorant receptor 1 (21.8%, 4e-10)	odorant receptor 10 (20.3%, 7e-11)	GAUT011101-RA (41.0%, 7e-52)	odorant receptor (81.0%, 9e-161)	odorant receptor (76.1%, 0.0)	GPAI009200-RA (98.1%, 3e-173)
OR12	odorant receptor 37 (25.9%, 1e-31)	odorant receptor 34 (27.9%, 2e-30)	odorant receptor 13 (29.0%, 2e-27)	GAUT005363-RA (94.6%, 1e-61)	odorant receptor (76.8%, 0.0)	GFUI012941-RA (91.6%, 3e-104)	GPAI009882-RA (24.9%, 2e-18)
OR13	odorant receptor (Or31) (26.8%, 2e-25)	odorant receptor 31 (24.7%, 2e-29)	olfactory receptor, putative (30.2%, 2e-35)	odorant receptor (81.7%, 0.0)	GBRI018811-RA (64.6%, 3e-112)	odorant receptor (89.0%, 2e-110)	GPAI024118-RA (84.1%, 0.0)
OR14	odorant receptor (Or31) (22.3%, 1e-17)	odorant receptor 31 (21.3%, 3e-18)	olfactory receptor, putative (21.7%, 7e-23)	odorant receptor (93.8%, 0.0)	odorant receptor (82.6%, 0.0)	odorant receptor (85.4%, 0.0)	GPAI041951-RA (95.8%, 0.0)
OR15	odorant receptor 8 (31.1%, 9e-20)	odorant receptor 8 (28.6%, 3e-15)	olfactory receptor, putative (25.7%, 1e-18)	GAUT022034-RA (97.9%, 0.0)	GBRI026647-RA (82.3%, 0.0)	GFUI032116-RA (98.0%, 3e-98)	GPAI026906-RA (98.9%, 0.0)
OR16	odorant receptor – partial Or39	odorant receptor 28 (22.3%, 1e-11)	odorant receptor OR161 (19.6%, 3e-	GAUT028238-RA (96.1%, 0.0)	GBRI008361-RA (69%, 0.0)	GFUI005658-RA (85.2%, 0.0)	GPAI014680-RA (96.5%,

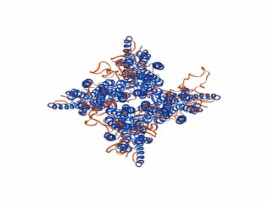
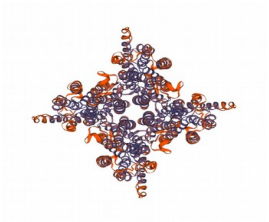
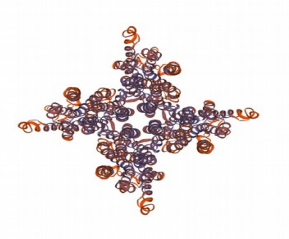
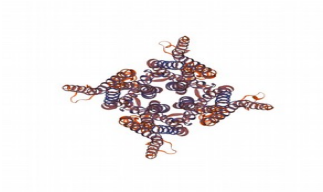
	(22.6%, 7e-10)		09)				0.0)
OR17	odorant receptor 28 (31.4%, 2e-21)	odorant receptor 8 (29.7%, 2e-18)	odorant receptor, putative (28.8%, 1e-18)	odorant receptor (93.1%, 0.0)	odorant receptor (76.2%, 0.0)	GFUI049134-RA (89.3%, 0.0)	odorant receptor (93.3%, 0.0)
OR18	olfactory receptor, putative (24.8%, 2e-25)	gustatory receptor (23.1%, 1e-24)	olfactory receptor, putati (22.8%, 2e-18)	GAUT003629-RA (56.0%, 95.7)	odorant receptor (71.9%, 0.0)	GFUI027054-RA (83.5%, 3e-145)	GPAI017649-(37.4%, 3e-87)
OR19	olfactory receptor, putative (28.1%, 3e-26)	gustatory receptor (26.7%, 5e-23)	olfactory receptor, putative (25.5%, 1e-19)	GAUT003629-RA (81.2%, 2e-51)	GBRI031534-RA (96.9%, 1e-136)	GFUI027050-RA (96.2%, 9e-88)	GPAI017649-RA (98.8%, 0.0)
OR20	No hits	putative GPCR class a orphan receptor 17 (41.2%, 3.6)	olfactory receptor, putative (24.3%, 1.3)	odorant receptor (47.1%, 0.16)	odorant receptor (50%, 0.01)	GFUI047908-RA (47.1%, 0.20)	odorant receptor (43.2%, 0.10)
OR21	odorant receptor 40 (30.4%, 2e-62)	odorant receptor 40 (30.1%, 2e-61)	odorant receptor 91 (29.7%, 9e-63)	odorant receptor (91.2%, 0.0)	odorant receptor (71.8%, 0.0)	GFUI037003-RA (91.6%, 0.0)	GPAI004056-RA (94.4%, 0.0)
OR22	odorant receptor (Or55) (26.1%, 4e-11)	odorant receptor 37 (26.2%, 2e-08)	Odorant receptor 94a, putative (24.3%, 1e-08)	odorant receptor (92.7%, 1e-131)	GBRI013056-RA (77.1%, 4e-78)	odorant receptor (90.3%, 3e-107)	GPAI027550-RA (91.6%, 6e-107)
OR23	lysosomal alpha-mannosidase (mannosidase alpha class 2b member 1) (41.6%, 0.0)	lysosomal alpha-mannosidase (39.5%, 0.0)	lysosomal alpha-mannosidase precursor (41.9%, 0.0)	gustatory and odorant receptor (56.3%, 0.0)	GBRI016960-RA (63.6%,0.0)	gustatory and odorant receptor (56.1%, 0.0)	gustatory and odorant receptor (55.8%, 0.0)
OR24	odorant receptor 8 (30.5%, 1e-47)	odorant receptor 8 (30.4%, 3e-51)	odorant receptor (31.7%, 3e-51)	odorant receptor (95.1%, 0.0)	odorant receptor (88.4%, 0.0)	GFUI047908-RA (92.6%, 0.0)	odorant receptor (94.3%, 0.0)
OR25	odorant receptor 8	odorant receptor 8	odorant receptor	odorant receptor	odorant receptor	odorant receptor	odorant

