

https://dspace.org/

The functional genomics of genitalia rotation in male Anopheles funestus

Zawada, Jacek Walerian

2021
https://hdl.handle.net/10539/33542

Downloaded from a DSpace repository. https://dspace.org/

The functional genomics of genitalia rotation in male *Anopheles funestus*

by

Jacek Walerian Zawada

**A thesis submitted to the Faculty of Health,
University of the Witwatersrand,
Johannesburg,
in fulfilment of the requirements for the degree of
Doctor of Philosophy**

Johannesburg, 2021

Declaration

I, Jacek Walerian Zawada, declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor in Philosophy at the University of Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



(*Signature of candidate)*

15th day of October 2021

Dedication

To my sorely missed cousins Bernd-Dieter Hauter and Birgit Du Toit, my father-in-law to be Derrick Griffiths, as well as to my namesake and grandfather Dr. Jacek Walerian Zawada.

May they rest in peace.

Presentations arising from this research project

Zawada, J.W., Dahan-Moss, Y.L., Koekemoer, L.L. **Genetic characterization of males, in a wild *Anopheles funestus* swarm in Zambia.** Presented at the MBRT Research Day on the 8th December 2016, held at the University of the Witwatersrand, Braamfontein, Johannesburg, Gauteng, South Africa.

Zawada, J.W., Dahan-Moss, Y.L., Mbanga, M., Hunt, R., Coetzee, M., Koekemoer, L.L. **Frequency of sexually matured males in a wild *Anopheles funestus* swarm in Zambia.** Presented at the 2nd Annual Malaria Research Conference on the 31st July - 2nd August 2016, held at the University of Pretoria, Groenkloof, Pretoria, Gauteng, South Africa.

Zawada, J.W., Dahan-Moss, Y.L., Mbanga, M., Hunt, R., Coetzee, M., Koekemoer, L.L. **Frequency of sexually matured males in a wild *Anopheles funestus* swarm in Zambia.** Presented at the 2016 Witwatersrand Research Day and Postgraduate Expo on the 1st September 2016, held at the University of the Witwatersrand, Braamfontein, Johannesburg, Gauteng, South Africa.

Zawada, J.W., Dahan-Moss, Y.L., Mbanga, M., Hunt, R., Coetzee, M., Koekemoer, L.L**. Frequency of sexually matured males in a wild *Anopheles funestus* swarm in Zambia.** Presented at the 2016 Molecular Biosciences Research Thrust on the 2nd December 2016, held at the University of the Witwatersrand, Braamfontein, Johannesburg, Gauteng, South Africa.

Dahan-Moss, Y.L., **Zawada, J.W**., Koekemoer, L.L**. From esterase's to mating in *Anopheles funestus.*** Presented at the Molecular and Population Biology of Mosquitoes and Other Disease Vectors conference on the 24th – 28th July 2017, held at the Orthodox Academy of Crete, Kolymbari, Chania, Crete, Greece.

Zawada, J.W., Dahan-Moss, Y.L., Koekemoer, L.L**. Genetic characterization of males in a wild *Anopheles funestus* swarm in Zambia.** Presented at the 3rd International 2017 PARSA conference, held on the 24th – 27th September 2017 at Skukuza in the Kruger National Park, Mpumalanga, South Africa.

Zawada, J.W., Dahan-Moss, Y.L., Koekemoer, L.L. **Functional genomics of male *Anopheles funestus* genitalia rotation.** Presented at the 3rd Annual Malaria Research Conference on the 7 - 9th November 2017, held at the National Institute for Communicable Diseases, Sandringham, Gauteng, South Africa.

Zawada, J.W., Dahan-Moss, Y.L., Koekemoer, L.L**. Genetic characterization of males in a wild *Anopheles funestus* swarm in Zambia.** Presented at the 2017 Molecular Biosciences Research Thrust on the 2nd December 2017, held at the University of the Witwatersrand, Braamfontein, Johannesburg, Gauteng, South Africa.

Zawada, J.W., Dahan-Moss, Y.L., Koekemoer, L.L. **Functional genomics of male *Anopheles funestus* genitalia rotation.** Presented at the South African Society of Biochemistry and Molecular Biosciences on the 8th – 11th of July 2018, held at the North West University, Potchefstroom, North West, South Africa.

Publications arising from this research project

Zawada, J. W., Dahan-Moss, Y. L., Muleba, M., Dabire, R. K., Maïga, H., Venter, N., Davies, C., Hunt, R.H., Coetzee, M. and Koekemoer, L. L. (2018). **Molecular and physiological analysis of *Anopheles funestus* swarms in Nchelenge, Zambia**. *Malaria Journal*, 17(1), 49. Malaria Journal (Impact factor: 3.109). https://doi.org/10.1186/s12936-018-2196-6

Role in publication: Conducted research, analysed data, assisted with the interpretation of data and wrote first and subsequent drafts of the paper.

Zawada, J.W., Dahan-Moss, Y.L and Koekemoer, L.L. **Adolescent transcriptomic study in male African malaria vector mosquitoes.** (In process).

Role in publication: Conducted research, analysed data, assisted with the interpretation of data and wrote first and subsequent drafts of the paper.

Publications indirectly linked to this research project

Dahan-Moss, Y., Hendershot, A., Dhoogra, M., Julius, H., **Zawada, J.,** Kaiser, M., Lobo, N.F., Brooke, B.D. and Koekemoer, L. L. (2020). **Member species of *the Anopheles gambiae* complex can be misidentified as *Anopheles leesoni***. *Malaria Journal*, *19*(1), 1-9. Malaria Journal (Impact factor: 3.109). https://doi.org/10.1186/s12936-020-03168-x

Role in publication: involved with the proficiency testing done at the laboratory and read and approved final manuscript.

Mozūraitis, R., Hajkazemian, M., **Zawada, J. W**., Szymczak, J., Pålsson, K., Sekar, V., Biryukova, I., Friedländer, M.R., Koekemoer, L.L., Baird, J.K., Borg-Karlson, A. and Noushin Emami, S. (2020). **Male swarming aggregation pheromones increase female attraction and mating success among multiple African malaria vector mosquito species**. *Nature Ecology and Evolution*, 1-7. Nature Ecology and Evolution (Impact factor: 12.541). https://doi.org/10.1038/s41559-020-1264-9

Role in publication: designed the field experiments, collected the field data, and provided comments on the manuscript.

Manach, C., Dam, J., Woodland, J.G., Kaur, G., Lutete P., Khonde, L.P., Brunschwig, C., Njoroge, M., Wicht, K.J., Horatscheck, A., Paquet, T., Boyle, G.A., Gibhard, L., Taylor, D., Lawrence, N., Yeo, T., Mok, S., Eastman, R.T., Dorjsuren, D., Talley, D.C., Guo, H., Simeonov, A., Reader, J., Van der Watt, M., Erlank, E., Venter, N., **Zawada, J.W.**, Aswat, A., Nardini, L., Coetzer, T.L., Lauterbach, S.B., Bezuidenhout, B.C., Theron, A., Mancama, D., Koekemoer, L.L., Birkholtz, L-M., Wittlin, S., Delves, M., Ottilie, S., Winzeler, E.A., Von Geldern, T.W., Smith, D., Fidock, D.A., Street, L.J., Basarab, G.S., Duffy, J., and Chibale, K. (2021). **Identification and Profiling of a Novel Diazaspiro [3.4] octane Chemical Series Active against Multiple Stages of the Human Malaria Parasite *Plasmodium falciparum* and Optimization Efforts**. *Journal of Medicinal Chemistry, 64(4),* 2291-2309. Journal of Medicinal Chemistry. (Impact Factor: 6.205). DOI: 10.1021/acs.jmedchem.1c00034

Role in publication: tested and characterized the transmission blocking activity of the compounds.

Abstract

Malaria is at the forefront of diseases transmitted by mosquito vectors worldwide with approximately 409, 000 deaths and over 229 million cases in 2019. Malaria is caused by the *Plasmodium* parasite and is transferred between human hosts by mosquito vectors. The reproductive cycle of the *Plasmodium* parasite takes place within the human host and the *Anopheles* mosquito vector. The *Anopheles funestus* malaria vector is anthropophilic and endophilic (i.e., preferring indoor human feeding). In several African countries a concern for the sustainable control of *An. funestus* is problematic due to the rapid development of its insecticide resistance. A lack of research of this species' biology limits the development of novel control interventions and this fact forms the rationale to study the biology of the *An. funestus* male mosquito. This includes understanding the adult male mosquito's sexual characteristics, physiology, and swarming behaviour, such as genitalia rotation, antennal fibrillae erection and male accessory gland development. The *An. funestus* male genitalia rotation is an important step prior to mating. The genitalia of the newly emerged *An. funestus* male has to rotate between 135° and 180° for the male to become sexually mature. Since the inversion of the genitals is of great importance to the process of sexual maturation and reproduction, an understanding of the underlying genetic mechanisms of male genitalia rotation in this mosquito could be fundamental to vector biology studies.

The aim of this study was to add to the scientific body of knowledge on this complex physiological change in the male mosquito. This was achieved by characterising the physical state of inversion in males that actively participated in swarming in a wild *An. funestus* population in Zambia. Wild *An. funestus* swarms have never been molecularly identified to a species level by using multiplex PCR and molecular assays nor has the rotation status of their genitalia been reported.

It was found that immature male anophelines do not participate in swarming until their genitalia have rotated to Stage 4 (90 - 135°), which indicates that behavioural change occurs with the onset of male sexual maturation. This study identified and validated putative transcripts associated with the male *An. funestus* sexual maturation process with the use of RNA-sequencing, next generation sequencing technologies and real-time quantitative PCR. RNA-sequencing found that

several cuticular transcripts that were down regulated in Stage 3, and this was validated with qPCR. A hypothesis was developed that the down regulation of the cuticular transcripts allows for the cuticles to remain soft in order to allow for the genitalia rotation to take place without the cuticle breaking. RNA interference (traditional as well as novel chitosan nanoparticles) was subsequently used as a gene silencing tool to knock-down specific transcripts (RPL19 and AFUN000965). The phenotypic effects of this knock-down were observed with regards to male sexual maturation development of the genitalia rotation, accessory glands, spermatogenesis, and male palp erection.