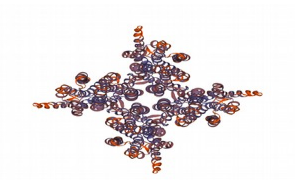
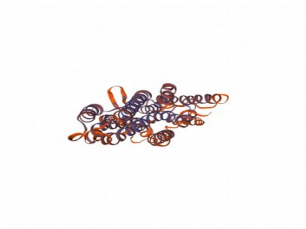
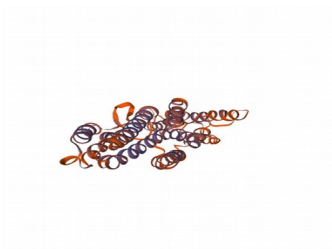
	(22.4%, 4e-14)	(24.6%, 6e-19)	(21.4%, 1e-13)	(93.3%, 0.0)	(63.2%, 1e-170)	(93.5%, 0.0)	receptor (95.8%, 0.0)
OR26	arginine-rich protein, putative (80.4%, 0.061)	inositol polyphosphate 1-phosphatase (91.7%, 0.63)	Kakapo (91.3%, 2.0)	GAUT043805-RA (78.0%, 3e-50)	GBRI027004-RA (77.8%, 0.0)	GFUI016705-RA (92.1%, 0.0)	GPAI001626-RA (94.9%, 3e-139)
OR27	cysteine-rich venom protein, putative (100%, 0.14)	AGAP001287-RC (89.3%, 1.4)	phd finger domain (95.2%, 1.3)	Odorant receptor (91.2%, 6e-152)	GBRI026158-RA (81.1%, 9e-131)	GFUI022126-RA (87.3%, 5e-173)	GPAI033169-RA (96.0%, 0.0)
OR28	polyadenylate-binding protein (91.3%, 1.0)	AGAP001287-RC (89.3%, 0.89)	zinc transporter (100%, 2.8)	Odorant receptor (94.9%, 9e-65)	GBRI026158-RA (87.6%, 4e-51)	GFUI022126-RA (93.1%, 2e-62)	GPAI033169-RA (98.1%, 1e-71)
OR29	shc transforming protein (91.3%, 2.9)	AGAP009554-RA (95.2%, 2.5)	bromodomain-containing protein 7 (100%, 2.3)	GAUT038273-RA (89.3%, 0.0)	GBRI002179-RA (80.7%, 5e-42)	GFUI051696-RA (85.7%, 0.0)	GPAI041241-RA (89.2%, 0.0)
OR30	casein kinase (79.5%, 0.67)	AGAP010825-RA (91.3%, 2.0)	bromodomain-containing protein 7 (100%, 1.9)	GAUT038273-RA (88.5%, 0.0)	GBRI002179-RA (66.6%, 3e-18)	GFUI051696-RA (85.7%, 0.0)	GPAI041241-RA (87.3%, 0.0)
OR31	AAEL012177-RA (95.8%, 0.067)	AGAP002942-RA (100%, 2.4)	conserved hypothetical protein (91.3%, 2.2)	Odorant receptor 24a (83.6%, 0.0)	Odorant receptor 24a (76.7%, 0.0)	GFUI032492-RA (87.3%, 0.0)	Odorant receptor 24a (86.9%, 0.0)
OR32	odorant receptor 8 (29.3%, 3e-37)	odorant receptor 8 (28.7%, 1e-37)	odorant receptor (26.4%, 7e-24)	odorant receptor (89.7%, 0.0)	GBRI045111-RA (86.5%, 0.0)	GFUI014938-RA (92.7%, 0.0)	odorant receptor (95.7%, 0.0)
OR33	odorant receptor 9 (25.2%, 2e-18)	odorant receptor 10 (32.1%, 4e-22)	odorant receptor 120 (30.2%, 7e-21)	odorant receptor (89.6%, 1e-88)	odorant receptor 49b (91.0%, 1e-92)	GFUI009255-RA (95.7%, 2e-41)	GPAI004557-RA (97.1%, 3e-42)
OR34	catrin, putative (91.7%, 0.61)	AGAP007646-RA (86.2%, 0.52)	superoxide dismutase 2 (95.2%,	GAUT039769-RA (95.8%, 0.049)	GBRI031439-RA (92.9%, 0.004)	mitochondrial Fe/S cluster exporter	GPAI001497-RA (90.1%,