The results obtained, identified a large number of transcripts associated with the biological sexual maturation in the male *An. funestus* mosquito and could lead to identifying potential targets for novel vector control interventions for malaria.

Acknowledgements

I would like acknowledge Prof. Lizette Koekemoer for being a fantastic supervisor and teacher. Without her support none of this would be possible.

I am also grateful to Dr. Yael Dahan-Moss for being a patient teacher; your teaching is highly valued.

I would like to thank the field teams in Zambia for all the hard work in finding swarms and catching the samples for analysis; it was an experience I will never forget.

I would like to thank everyone at both the Vector Control Reference Laboratory and the WITS Research Institute for Malaria for their help with rearing the mosquitoes and all the advice and help with my studies.

I have been blessed with an exceptionally supportive family and none of this would have been possible without their support, especially Britta Zawada for all of the moral and editing support. Thanks Mom.

Thank you, Megan, for being so patient and supportive of me during this time. I want to especially thank you for pushing me when I needed to be pushed and consoling me when consoling was needed. I want to thank you for having the wisdom of being able to tell when to do which.

I would not have been able to finish my studies if it wasn't for the financial support of the Department of Science and Technology (DST)/National Research Foundation (NRF) Professional Development Programme (PDP) to Lizette Koekemoer; DST/NRF Research Chair Initiative grant to MC, National Research Foundation Incentive Funding ((UID) 85538) to Lizette Koekemoer. A special thanks goes to the NIH International Centre of Excellence for Malaria Research (ICEMR) for the grant to the Johns Hopkins Malaria Research Institute (grant no. U19 AI089680) for the opportunity to collect the wild samples.

Ethics

ANIMAL RESEARCH ETHICS COMMITTEE
Registration number: AREC-101210-002

March 23, 2018

To whom it may concern,

Re: Waiver from the Animal Ethics Research Committee of the University of the Witwatersrand

Jacek Zawada is an MSc candidate undertaking a study titled "Functional Genomics of Genitalia Rotation in Male *Anopheles Funestus*". He is upgrading to a PhD.
The study will be performed at the Vector Control Reference Laboratory at the National Institute for Communicable Diseases on the National Health Laboratory Services campus in Sandringham, Johannesburg.
The research is being conducted under the auspices of the Research Institute for Malaria, affiliated to the School of Pathology, 10th floor of the Faculty of Health Sciences of the University of the Witwatersrand.
The research is being supervised by Prof. Lizette Koekemoer (Malaria Research Institute) and Dr. Yale Dahan-Moss's (Vector Control Reference Laboratory, NICD, Sandringham).

This letter is to confirm that the research being conducted on mosquitoes and does not require full Animal Ethics Research Committee clearance to undertake the work.

Please contact me should you require further information,

Yours sincerely

GP Candy

Geoffrey Candy
Chair : Animal Ethics Research Committee
University of the Witwatersrand

Reference : Jacek Zawada*: Waiver 20180215O*

Geoffrey P Candy PhD
Professor
Department of Surgery
Faculty of Health Sciences, University of the Witwatersrand, *7 York Road Parktown 2193, Johannesburg*
Tel: 27-11-717-2574; Email: geoffrey.candy@wits.ac.za

Table of Contents

The functional genomics of genitalia rotation in male *Anopheles funestus* i
Declaration i
Dedication ii
Presentations arising from this research project iii
Publications arising from this research project iv
Publications indirectly linked to this research project v
Abstract vi
Acknowledgements viii
Ethics ix
Table of Contents x
List of Figures xiv
List of Tables xv
1 Introduction 1
1.1 General introduction 1
1.1.1 Malaria 1
1.1.2 Malaria vector control and insecticide resistance 2
1.1.3 African malaria vectors 4
1.1.3.1 The Anopheles gambiae complex 4
1.1.3.2 The Anopheles funestus group 5
1.1.4 Male mating biology and sexual development 10
1.1.4.1 Arthropod cuticles 10
1.1.4.2 Male accessory glands 12
1.1.4.3 Antennal fibrillae erection 14
1.1.4.4 Genitalia rotation 16
1.1.5 Functional genomics 19
1.2 Research aims and hypotheses 21
1.3 Objectives 22
2 Molecular species identification and physiological characterization of wild *Anopheles funestus* swarms in Zambia 23
2.1 Introduction 23
2.1.1 Anopheline swarming 23

2.2 Rationale 25
2.3 Material and methods 26
2.3.1 Biological material and swarm collection 26
2.3.2 Species identification 27
2.3.2.1 Anopheles funestus group PCR identification 27
2.3.2.2 Anopheles gambiae complex PCR identification 29
2.3.2.3 Anopheles funestus genotyping 30
2.3.2.4 Differentiation between An. gambiae s.s. and An. coluzzii using RFLP identification assay. 32
2.3.3 Clade identification 32
2.3.4 Genitalia rotation observations 34
2.3.5 Spermathecae dissection 34
2.4 Results 35
2.4.1 Biological material preservation and swarm collection 35
2.4.2 Species identification 36
2.4.3 Species population identification and clade characterization 37
2.4.4 Genitalia rotation observations 38
2.4.5 Spermathecae dissection 39
2.5 Discussion 40
3 Transcriptomic studies of male *An. funestus* during sexual maturation 44
3.1 Introduction 44
3.2 Rationale 46
3.3 Materials and methods 47
3.3.1 Biological material, RNA extraction and preparation 47
3.3.2 RNA-sequencing and data production 48
3.3.3 Reference mapping and assembly 49
3.3.4 Differentially expressed gene analysis 50
3.4 Results 52
3.4.1 Assessment of RNA quality and degradation 52
3.4.2 RNA-Sequencing and data production 54
3.4.3 Reference mapping and assembly 56
3.4.4 Differentially expressed gene analysis 59
3.5 Discussion 67

4 Confirmation of transcript abundance of selected genes....72
4.1 Introduction 72
4.2 Rationale 76
4.3 Material and methods 77
4.3.1 Biological material, RNA extraction and preparation 77
4.3.2 Primer design and reference gene selection 77
4.3.3 cDNA synthesis 78
4.3.4 Relative-quantitative Real-time PCR 79
4.4 Results 80
4.4.1 Assessment of RNA quality 80
4.4.2 Primer design and reference gene selection 80
4.4.3 cDNA synthesis 81
4.4.4 Relative-quantitative real-time PCR....82
4.5 Discussion.... 84
5 Investigation of the potential biological function of selected transcripts....86
5.1 Introduction 86
5.2 Rationale 92
5.3 Material and methods 93
5.3.1 The effect of exogenous Mβ2M siRNA on transcript abundance of RPL19 and UNN1 with and without nanoparticles. 93
5.3.1.1 Experimental set up 93
5.3.1.2 siRNA design 93
5.3.1.3 Preparation of chitosan/siRNA nanoparticles.... 94
5.3.1.4 siRNA delivery of nanoparticles 95
5.3.1.5 Knock-down effects on phenotype 96
5.3.1.6 RNA extraction, cDNA synthesis and qPCR.... 97
5.3.2 Effect of RPL19 siRNA and UNN1 siRNA on transcript abundance with and without nanoparticles....98
5.3.2.1 Experimental set up 98
5.3.2.2 siRNA design 98
5.3.2.3 Preparation of chitosan/siRNA nanoparticles.... 99
5.3.2.4 siRNA delivery of nanoparticles 99
5.3.2.5 Knock-down effects on phenotype 99

5.3.2.6 RNA extraction, cDNA synthesis and qPCR 99
5.4 Results 100
5.4.1 The effect of exogenous Mβ2M siRNA treatment on transcript abundance of RPL19 and UNN1 with and without nanoparticles 100
5.4.1.1 Phenotypic observations 100
5.4.1.2 Relative normalized expression of the RPL19 and UNN1 (AFUN000965) transcript treated with Mβ2M siRNA and Mβ2M siRNA-NP compared to the Untreated sample group 101
5.4.2 Effect of RPL19 siRNA and UNN1 siRNA on transcript abundance with and without nanoparticles 103
5.4.2.1 Phenotypic observations 103
5.4.2.2 RNE of the RPL19 and UNN1 (AFUN000965) transcript treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP compared to the sample group treated with Mβ2M siRNA and Mβ2M siRNA-NP, respectively 104
5.5 Discussion 107
6 Conclusions and future research 110
7 References 113
8 Appendix 136
8.1 E.g. *An. funestus* PCR results of swarm 4. ..... **Error! Bookmark not defined.**
8.2 Bioanalyzer results: Bioanalyzer graphs of samples 136
8.3 cDNA concentrations: Working out of cDNA concentrations for qPCR 138
8.4 RNA-seq data 139
8.5 Plagiarism declaration 168