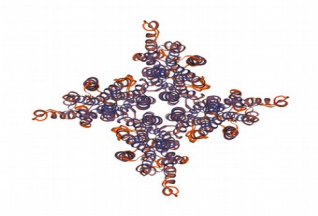
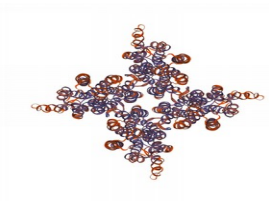
			1.7)			(92.3%, 0.059)	0.0)
OR35	odorant receptor Or9 (77.1%, 8e-13)	odorant receptor 2 (87.1%, 0.052)	isoleucyl tRNA synthetase (86.7%, 0.17)	odorant receptor 43a (96.0%, 0.0)	Odorant receptor 43a (90.9%, 0.0)	GFUI003104-RA (96.2%, 0.0)	GPAI039623-RA (84.7%, 0.0)
OR36	Toll-like receptor (95.5%, 0.64)	dynein heavy chain 2, cytosolic (89.3%, 0.16)	t-diRNAhydrouridine synthase (100%, 1.8)	GAUT000836-RA (95.7%, 0.0)	GBRI002449-RA (78.3%, 7e-64)	GFUI003105-RA (89.6%, 0.0)	GPAI039631-RA (97.4%, 0.0)
OR37	AAEL008608-RA (89.7%, 0.23)	NADPH cytochrome P450 reductase (91.3%, 2.4)	malic enzyme (81.1%, 2.2)	Odorant receptor 74a (95.5%, 0.0)	Pitslre (88.6%, 0.07)	GFUI022472-RA (93.3%, 0.0)	GPAI026228-RA (100%, 0.024)
OR38	cytochrome P450 (86.5%, 2.4)	isoleucyl-tRNA synthetas (100%, 7.2)	conserved hypothetical protein (88.5%, 1.9)	GAUT016620-RA (95.6%, 0.0)	GBRI026891-RA (82.8%, 0.0)	GFUI045476-RA (92.8%, 0.0)	GPAI039539-RA (96.7%, 0.0)
OR39	odorant receptor (Or5) (22.2%, 2e-15)	odorant receptor 6 (Or6) (22.8%, 2e-11)	olfactory receptor, putative (29.5%, 4e-11)	Odorant receptor (94.6%, 0.0)	Uncharacterised protein (77.8%, 7.7e-80)	Uncharacterised protein (92.9%, 0.0)	Uncharacterised protein (92.9%, 0.0)
OR40	raf (88.9%, 0.52)	kinesin family member 20/23 (91.3%, 1.6)	lipase member I precursor (83.9%, 1.4)	GAUT042364-RA (94.2%, 3e-149)	GBRI011898-RA (85.7%, 0.0)	GFUI038138-RA (92.5%, 1e-143)	GPAI045424-RA (96.5%, 7e-159)
OR41	odorant receptor 45 (24.7%, 2e-30)	odorant receptor 44 (22.7%, 3e-21)	odorant receptor 42 (28.2%, 1e-31)	odorant receptor (77.9%, 0.0)	GBRI017432-RA (65.8%, 6e-94)	GFUI007388-RA (76.5%, 2e-123)	GPAI012943-RA (66.1%, 2e-95)
OR42	odorant receptor 45 (23.9%, 1e-31)	odorant receptor 44 (23.6%, 6e-24)	odorant receptor 42 (28.2%, 4e-33)	odorant receptor (87.3%, 0.0)	GBRI017432-RA (67.9%, 2e-93)	GFUI007388-RA (84.8%, 3e-137)	GPAI012943-RA (87.6%, 2e-120)
OR43	AAEL001272-RA (91.7%, 0.72)	PH-sensitive chloride channel (95.8%, 0.051)	fibroblast growth factor receptor 3 precursor (95.5%, 0.57)	odorant receptor (91.0%, 0.0)	GBRI017432-RA (79.4%, 2e-122)	GFUI007397-RA (88.7%, 0.0)	GPAI012943-RA (95.3%, 0.0)

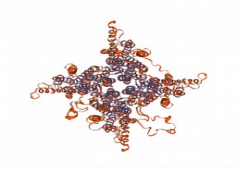
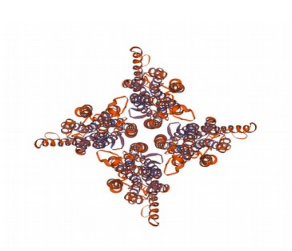
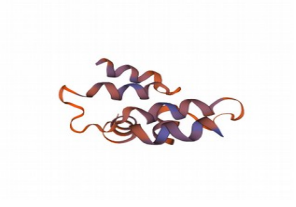
OR44	odorant receptor 45 (25.4%, 7e-36)	odorant receptor 72 (23.2%, 7e-33)	odorant receptor 161 (27.8%, 6e-51)	odorant receptor (35.0%, 3e-70)	GBRI037236-RA (78.7%, 6e-83)	odorant receptor (91.6%, 5e-166)	GPAI046202-RA (92.6%, 0.0)
OR45	ras-related protein Rab-7 (85.3%, 0.83)	pre-mRNA-splicing factor CDC5/CEF1 (83.8%, 0.20)	ADP, ATP carrier protein 2 (87.9%, 0.054)	Odorant receptor 67d (92.8%, 0.0)	Odorant receptor 67d (81.2%, 0.0)	GFUI036188-RA (90.9%, 0.0)	GPAI002749-RA (94.4%, 0.0)
OR46	exocyst complex protein exo70 (95.7%, 0.19)	Adaptor-related protein complex 2, beta 1 subunit (81.1%, 0.56)	coatomer, gamma-subunit, putative (78.4%, 6.2)	GAUT029743-RA (94.1%, 3e-107)	odorant receptor (75.7%, 1e-149)	odorant receptor (91.2%, 4e-125)	GPAI042230-RA (96.0%, 0.0)

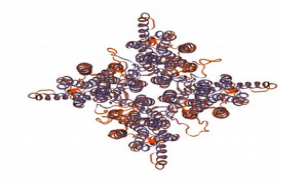
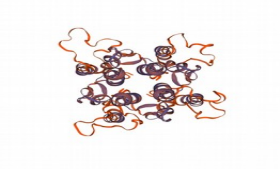
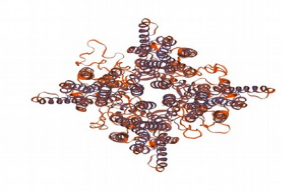
Appendix ii: Protein Structure of the odorant receptors and their respective ligands

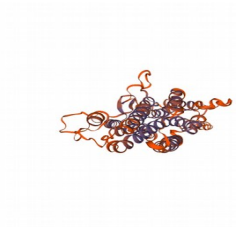
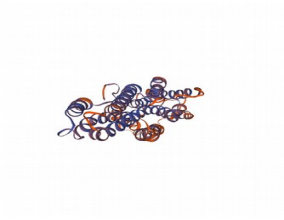
Gmm OR	Gene ID Accession No	Protein Structure	Ligand(s)
OR1	GMOY005610		Cadmium ion, Sulphate ion, Fe(III) ion, 4-(2-Hydroxyethyl)-1-Piperazine ethanesulfonic acid, Iron (II) oxide.
OR2	GMOY005796		Manganese (II) ion, Fe (III) ion, Zinc ion, 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID, PROTOPORPHYRIN IX CONTAINING FE
OR3	GMOY004772		Sulphate ion, Iron (II) ion, PROTOPORPHYRIN IX CONTAINING FE, Potassium ion.
OR4	GMOY012195		No predicted binding pockets
OR5	GMOY012018		otassium ion, Magnesium ɔn, Fe (III) ion, Fe (II) ion, rotoporphyrin IX Containing e
OR6	GMOY009475		Sodium ion, Magnesium ion, Fe (II) ion, Barium ion, Cadmium ion

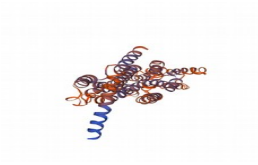
OR7	GMOY013231		No predicted binding pockets
OR8	GMOY012193		No predicted binding pockets
OR9	GMOY012276		Sodium ion, Fe (III) ion, Magnesium ion, TERBIUM(III) ION
OR10	GMOY012214		Sulphate ion, Fe (III) ion, PROTOPORPHYRIN IX CONTAINING FE, Potassium ion, Manganese (II) ion
OR11	GMOY010761		No predicted binding pockets
OR12	GMOY009271		No predicted binding pockets

OR13	GMOY003312		Iron (III) ion, Magnesium ion, Sodium ion, 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID
OR14	GMOY001365		4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID, Fe (III) ion, Calcium ion, Potassium ion, Fe (II) ion
OR15	GMOY012287		Iron (II) ion, Iron (III) ion, Potassium ion, PROTOPORPHYRIN IX CONTAINING FE, Sulphate ion.
OR16	GMOY012240		Zinc ion, Potassium ion, Sulphate ion, PROTOPORPHYRIN IX CONTAINING FE, Fe (III) ion
OR17	GMOY005386		PROTOPORPHYRIN IX CONTAINING FE, Fe (III) ion, Fe (II) ion, Potassium ion
OR18	GMOY012323		Adenosine-5' monophosphate, Guanosine-5'-monophosphate, Cytidine-5'-monophosphate, Uridine-5'-monophosphate

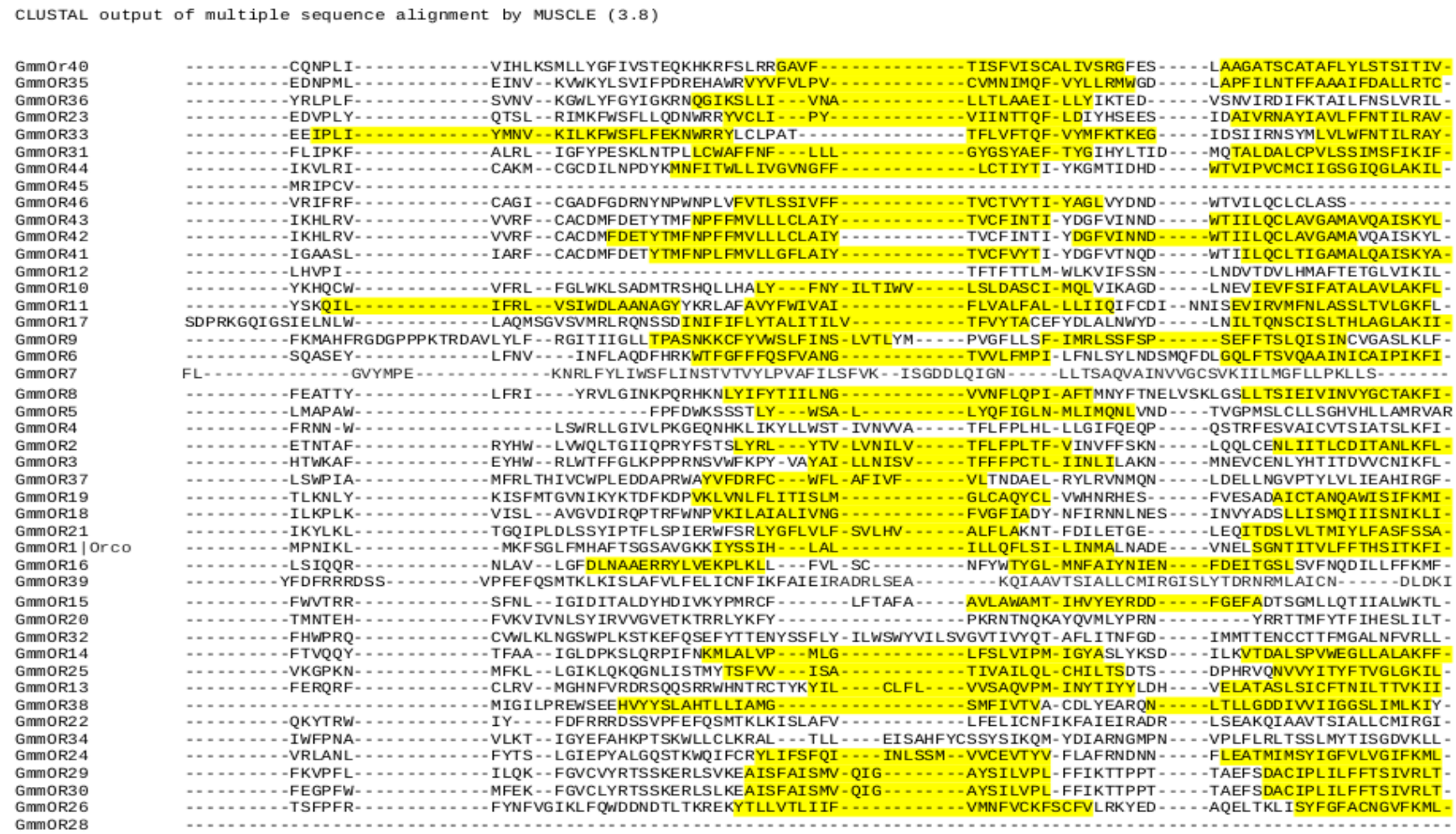
OR19	GMOY012322		Fe (II) ion, PROTOPORPHYRIN IX CONTAINING FE, Magnesium ion, Potassium ion, Sulphate ion
OR20	GMOY012277		No predicted binding pockets
OR21	GMOY011399		Fe (II) ion, Magnesium ion, Meso-2,3-Butanediol, Silver ion
OR22	GMOY012256		
OR23	GMOY011902		No predicted binding pockets
OR24	GMOY010839		Magnesium ion, Sodium ion, Fe (III) ion, Sulphate ion, Glycerol, Fe (II) ion.
OR25	GMOY012357		No predicted binding pockets
OR26	GMOY012255		Protoporphyrin IX containing