List of Figures

Figure 1.1 Genotypic distribution of *An. funestus* in the Great Rift Valley 8
Figure 1.2 Diagram of male accessory glands 13
Figure 1.3 Example of male accessory glands 14
Figure 1.4 Example of antennal male fibrillae erection 15
Figure 1.5 Inversion/genitalia rotation of the male anopheline terminalia 16
Figure 1.6 Disposition of the two pairs of rotational muscles of an adult male *Aedes aegypti* 17
Figure 1.7 The different stages of male genitalia rotation in *An. funestus* 18
Figure 2.1 Swarming location 27
Figure 2.2 Restriction fragment length polymorphism banding pattern 31
Figure 2.3 Spermathecae dissection 34
Figure 2.4 Photo of swarming area 36
Figure 2.5 Molecular species identification of six wild *An. funestus* swarms 37
Figure 3.1 Native RNA gel of total RNA 52
Figure 3.2 Agilent Technologies 2100 Bioanalyzer graph of St. 0.1, *An. funestus* total RNA 54
Figure 3.3 Expression density plot per sample 57
Figure 3.4 Correlation matrix of all the samples 57
Figure 3.5 Multidimensional Scaling analysis plot 58
Figure 3.6 Heatmap of one-way hierarchical clustering analysis 59
Figure 3.7 Significant count by $2 \leq |FC| \leq -2$ of Stage 3 compared to Stage 0 60
Figure 3.8 Expression level of genes 61
Figure 3.9 Volume plot of expression 62
Figure 3.10 Scatter plot of Avg. fold-change of differentially expressed genes 64
Figure 4.1 Theoretical amplification curve of a single PCR product 74
Figure 4.2 Validation of the normalized expression of RNA-Seq transcripts 83
Figure 5.1 RNAi pathway 88
Figure 5.2 Chitosan nanoparticle and siRNA interaction 90
Figure 5.3 Relative Normalized Expression (RNE) in log2 fold change of the RPL19 and UNN1 transcript in *An. funestus* males treated with Mβ2M siRNA and Mβ2M siRNA-NP 102
Figure 5.4 Relative Normalized Expression (RNE) of the RPL19 and UNN1 transcripts in *An. funestus* males treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP 106

List of Tables

Table 1.1 Table of *An. funestus* group species found on the African continent* 5
Table 2.1 Table of primers used in the *An. funestus* multiplex PCR 28
Table 2.2 Table of primers used in the *An. gambiae* multiplex PCR 29
Table 2.3 Table of primers used in the D3 amplification for RFLP of the *An. funestus* species .. 32
Table 2.4 Table of primers and probes used in the *An. funestus* clade characterization assay 33
Table 2.5 Samples collected from Nchelenge swarms 38
Table 2.6 The frequency of genitalia rotation of swarming males 39
Table 3.1 Nanodrop 2000 results 53
Table 3.2 Agilent technologies 2100 bioanalyzer results 53
Table 3.3 Table of raw data statistics of the RNA-sequencing data of each replicate 55
Table 3.4 Table of mapping data statistics 56
Table 3.5 Number of cuticular transcripts that were differentially expressed and their families.. 63
Table 3.6 Selected differentially expressed transcripts from the RNA-Seq data between Stage 0 and Stage 3 of the male *An. funestus* mosquito 66
Table 4.1 Table of candidate reference genes 78
Table 4.2 Nanodrop 2000 results RNA concentration 80
Table 4.3 Primer sequences for the CPRR1, UNN1, UNN2, UNN3, UNN4 and UNN5 genes ... 81
Table 4.4 Stock cDNA concentrations 82
Table 4.5 qPCR normalized expression results 83
Table 5.1 Table of siRNA designed for the knock-down of Mβ2M 94
Table 5.2 Table of siRNA designed for the knock-down of RPL19 and UNN1 (AFUN000965) transcript 98
Table 5.3 Effect of male sexual maturity phenotype after siRNA treatment 100
Table 5.4 RNE of the RPL19 and UNN1 (AFUN000965) transcript in *An. funestus* males treated with Mβ2M siRNA and Mβ2M siRNA-NP compared to untreated males 102
Table 5.5 Effect of male sexual maturity phenotype after siRNA treatment 104
Table 5.6 RNE of the RPL19 and UNN1 (AFUN000965) transcript in *An. funestus* males treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP when compared to samples treated with an exogenous negative control 106

1 Introduction

1.1 General introduction

1.1.1 Malaria

Malaria is at the forefront of diseases transmitted by mosquito vectors worldwide with approximately 409, 000 deaths and over 229 million cases in 2019 (WHO, World Malaria Report, 2020). Most of these cases (94%) occurred in Africa, although Asia and Latin America, and to a lesser extent the Middle East and Europe are also affected (WHO, World Malaria Report, 2020).

In 1897, the mosquito belonging to the sub-family *Anophelinae* was suggested as the intermediate vector for the malaria parasite (Boyce, 1911). During this time, it was shown that the *Plasmodium falciparum* parasite is transmitted through the infective bite of an *Anopheles* mosquito to humans. This confirmed that malaria is caused by the parasitic protist that belongs to the genus *Plasmodium.* The genus *Plasmodium* has five parasitic species, *Plasmodium falciparum, P. knowlesi, P. malariae, P. ovale* and *P. vivax* that cause malaria in humans (Boyce, 1911; Garnham, 1966). *Plasmodium falciparum* is the main species responsible for malaria in humans and could be fatal if not treated (WHO, World Malaria Report, 2020). Symptoms usually appear 10 - 15 days after the infective mosquito bite and are characterized by fever, chills, headache, and vomiting (WHO, World Malaria Report, 2020). Multiple organs (liver, kidneys, lungs, spleen, and brain) can also be affected by the infection (WHO, World Malaria Report, 2020). Children under the age of five are severely affected by the parasite and account for 67% of malaria deaths in 2019 (WHO, World Malaria Report, 2020). The symptoms for children can be severe anaemia, cerebral infection, or respiratory distress in relation to metabolic acidosis. Other groups that are at risk include pregnant women, people that are HIV positive or have AIDS, and international travellers that are from non-endemic areas due to their reduced immune tolerance (Askling *et al.,* 2005).

Apart from the direct impact on human mortality, health and well-being, malaria has an impact on the economy. Chima *et al.* (2003) showed that there was decreased levels of productivity within home and work environments in countries with malaria. It was also shown that there was

an increased burden on the health care systems (Chima *et al.,* 2003). The disease has been implicated as one of the factors that contribute to the lack of development in African and other third-world countries, also placing an economical strain on them (Sachs and Malaney, 2002). Malaria control interventions are imperative to alleviate the burden on the affected countries. These interventions include timely diagnosis and treatment of the disease in infected people, the use of effective preventative antimalarial medication and the control of the mosquito vectors that transmit the disease, the latter of which is pertinent to this study.

There are serious concerns regarding the parasite's development of resistance to antimalarial medicines, such as sulfadoxine-primethamine and chloroquine (WHO, World Malaria Report, 2020). In the face of chloroquine resistance, artemisinin combined therapy (ACT) is readily available for treatment (Gatton *et al.,* 2004). Currently, there are no licensed vaccines that are effective against malaria (WHO, World Malaria Report, 2020). As with the parasite, the vectors have also developed resistance to the insecticides that are used to control them (Collins and Paskewitz, 1995; Sokhna *et al.,* 2013; WHO, World Malaria Report, 2020). Therefore, integrated, and combined malaria control programs are recommended and necessary to circumvent the currently emerging and ongoing drug resistance of the plasmodium parasite and insecticide resistance of the mosquito vectors.

1.1.2 Malaria vector control and insecticide resistance

Indoor resting mosquitoes were targeted from the beginning of the 1930's for malaria vector control by spraying the inside walls and surfaces of houses with pyrethrum extracted from flowers (Ross, 1936). Subsequently 1,1'-(2,2,2-Trichloroethane-1,1-diyl)bis(4-chlorobenzene) (DDT) replaced pyrethrum and was the main insecticide used during World War II to protect soldiers from arthropods such as lice, mites, fleas and mosquitoes (Berry-Cabán, 2011).

Even today, the main malaria vector control strategies are the use of insecticide-based indoor residual spraying (IRS) and insecticide-treated nets (ITNs) (WHO, World Malaria Report, 2018; Oguna *et al.,* 2020). Both methods are used mainly indoors to reduce human-vector contact and lead to a reduction in transmission of the malaria parasite (Durnez and Coosemans, 2013; Killeen, 2014; Benelli and Beier, 2017; Sougoufara *et al.,* 2017). Other intervention methods such as larvicides are also used (Center for Disease Control, CDC, 2018). All of these methods

are primarily dependent on the use of chemical insecticides. There are five classes of insecticides approved by the WHO (carbamates, neonicotinoids, organochlorines, organophosphates, and pyrethroids) which are the corner stones for vector control programs (Ranson, 2017; WHO, World Malaria Report, 2020). Pyrethroids and organochlorines cause a depolarization of the axonal membrane by preventing the closure of voltage-gated sodium and causing an influx of sodium ions during excitation which in turn leads to repetitive nerve activity and can result in hyperexcitation and death of the mosquito (Coats, 1990; Hougard *et al.*, 2002; Soderlund *et al.*, 2002). Organophosphates and carbamates, on the other hand, inhibit the acetylcholinesterase enzyme. This prevents the breakdown of the neurotransmitter acetylcholine into choline and acetate, which increases the level and duration of the acetylcholine in the neuromuscular junction. This results in parasympathetic and sympathetic overstimulation and causes skeletal muscle paralysis of the mosquito (Coats, 1990; Fukuto, 1990; Soderlund *et al.*, 2002; Mladenovic *et al.*, 2018).

Due to the ability of anophelines to develop diverse resistance mechanisms, these methods have, however, resulted in many resistant populations of anopheline mosquitoes (Brogdon and McAllister, 1998; Blandin *et al.*, 2009; Ranson *et al.*, 2011; Coetzee and Koekemoer, 2013; Sokhna, *et al.*, 2013; Sougoufara *et al.*, 2017).

According to the Insecticide Resistance Mapper (IRMapper), there has been a steady increase of resistance mechanisms across species, especially since 2002 (https://www.irmapper.com/). Novel insecticides need to be developed rapidly in order to get ahead of the insecticide resistance curve. Monitoring of the insecticide resistance within populations and geographical regions can help determine which insecticides are the most effective. This requires stringent and ongoing surveillance of the malaria vectors within these areas (Kelly-Hope *et al.*, 2008; Brooke *et al.*, 2015; Hemingway *et al.*, 2016; Hancock *et al.*, 2018; Vontas and Mavridis, 2019). Continuous surveillance of the malaria vectors also allows countries to implement insecticide resistance management plans on the malaria vector populations (Hemingway *et al.*, 2013; Hemingway *et al.*, 2016; WHO, World Malaria Report, 2020).