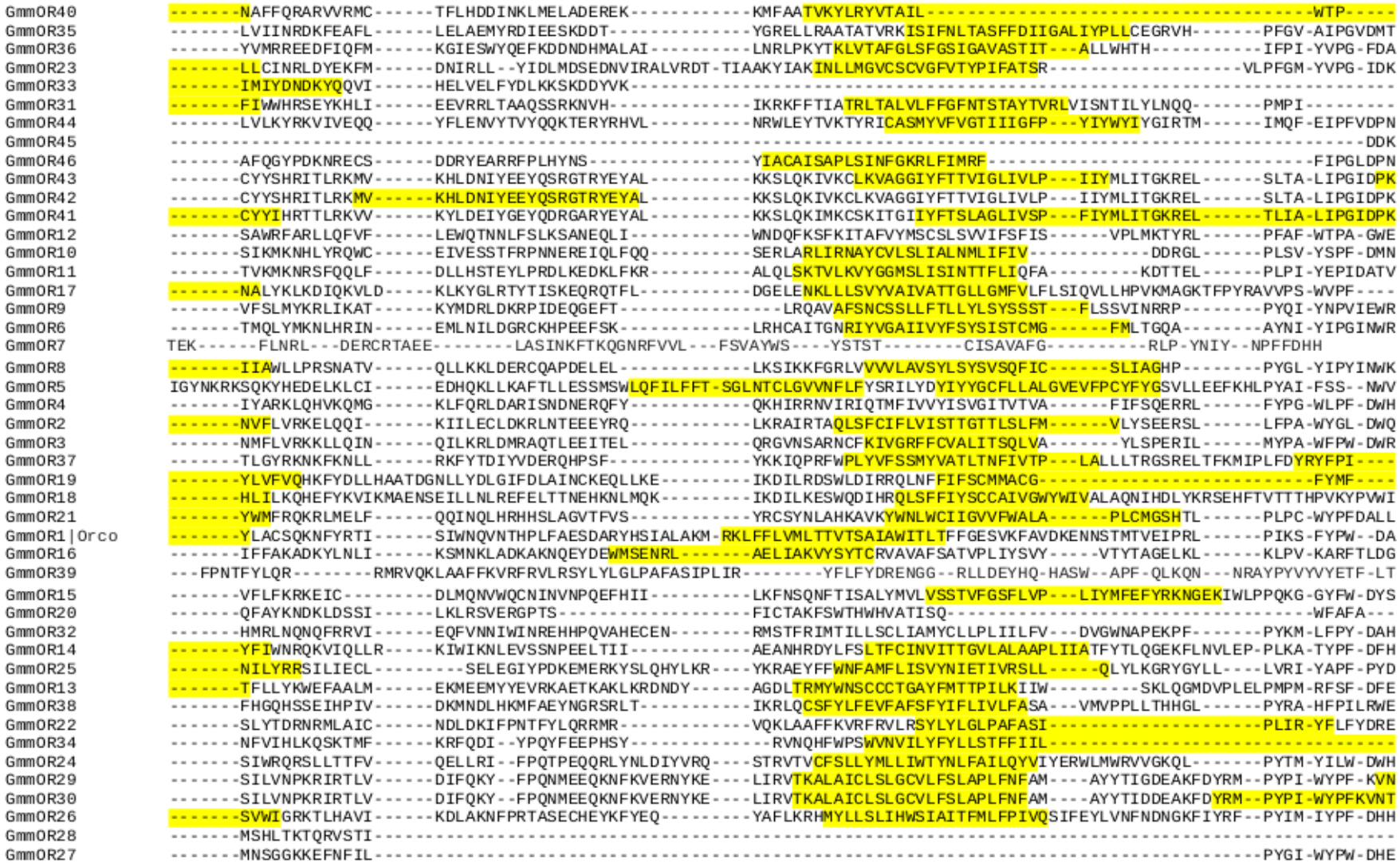
			Fe, Iron (III) ion and Iron (II) ion.
OR27	GMOY008038		THYMIDINE-5'-MONOPHOSPHATE, 2'-DEOXYGUANOSINE-5'-MONOPHOSPHATE, 2'-DEOXYCYTIDINE-5'-MONOPHOSPHATE, Chloride ion.
OR28	GMOY012199		No predicted binding pockets
OR29	GMOY012283		Magnesium ion, Fe (III) ion, Zinc ion, Sodium ion, PROTOPORPHYRIN IX CONTAINING FE.
OR30	GMOY012282		No predicted binding pockets
OR31	GMOY012239		No predicted binding pockets
OR32	GMOY005084		No predicted binding pockets

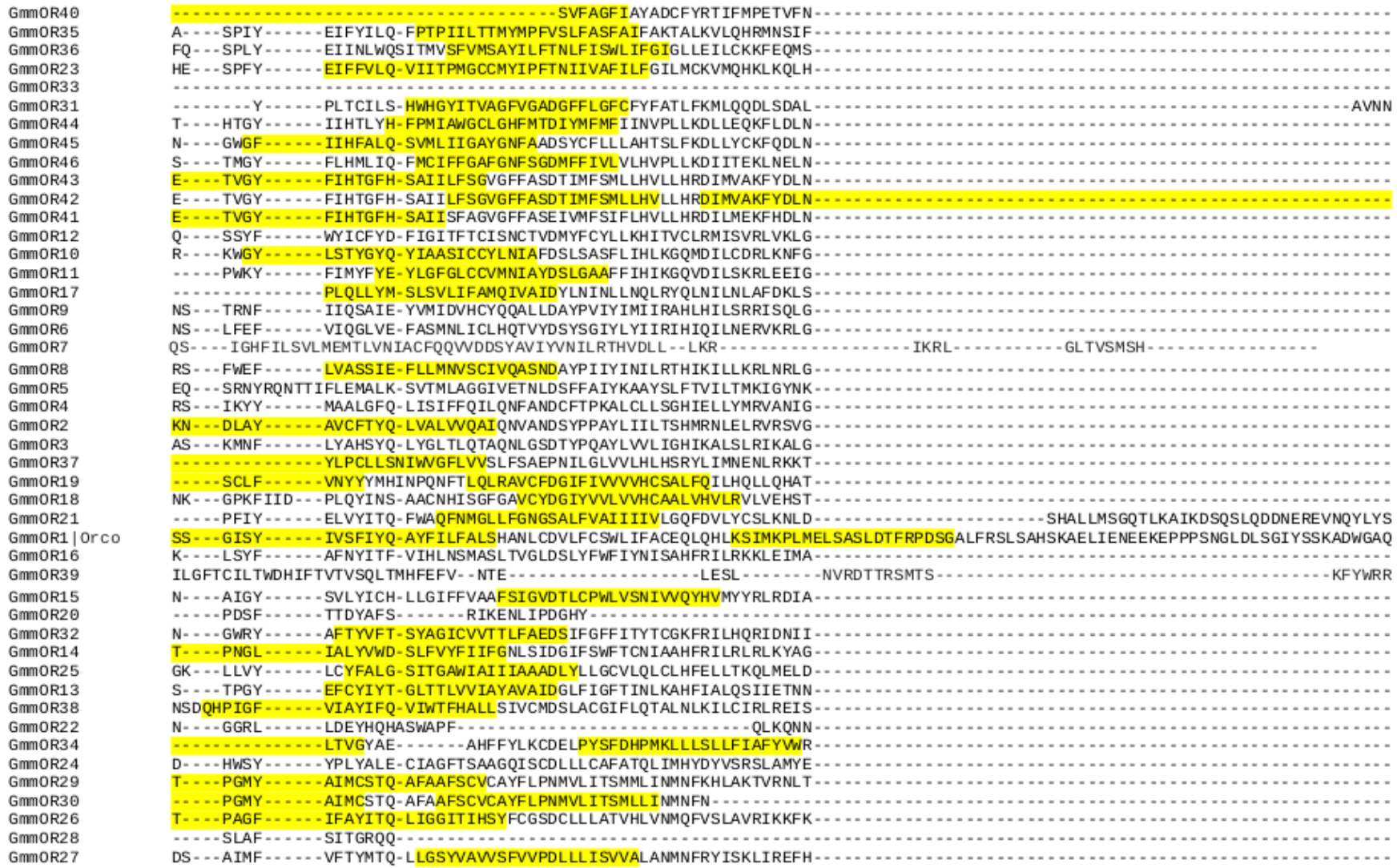
OR33	GMOY005479		Fe (III) ion, Potassium ion, Magnesium ion, Sodium ion, Fe (II) ion
OR34			
OR35	GMOY012253		PROTOPORPHYRIN IX CONTAINING FE, 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID, Iron (III) ion, Iron (II), Calcium ion.
OR36	GMOY012254		No predicted binding pockets
OR37	GMOY012218		No predicted binding pockets
OR38	GMOY012208		Iron (II) ion, Potassium ion, PROTOPORPHYRIN IX CONTAINING FE, ARACHIDONIC ACID, Iron (III) ion.
OR39	GMOY004392		No predicted binding pockets

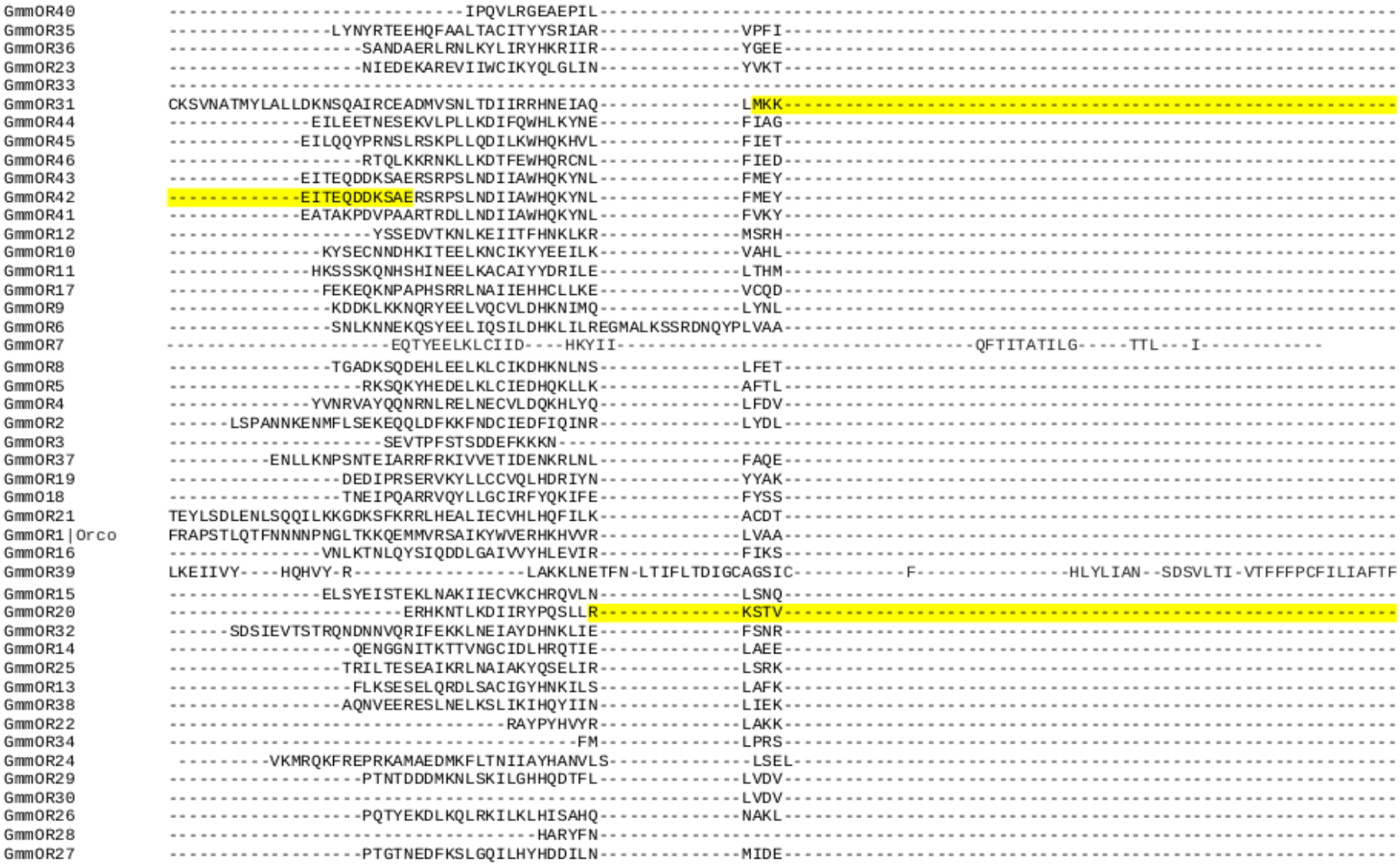
OR40	GMOY012356		No predicted binding pockets
OR41	GMOY006480		No predicted binding pockets
OR42	GMOY006479		TERBIUM(III) ION, 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID, Calcium ion.
OR43	Not Available		
OR44	GMOY006265		Zinc ion, 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID, Iron (III) ion, PROTOPORPHYRIN IX CONTAINING FE, Calcium ion.
OR45	GMOY007896		Protoporphyrin IX containing Fe, Iron (III) ion and Iron (II) ion, Magnesium ion, Sulphate ion, Iron (II) oxide, Adenosine-5’-monophosphate, Zinc ion.
OR46	GMOY003305		Potassium ion, iron (III) ion, 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID, TRIETHYLENE GLYCOL, Protoporphyrin IX containing Fe.