1.1.3 African malaria vectors

The primary vectors for malaria are species within the genus *Anopheles* (Gillies and DeMeillon, 1968; Coetzee and Fontenille, 2004). The two main malaria taxonomic groups for the vector species in Africa are the *An. gambiae* complex and the *An. funestus* group. Although other vectors such as *An. nili, An. marshallii, An. moucheti, An. bancroftii,* etc. (Giles, 1902; Theobald, 1903; Gillies and DeMeillon, 1968; Gillies and Coetzee, 1987; Irish *et al.,* 2020), are also important, they will not be discussed further in more detail.

1.1.3.1 The Anopheles gambiae complex

There are three main vectors within the *An. gambiae* species complex, namely, *An. arabiensis, An. coluzzii* and *An. gambiae.* The saltwater breeders such *An. melas* and *An. merus* are minor vectors (Sinka *et al.,* 2011; Coetzee *et al.,* 2013). Other members in this group are *An. bwambiae, An. quadriannulatus, An. Amharicus,* and *An. fontenillei* which have not been implicated in malaria transmission (Giles, 1902; White 1985; Brunhes and Geoffroy 1997; Coetzee *et al.,* 2013a; Barrón *et al.,* 2019; Irish *et al.,* 2020).

At least nine species make up the *An. gambiae* species complex (Giles, 1902; White 1985; Brunhes and Geoffroy 1997; Coetzee *et al.,* 2013a; Barrón *et al.,* 2019; Irish *et al.,* 2020). These nine species have been reviewed and classified as morphologically indistinguishable (Giles, 1902; White 1985; Gillies and Coetzee, 1987; Brunhes and Geoffroy 1997; Coetzee *et al.,* 2013a; Barrón *et al.,* 2019; Irish *et al.,* 2020). It has been shown that there are five chromosomal forms of the *An. gambiae sensu stricto* (Bamako, Bissau, Forest, Mopti, and Savannah) (Coluzzi *et al.,* 1985). Based on polymorphic chromosomal inversions, the five chromosomal forms, are on the arm of chromosome 2 (Coluzzi *et al.,* 1985). Two molecular forms are in the intergenic spacer (IGS) and internal transcriber spacer (ITS) regions at the centromere of the X chromosome based on specific ribosomal DNA (rDNA) single nucleotide polymorphisms (SNPs) (Gentile *et al.,* 2001; Della Torre *et al.,* 2002). Due to binomial evidence such as chromosomal inversions and pre-mating reproductive isolation mechanisms, the two molecular forms, the Mopti (M-form) and the Savannah (S-form) have recently been renamed (Coetzee *et al.,* 2013a). The S- form, *Anopheles gambiae* Giles kept its name for the *Anopheles gambiae s.s.* and the M- form was renamed *An. coluzzii* Coetzee and Wilkerson sp. n (Coetzee *et al.,* 2013a). For the remainder of

the thesis, *An. coluzzii* will refer to the M- molecular form whereas *An. gambiae s.s.* will refer to the S- molecular form.

1.1.3.2 The Anopheles funestus group

The *An. funestus* group has at least 11 recognized African species: *Anopheles aruni*, *An. brucei*, *An. confusus*, *An. funestus s.s.*, *An. funestus*-like, *An. fuscivenosus*, *An. leesoni*, *An. parensis*, *An. vaneedeni*, *An. rivulorum*, *An. rivulorum*-like and is broadly distributed across the African continent, as can be seen in Table 1.1 (Gillies and DeMeillon, 1968; Gillies and Coetzee, 1987; Hackett *et al.*, 2000; Cohuet *et al.*, 2003; Harbach, 2004; Spillings *et al.*, 2009; Vezenegho *et al.*, 2013; Kyalo *et al.*, 2017; Irish *et al.*, 2020).

Table 1.1 Table of *An. funestus* group species found on the African continent*

Species name	Vector status	Main blood meal source	Distribution
An. aruni (Sobti, 1968)	non-vector	Zoophilic	Extremely rare and reported only in Zanzibar
An. brucei (Service, 1960)	non-vector	Zoophilic	Extremely rare and only recorded from Nigeria
An. confusus (Evans and Leeson, 1935)	non-vector	Zoophilic	Eastern and southern Africa
An. funestus s.s. (Giles, 1900)	primary vector	anthropophilic and endophilic	Broadly distributed across sub-Saharan Africa
An. funestus-like (Spillings *et al.*, 2009)	non- vector	Zoophilic	Only found in Malawi
An. fuscivenosus (Leeson, 1930)	non-vector	Zoophilic	Extremely rare and described only from Zimbabwe
An. leesoni (Evans, 1931)	secondary vector	Zoophilic	Broadly distributed across sub-Saharan Africa
An. parensis (Gillies, 1962)	secondary vector	Zoophilic	Broadly distributed across sub-Saharan Africa
An. vaneedeni (Gillies and Coetzee, 1987)	secondary vector	anthropophilic and zoophilic	Found in South Africa and Kenya
An. rivulorum (Leeson,1935)	secondary vector	anthropophilic and zoophilic	Broadly distributed across sub-Saharan Africa
An. rivulorum-like (Cohuet *et al.*, 2003)	non-vector	Zoophilic	Cameroon, Burkina Faso, and South Africa but might extend to western and central Africa

*The data presented in Table 1.1 was taken from the following references: Gillies and DeMeillon, 1968; Gillies and Coetzee, 1987; Cohuet *et al.*, 2003; Harbach, 2004; Spillings *et al.*, 2009; Vezenegho *et al.*, 2013; Kyalo *et al.*, 2017 and Irish *et al.*, 2020.

The ability to transmit the *Plasmodium* parasite efficiently through its feeding and resting behaviours is what makes *An. funestus s.s.* a major vector of malaria in Africa (Gillies and DeMeillon, 1968; Gillies and Coetzee, 1987; Irish *et al.,* 2020). This vector, like others, bites and feeds between the dusk and dawn period. Only the female anopheline mosquito has the ability to feed on blood, thus only the female can carry the *Plasmodium* parasite. These species are highly anthropophilic (feeds on humans) and endophilic (rests indoors) and are therefore efficient vectors for malaria transmission (Gillies and DeMeillon, 1968; Gillies and Coetzee, 1987). The *Anopheles funestus* group breeds in permanent water bodies and is widely distributed over the African tropics (Gillies and DeMeillon, 1968; Gillies and Coetzee, 1987; Irish *et al.*, 2020). One of the difficulties in controlling the African anopheline vector population, is the difficulty in identifying the vectors to a species level (Gillies and Coetzee,1987). The difficulty in identifying the vectors within groups or complexes to a species level is that morphologically they are very similar and are sometimes identical to each other.

Gillies and DeMeillon (1968) first developed keys to identify anophelines in Africa, which were later updated by Gillies and Coetzee (1987). Gillies and Coetzee (1987) developed dichotomous keys that were recently updated to include the newly named species (Coetzee, 2020). However, mosquitoes belonging to the *An. funestus* group can only be partially identified using these keys and hence species identification is mainly through the use of a standard multiplex PCR that is able to identify the main members of the *An. funestus* group (Koekemoer *et al.,* 2002; Cohuet *et al.,* 2003). This is also true for the identification of members of the *An. gambiae* complex (Scott *et al*., 1993).

The standard multiplex PCR is designed to amplify six different species of the *An. funestus* group by using a unique set of primers that target the different species-specific internal transcribed spacer 2 (ITS2) regions (Koekemoer *et al.*, 2002; Cohuet *et al.,* 2003). The following member species of the *An. funestus* group, namely, *An. funestus, An. rivulorum, An. rivulorum*-like, *An. parensis, An. vaneedeni*, as well as *An. leesoni* will be targeted in this study (Koekemoer *et al.,* 2002; Cohuet *et al.,* 2003). The multiplexed assay allows for amplification of a DNA fragment that has a unique size in each of the species (Koekemoer *et al.,* 2002). The ability to identify the anopheline species allowed researchers to determine spatial geographical data and trends of specific anopheline species populations (Ayala *et al.,* 2011). These trends demonstrated that

population structuring was occurring between populations of the same species and not only between species of anophelines (Ayala *et al.*, 2011).

Anopheles funestus s.s. has shown population structuring within chromosomal rearrangement (reviewed in Coetzee and Koekemoer, 2013), mitochondrial microsatellites and mitochondrial clades (reviewed in Coetzee and Koekemoer, 2013). In addition to this, sequence variation in the ribosomal DNA (rDNA) region was demonstrated using restriction fragment length polymorphism (RFLP) genotyping (Garros *et al*., 2004; Koekemoer *et al*., 2006), as well as sequence variation in the mitochondrial genes that revealed two specific clade structures (Michel *et al*., 2005; Choi *et al*., 2013).

Anopheles funestus s.s. can be differentiated by rDNA RFLP typing, mainly into M-; W-; MW-; Y-; and Z-type populations and these populations are associated with the species geographical distribution (Figure 1.1). The eastern side of the Great Rift Valley contains the M-type *An. funestus* and has been reported in Kenya, Mali, and Madagascar (Koekemoer *et al.*, 2006). Angola, Malawi, Mozambique, and South Africa and have reported that they contain the MW-type. The western side of the Great Rift Valley that spans into Kenya contains the widespread W-type. The Y-type can be found at the southern tip of the Great Rift Valley in Malawi. One Z-type specimen was found in western Kenya (Koekemoer *et al.*, 2006).