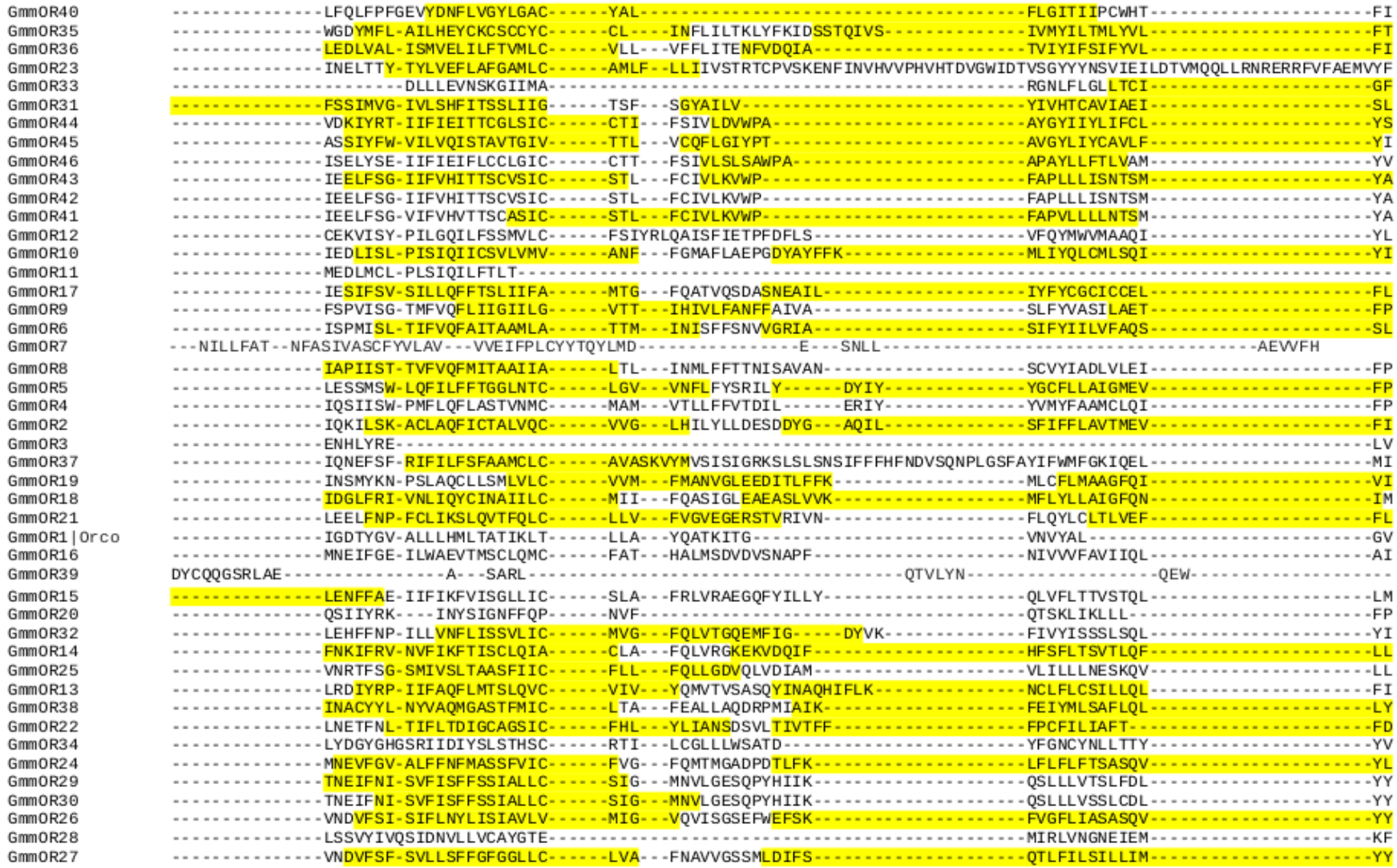
Appendix 3: Multiple Sequence alignment




GmmOR40
GmmOR35
GmmOR36
GmmOR23
GmmOR33
GmmOR31
GmmOR44
GmmOR45
GmmOR46
GmmOR43
GmmOR42
GmmOR41
GmmOR12
GmmOR10
GmmOR11
GmmOR17
GmmOR9
GmmOR6
GmmOR7
GmmOR8
GmmOR5
GmmOR4
GmmOR2
GmmOR3
GmmOR37
GmmOR19
GmmOR18
GmmOR21
GmmOR1|Orco
GmmOR16
GmmOR39
GmmOR15
GmmOR20
GmmOR32
GmmOR14
GmmOR25
GmmOR13
GmmOR38
GmmOR22
GmmOR34
GmmOR24
GmmOR29
GmmOR30
GmmOR26
GmmOR28
GmmOR27


GmmOR40
GmmOR35
GmmOR36
GmmOR23
GmmOR33
GmmOR31
GmmOR44
GmmOR45
GmmOR46
GmmOR43
GmmOR42
GmmOR41
GmmOR12
GmmOR10
GmmOR11
GmmOR17
GmmOR9
GmmOR6
GmmOR7
GmmOR8
GmmOR5
GmmOR4
GmmOR2
GmmOR3
GmmOR37
GmmOR19
GmmOR18
GmmOR21
GmmOR1|Orco
GmmOR16
GmmOR39
GmmOR15
GmmOR20
GmmOR32
GmmOR14
GmmOR25
GmmOR13
GmmOR38
GmmOR22
GmmOR34
GmmOR24
GmmOR29
GmmOR30
GmmOR26
GmmOR28
GmmOR27


GmmOR40
GmmOR35
GmmOR36
GmmOR23
GmmOR33
GmmOR31
GmmOR44
GmmOR45
GmmOR46
GmmOR43
GmmOR42
GmmOR41
GmmOR12
GmmOR10
GmmOR11
GmmOR17
GmmOR9
GmmOR6
GmmOR7
GmmOR8
GmmOR5
GmmOR4
GmmOR2
GmmOR3
GmmOR37
GmmOR19
GmmO18
GmmOR21
GmmOR1|Orco
GmmOR16
GmmOR39
GmmOR15
GmmOR20
GmmOR32
GmmOR14
GmmOR25
GmmOR13
GmmOR38
GmmOR22
GmmOR34
GmmOR24
GmmOR29
GmmOR30
GmmOR26
GmmOR28
GmmOR27


GmmOR40
GmmOR35
GmmOR36
GmmOR23
GmmOR33
GmmOR31
GmmOR44
GmmOR45
GmmOR46
GmmOR43
GmmOR42
GmmOR41
GmmOR12
GmmOR10
GmmOR11
GmmOR17
GmmOR9
GmmOR6
GmmOR7
GmmOR8
GmmOR5
GmmOR4
GmmOR2
GmmOR3
GmmOR37
GmmOR19
GmmOR18
GmmOR21
GmmOR1|Orco
GmmOR16
GmmOR39
GmmOR15
GmmOR20
GmmOR32
GmmOR14
GmmOR25
GmmOR13
GmmOR38
GmmOR22
GmmOR34
GmmOR24
GmmOR29
GmmOR30
GmmOR26
GmmOR28
GmmOR27

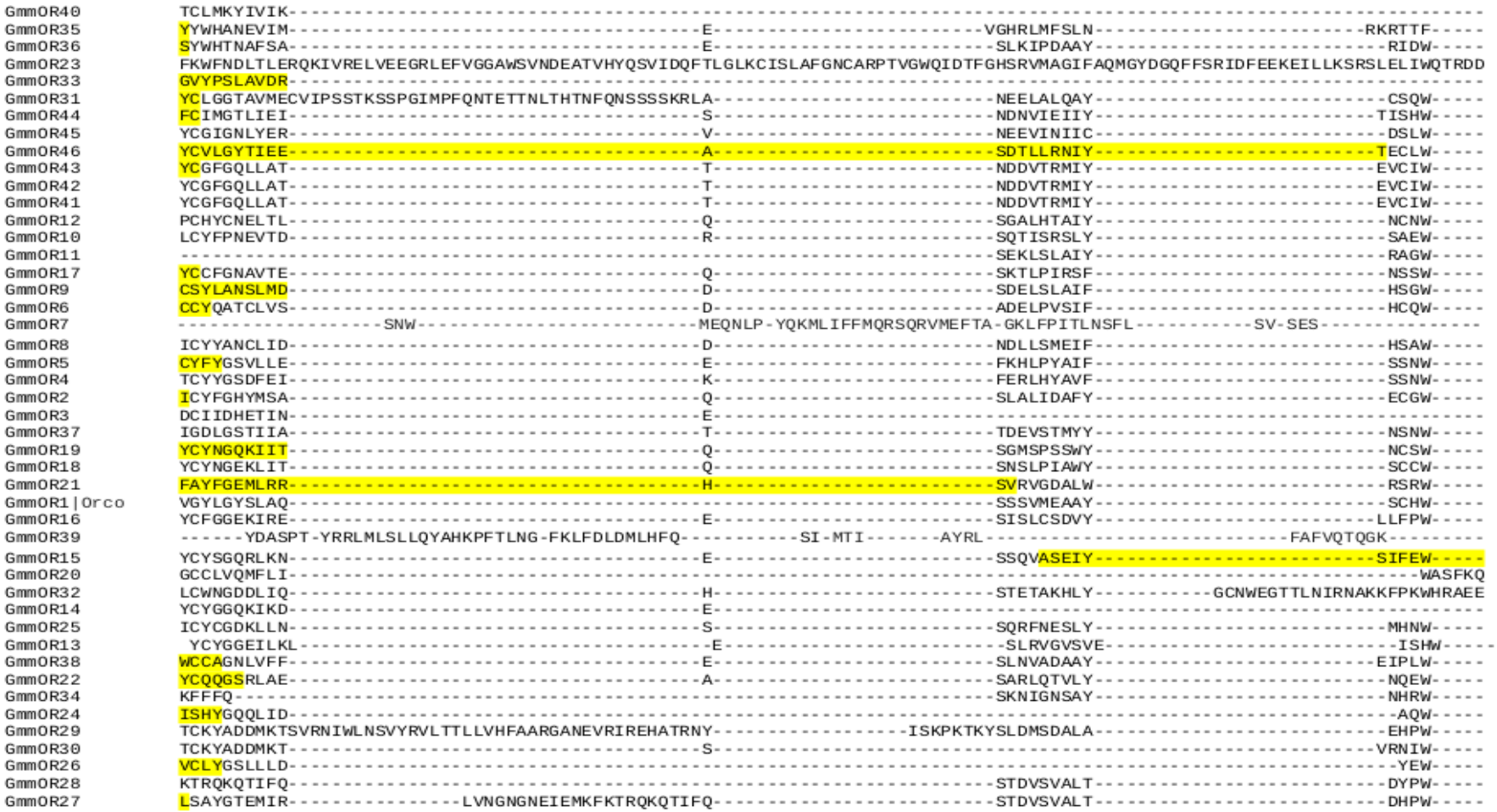


Figure: CLUSTAL output of sequence alignment of *G. m. morsitans* protein sequences by MUSCLE. The highlighted regions in yellow are the conserved transmembrane domains that coincides with the helix coordinates from TMHMM tool.