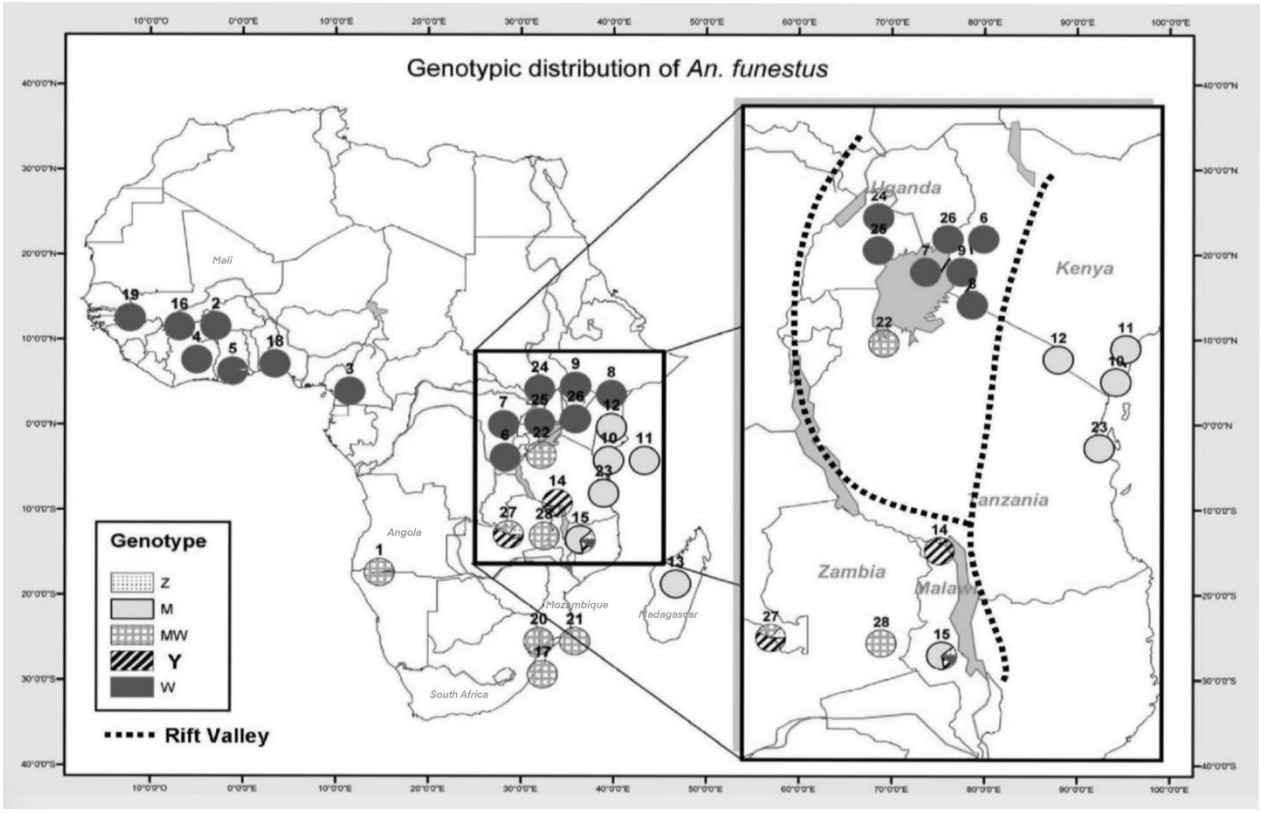


Figure 1.1 Genotypic distribution of *An. funestus* in the Great Rift Valley

Genotypic distribution of *An. funestus* in the Great Rift Valley. The figure shows the distribution of the Z, Y, M, W, and MW molecular types across the Great Rift Valley region in Africa (figure directly taken from Koekemoer *et al.*, 2006).

The two mitochondrial DNA lineages of *An. funestus s.s.* that have evolved independently for ~1 million years, designated clade I and clade II, were described in Michel *et al.* (2005) and Choi *et al.* (2013). Mitochondrial clade analysis compared to the RFLP analysis, showed similar population structuring with regards to geospatial organisation (Michel *et al.*, 2005; Koekemoer *et al.*, 2006; Choi *et al.*, 2013). Choi *et al.* (2013) subsequently described a hydrolysis probe assay to easily differentiate between the clades (Choi *et al.*, 2013). There were 12 widespread African countries that reported clade I, while a combination of clade I and clade II has only been reported from Madagascar, Mozambique, and Zambia (Choi *et al.*, 2013; Choi, 2015; Michel *et al.*, 2005). The proportion varies, of clade I : clade II, between the countries (Choi *et al.*, 2013; Choi, 2015; Michel *et al.*, 2005). There was a ratio of 80% to 20% for clade I and clade II in the *An. funestus* population from Zambia (Choi, 2015). In contrast, 60% of the *An. funestus* population in Mozambique and Madagascar were clade II, with clade I the remaining 40% (Choi *et al.*, 2013;

Choi, 2015). It was thought that clade type analysis could be used in assessing any association between the clade types and resistance to insecticides (Choi, 2015).

The first reports of the *An. funestus* species showing insecticide resistance, specifically to organophosphates, appeared in 1970 (Brown, 1986). Countries within southern Africa where resistance to pyrethroids and permethrin occur are Malawi, Mozambique, South Africa, and Zambia (Hargreaves *et al.*, 2000; Coetzee and Koekemoer, 2013; Knox *et al.*, 2014; Wiebe *et al.*, 2017). There is continued resistance developing to the pyrethroids, carbamates and permethrin in the African *An. funestus* populations (Coetzee and Koekemoer, 2013; Knox *et al.*, 2014; WHO, World Malaria Report, 2020).

Due to the success of DDT during World War II, pyrethrum insecticides were replaced in South Africa with DDT. The use of IRS with DDT eradicated the *An. funestus* species from the region during that period (DeMeillon, 1986; Sharp and Le Sueur, 1996; Coetzee *et al.*, 2013b). Continued pressure to discontinue the use of DDT due to environmental concerns, resulted in the switch to using a pyrethroid, deltamethrin, in 1996 (Govere *et al*., 2000; Coetzee and Fontenille, 2004). At the change of the millennium, there was a rapid increase of malaria incidence in the KwaZulu-Natal region. Both female and male members of *An. funestus* were collected in the northern KwaZulu-Natal region and their progeny were reported as having varying levels of pyrethroid resistance (Hargreaves *et al.*, 2000). This resistance to pyrethroids allowed for the re-establishment of *An. funestus* in South Africa, allowing for the country's worst epidemic in 50 years (Govere *et al*., 2000), in a pre-COVID 19 era. The use of DDT was re-instated in 2000 which resulted in a reported decrease of 91% of malaria incidence in the KwaZulu-Natal region by 2002 (Department of Health, unpublished data, KZN; Maharaj *et al.*, 2005; Coetzee et al., 2013b). *Anopheles funestus* populations from Benin, Cameroon, Ghana, Kenya, Mali, and Nigeria have been shown to have resistance to DDT (Brown, 1986). Due to the increase in insecticide resistance in *An. funestus* and other *Anopheles* species across Africa, alternative vector control strategies need to be developed and implemented.

The development of various novel control strategies such as eave tubes, Attractive Targeted Sugar Bait (ATSB), the use of ivermectin and the Sterile Insect Technique (SIT) are currently under investigation (Oliva *et al*., 2012; Oliva *et al*., 2013; Bouyer and Lefrançois, 2014; Sternberg *et al.*, 2015; Knols *et al.*, 2016; Benelli and Beier, 2017; Waite *et al.*, 2017).

All vector control methods to date have mainly targeted females that exhibit blood seeking behaviour or blood-fed females resting on sprayed walls. Strategies are currently under investigation that do not only target the females and their blood-seeking behaviour. Current approaches to control the male mosquito, include the targeting of male swarms (Diabaté *et al.,* 2004; Ageep *et al.,* 2014; Sawadogo *et al.,* 2017; Kaindoa *et al.,* 2019). These methods include the use of the Sterile Insect Technique, which require the release of large numbers of sterile males (Benedict and Robinson, 2003; Dyck *et al.,* 2005; Hendrichs *et al.,* 2005; Lees *et al.,* 2014) and gene drives (Sinkins and Gould, 2006; Gantz *et al.,* 2015; Hammond *et al.,* 2016; Kyrou *et al.,* 2018). Studies on the *An. funestus* bionomics, mating biology and sexual development of male mosquitoes are, however, limited.

1.1.4 Male mating biology and sexual development

There is limited knowledge about the biology of the *Anopheles* male species, specifically when it comes to their mating biology. The understanding of the anopheline adult male's biology and their physiology is still limited. This knowledge is important to assist in developing additional control strategies that target the mosquitoes' mating and sexual developmental processes. Male anophelines go through several sexual developmental stages after pupation to reach sexual maturation. Some of the physical changes include the rotation of the genitalia by 180°, development of the male accessory glands and the ability for their antennae fibrillae to erect (Clements, 1999). Another important developmental characteristic is cuticle formation which is important for the maturation of the adult male anopheline (Richards, 1958; Mitraka *et al.,* 2013; Moussian, 2013). Once mature, the males are ready to copulate with a female. Virgin females will enter a male swarm and mate with a suitable male. Female anophelines are mated for life and will continue to find hosts to blood feed on, which enable the female to produce eggs until she dies (Roth, 1948; Rees and Onishi, 1951; Clements, 1999; Huho *et al.,* 2006). Swarming is discussed in more detail in Chapter 2, with the various developmental aspects discussed in more detail below.

1.1.4.1 Arthropod cuticles

Arthropod cuticles function as an exoskeleton which serves as a barrier to the environment, maintains body structure and inhibits water from evaporating and can even provide protection for the internal organs and protection against insecticides (Lewis, *et al.,* 1980; Togawa *et al.,* 2008;

Neville, 2012; Yahouédo *et al.,* 2017; Balabanidou *et al.,* 2018). The cuticle is an assembly of polysaccharide chitin and cuticular proteins, which make up the exoskeleton of the anopheline mosquito (Andersen *et al.,* 1995; Willis *et al.,* 2005; Charles, 2010).

The properties of the cuticle are determined largely by the proteins it contains. Depending on the type of cuticle (flexible or stiff), there are specific cuticular proteins that occur at different stages of the insect's development (Willis, 2010; Yang *et al.,* 2017). Cuticular proteins have been categorized into twelve families and are characterized by their conserved amino acid motifs (Willis, 2010 and Yang *et al.,* 2017). The twelve cuticular proteins (CP) families are named: CPRR (named after the Rebers and Riddiford Consensus sequence); CPF (named because it has a highly conserved region consisting of about forty-four amino acids); CPFL (like the CPFs in a conserved C-terminal region); the TWDL family (has a low complexity sequence); CPLCA (named diagnostically); CPLCG (named diagnostically); CPLCW (named diagnostically); CPLCP (named diagnostically); CPG (glycine-rich), the apidermin family (named after three proteins in *Apis mellifera)*; and CPAP1 and CPAP3 (named because they have features similar to peritrophins).

The Cuticular Protein RR (CPRR) family is the largest known family of cuticular proteins in insect species (Yang *et al.,* 2017). The CPRR family is recognized by its Rebers and Riddiford (R and R) consensus sequence, which is a chitin-binding domain (Rebers and Willis, 2001). There are three sub-families of the CPRR family, RR-1, RR-2, and RR-3. The R and R consensus is a conserved region of 35-36 amino acids. It has been shown that the RR-1 proteins predominantly form the soft (flexible) cuticle while the hard (rigid) cuticles mostly consist of the RR-2 proteins (Willis, 2010; Yang *et al.,* 2017). Anderson *et al.* (1995) suggested that the endocuticle is predominantly made up of RR-1 proteins whilst the RR-2 proteins were mainly found in the exocuticle.

The arthrodial membrane is the tough flexible cuticle present between skeletal elements in insects and helps in relative movement (Chevone and Richards, 1976 and Anderson, *et al.* 1995). This membrane (mostly consisting of CPRR-1 proteins) becomes greatly stretched during the rotation of the genitalia (Chevone and Richards, 1976). The mesocuticle (a layer within the procuticle, between the exocuticle and the endocuticle) are made up of different cuticular proteins (mostly CPRR-1 proteins) of the adjacent sclerites which show a realignment of fibres during genitalia

rotation (Rebers and Willis, 2001; Vannini and Willis, 2017). After genitalia rotation, the disrupted regions of cuticle become ultra-structurally reorganized (Chevone and Richards, 1976).

The expression of the proteins in the exoskeleton of the anopheline mosquito is regulated through its genome according to its developmental stage and body regions (Charles, 2010).

1.1.4.2 Male accessory glands

The male accessory glands (MAGs) arise in the 9th abdominal segment of the organism as parts of the imaginal disks during the larval stage. The MAGs are located on either side and anteriorly to the ejaculatory pump and attached to the two seminal vesicles (Figure 1.2). The MAGs enlarge during the pupal stage and continue to enlarge for two days after adult emergence, with the cells starting to fill with secretions (Clements, 1999). The MAGs can be seen as yellow sacs near the terminalia (Figure 1.3, A). The combination of the secretions and spermatozoa forms the semen that is passed from male to female during copulation. Studies have shown that MAG development is related to the sexual maturity of the *An. gambiae* male mosquito (Huho *et al.*, 2006). Sexually immature male mosquitoes tend to have a clear area in the MAG and small sperm reservoirs, whereas a non-transparent area on the edge of the MAG and large sperm reservoir show that the male is sexually mature, as seen in Figure 1.3, B and C. Sexually immature individuals were found to have shorter testes compared to sexually mature individuals which had longer testes (Figure 1.3, D, E and F). The testes lengthen with the aging of the mosquito but start to shorten again when the mosquito becomes too old to be a sexually viable mate (Mahmood and Reisen, 1982).

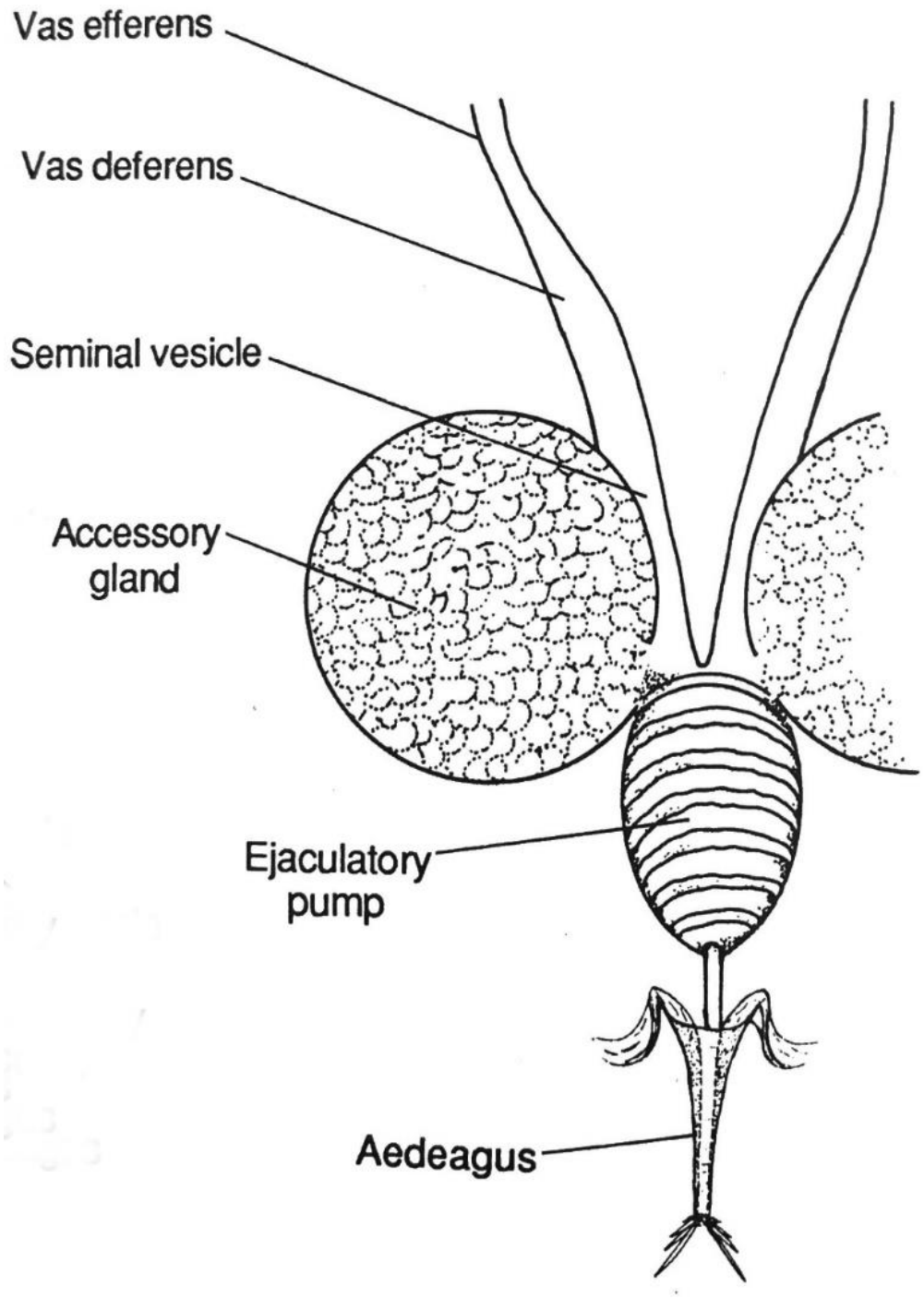


Figure 1.2 Diagram of male accessory glands

Posterior region of the male gonoducts and associated organs of *Anopheles.* The diagram shows the vas efferens, vas deferens, seminal vesicles, the male accessory glands, the ejaculatory pump and the aedeagus (figure taken directly from Clements, 1999).

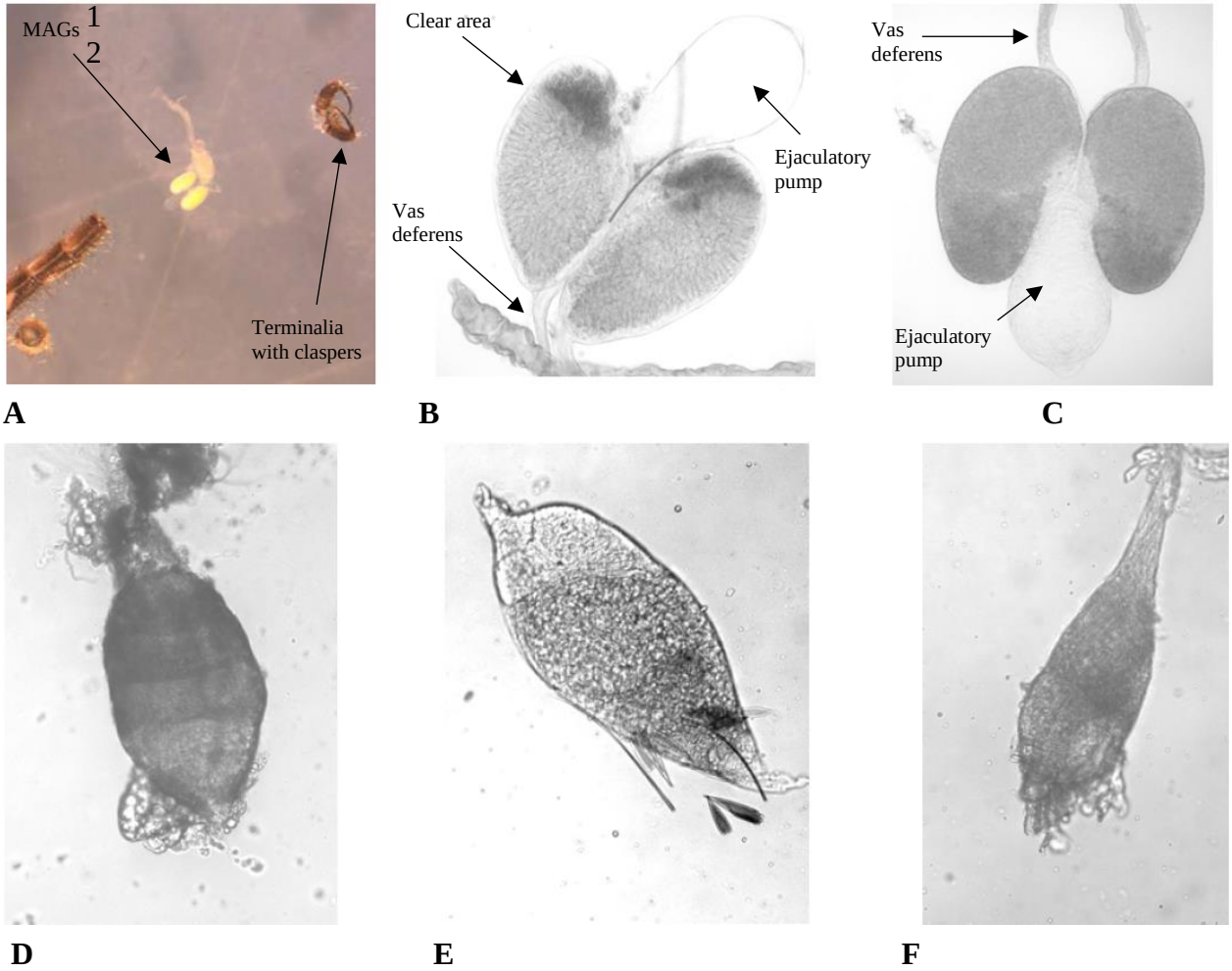


Figure 1.3 Example of male accessory glands

(A) Dissected male anopheline mosquito, the MAGs are visible as the yellow sacs near the terminalia with the claspers seen in the top right corner. (B) Sexually immature male accessory glands, with a clearly defined clear area or transparent area around the edge of the gland, showing that the male is sexually immature. The ejaculatory pump can be seen in between the male accessory glands, the vas deferens can also be seen attached to the two accessory organs. (C) Sexually mature male accessory glands, without a clear band or transparent area around the edge of the gland, ejaculatory pump, and vas deferens can also be seen. (D) Testis of a 1-day old *Anopheles gambiae* mosquito, (E) Testis of a five-day old *An. gambiae* mosquito, (F) Testis of 7-day old *An. gambiae* mosquito (figure taken directly from Huho *et al.*, 2006).

1.1.4.3 Antennal fibrillae erection

The antennae have evolved to become organs of hearing and smelling that allow for the male to detect sounds emanating from conspecific females that are ready for mating within swarms (Roth, 1948; Rees and Onishi, 1951; Nielsen, 1964; Dingle, 1972; Foster and Lea, 1975; Nijhout, 1977; Charlwood and Jones, 1979; Klowden and Beach, 1980; Clements, 1999; Anderson *et al.*, 2000 and Da Silva *et al.*, 2009). The *Anopheles* male antennae consist of a flagellum which has many hair-like structures (antennal fibrillae) attached. Within a day or two after emergence, the male anopheline mosquito has the capability of erecting the fibrillae to an almost perpendicular

degree to the flagellum, as can be seen in Figure 1.4. The fibrillae generally stay recumbent except during swarming activities (Roth, 1948; Nielsen, 1964; Dingle, 1972; Foster and Lea, 1975; Nijhout, 1977; Charlwood and Jones, 1979; Klowden and Beach, 1980; Anderson *et al.*, 2000).

The antennal flagella are also used for olfaction and comprise multiparous hair and peg sensilla which detect volatile compounds and detect nectar sources for both male and female *An. funestus* mosquitoes (Davis, 1976). In a recent study, it was shown that the sexually mature male *Anopheles* release aggregation pheromones which mediate the formation and maintenance of mosquito swarms (Mozūraitis *et al.*, 2020). This study also showed that there was an increase in mating success when the aggregation pheromones were released artificially, indicating that the female anopheline was attracted to the pheromones through the olfactory pathway (Mozūraitis *et al.*, 2020). It was found that these pheromones were released during swarming activity which is the same period as when the antennal fibrillae are erect.

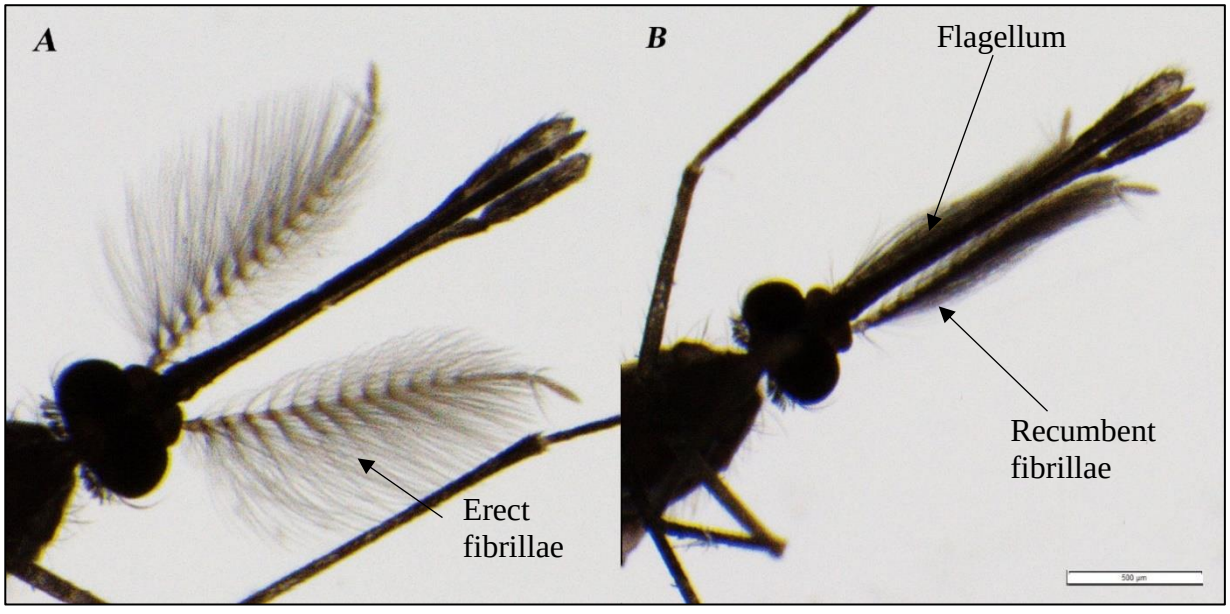


Figure 1.4 Example of antennal male fibrillae erection

Image of *An. funestus* male antennae. (A) antennal fibrillae that are erect, where fibrillae are almost perpendicular to the flagellum, which occurs during the swarming period (sunset). (B) Recumbent antennal fibrillae. (Images by Jacek Zawada).

1.1.4.4 Genitalia rotation

The *Culicidae* male genitalia or terminalia consists of the abdominal sections 8 through 10, the 10th section being the claspers which are attached to the 9th segment, as seen in Figure 1.5 (Clements, 1999). Sections 1 to 7 make up the mosquito abdomen. Genitalia rotation is the hallmark event for the sexual maturation of the *Culicidae* male (Roth, 1948; Rees and Onishi, 1951; Provost *et al.,* 1961; Akey and Jones, 1968; Chevone and Richards, 1976; O'Donnell and Klowden, 1997 and Dahan and Koekemoer, 2014). After emergence, a 135° - 180° rotation of the genitalia is required by the 8th abdominal segment so that the claspers are orientated downwards (Figure 1.5), in order for the male anopheline to sexually copulate. This allows the male to grasp the female during flight in order to copulate (Rees and Onishi, 1951, Clements, 1992; Dahan and Koekemoer, 2014). This was shown with earlier work by Roth (1948) where the 7th and the 8th posterior segments were fused, with wax in *Aedes aegypti* males which prevented rotation and subsequent insemination (Roth, 1948).

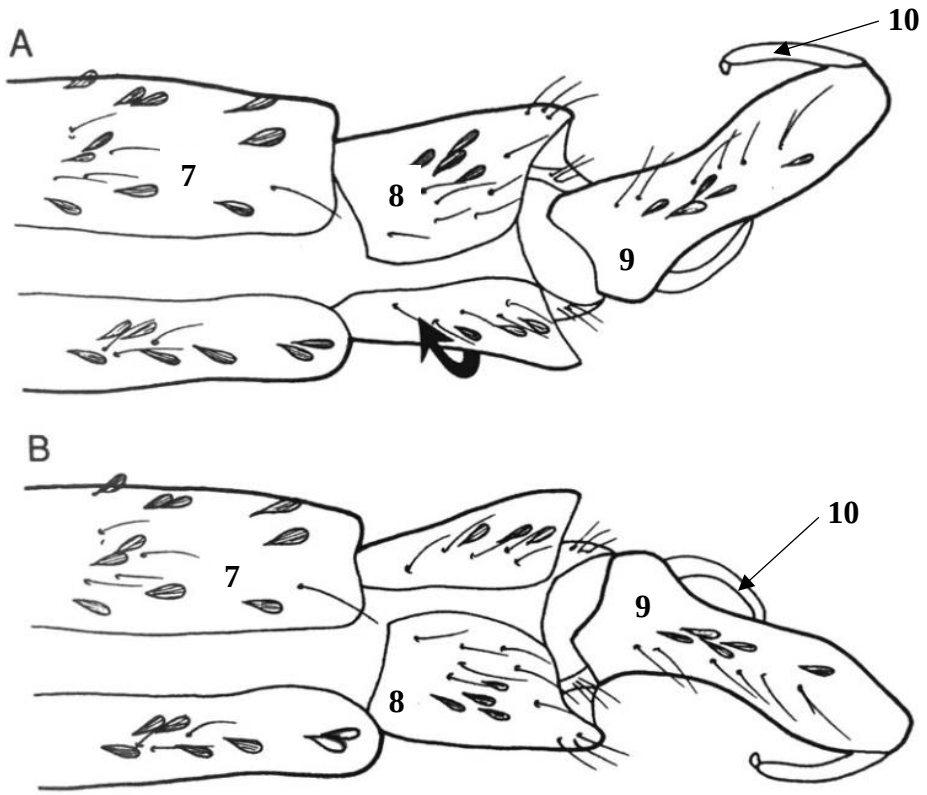


Figure 1.5 Inversion/genitalia rotation of the male anopheline terminalia

(A) The orientation of the 8th posterior abdominal segment at emergence; (B) at 24h post-emergence showing how the claspers tipped with claws are orientated downwards allowing for the male to grasp the female during flight. The numbers on the figure indicate the abdominal segments, abdominal segments 8 through 10 are considered the genitalia and rotate through to 135° – 180 °, whilst segment 7 remains stationary (figure directly taken from Clements,1992).

Genital rotation involves two pairs of opposed abdominal muscles. The first pair of muscles is situated ventrally between the 7th and 8th sterna (Figure 1.6) while the second pair is located dorsally between the 7th and 8th terga (Chevone and Richards, 1976). During the rotation event, one muscle between the 7th and 8th sterna contracts whilst the other relaxes. The 7th and 8th terga muscle can be seen in Figure 1.6 (Chevone and Richards, 1976; Clements, 1992). This rotation can be either clockwise or counter-clockwise (Clements, 1992).

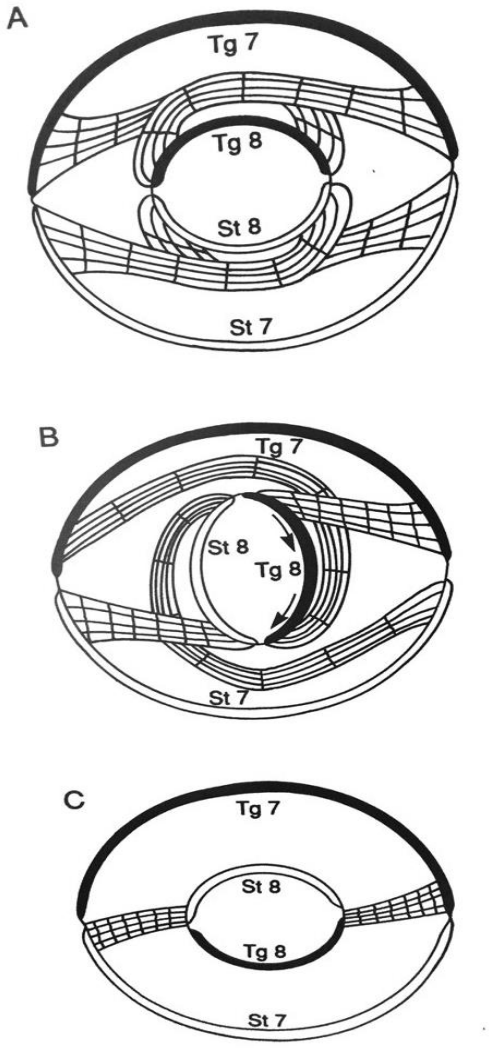


Figure 1.6 Disposition of the two pairs of rotational muscles of an adult male *Aedes aegypti*

Before (A), during (B) and after (C) inversion of the terminal abdominal segments. Thick transverse sections through the posterior region of the 7th abdominal segment, into which the 8th segment is telescoped. (A) In pharate adult. (B) At 90° rotation. (C) Twenty-four hours after completion of inversion. St, sternum; Tg, tergum. Arrows, directions of rotation; numbers indicate abdominal segment number (figure directly taken from Clements, 1992).

Dahan and Koekemoer (2014), showed that between 14 to 36 h post-emergence at 23 ±1°C, full rotation in *An. funestus* males is reached. Furthermore, laboratory colonized, and wild *An. funestus* males had a similar rate of genitalia rotation, reared under controlled conditions (Dahan and Koekemoer, 2014). The rate of genitalia rotation is influenced by temperature (Provost *et al.*, 1961; Oliva *et al.*, 2011; Dahan and Koekemoer, 2014) and this might have implications for the debate on global warming. The hypothesis is that as the global temperature increases, young male mosquitoes will sexually mature earlier with mating happening earlier which will indirectly result in the population growth of *An. funestus* (Dahan and Koekemoer, 2014).

Dahan and Koekemoer (2014) described the five stages of genitalia rotation of *An. funestus*. Figure 1.7 shows the rotation through these 5 stages from 0° - 180° rotation. The morphological changes are easily visualised, however, the molecular mechanisms or pathways involved in this complex biological process need to be elucidated. Functional genomics will provide valuable insight into this process.

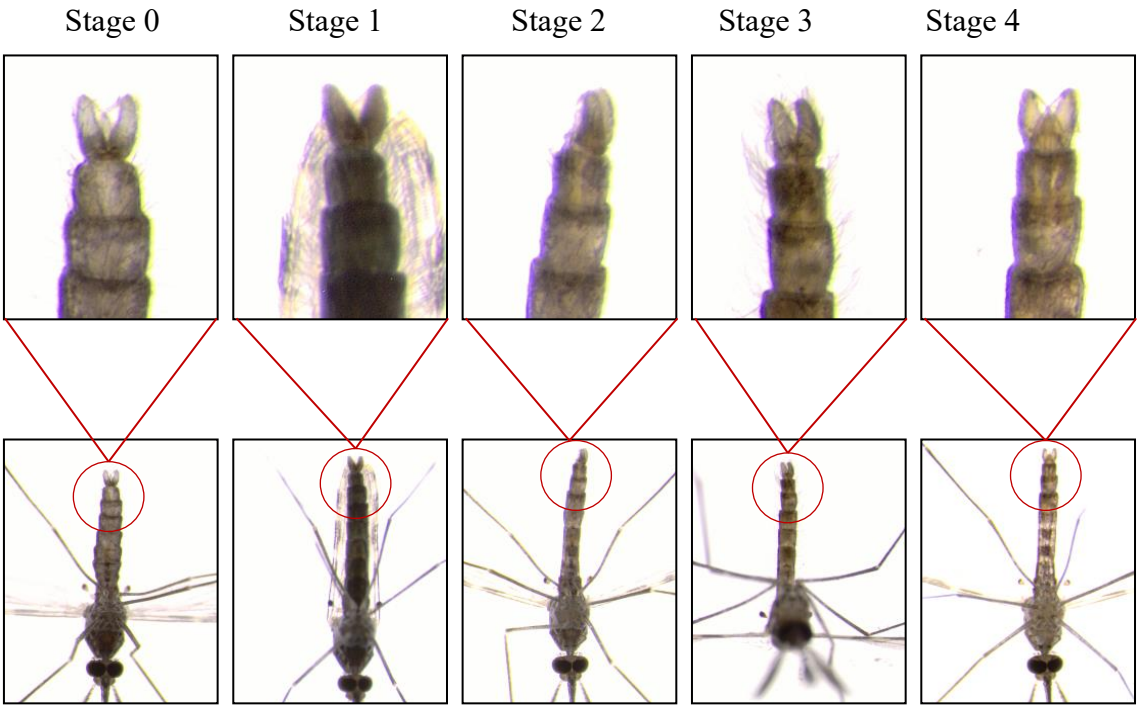


Figure 1.7 The different stages of male genitalia rotation in *An. funestus*

Stage 0: showing 0° - 44° clasper rotation; Stage 1: showing 45° - 89° clasper rotation; Stage 2: showing 90° - 134° clasper rotation; Stage 3: showing 135° - 179° clasper rotation; Stage 4: showing 180° clasper rotation (figure taken directly from Dahan and Koekemoer, 2014).

1.1.5 Functional genomics

Functional genomics is undertaken to understand the functions of genes or proteins. In 1957, Francis and Crick stated that there are three main transfers of genetic information flow, DNA (deoxyribonucleic acid) can be copied to DNA (DNA replication), DNA can be copied into messenger ribonucleic acid or mRNA (transcription), and the mRNA can be used as a template to synthesize proteins (translation) (Cobb, 2017). There are other special transfers of genetic information flow, such as RNA replication and reverse transcription.

There are a number of tools used for the study of functional genomics. Some of the tools used to study functional genomics are microarrays, Next-Generation Sequencing, quantitative real-time PCR and RNA interference. Differential gene expression techniques, such as microarrays, Next Generation Sequencing, quantitative real-time PCR will be discussed further in Chapter 3 and 4. Loss of function techniques, such as RNA interference will be discussed further in Chapter 5. Using these tools will allow us to study *An. funestus* functional genomics, by comparing the expression of the transcripts and genes of the *An. funestus* male at one developmental time point and comparing the expression levels of those same transcripts at different developmental time points, such as in RNA-sequencing. This would give a suggestion as to which transcripts and genes are putatively involved during the sexual development of the *An. funestus* male. These differentially expressed genes should be verified by another quantitative method such as quantitative real-time PCR. This could be followed up with a loss of function technique such as RNA interference to target an expression of specific gene and downregulate its expression. This can then be used to examine the phenotypic changes that occur with the loss of the function of this particular transcript or gene, which could, in turn, indicate the function or the purpose of that transcript or gene.

In order for a phenotype to be changed, the organism's genotype will have to transcribe the mRNA for that phenotype after which the mRNA gets translated to that phenotype (protein), giving rise to the organism having that phenotype. Measuring the quantity of mRNA of a specific gene (transcript) that is associated with a particular phenotype, can give evidence of the expression level of that gene and its potential association with that particular phenotype. This becomes valuable for this study in that the expression levels between two developmental time points, such as sexually mature versus sexually immature *An. funestus* males, can be compared.

Transcriptomic studies can provide valuable information such as the identification of potential gene regulation elements and pathways, the identification of unknown genes, the functions of genes, and identifying genes that are responsible for particular phenotypes. These approaches are discussed in more detail in Chapter 3 and 4. The abundance of a specific transcript(s) in an organism at a particular time could indicate that those transcript(s) may have a direct or indirect biological function during the particular time being evaluated. In order to study the biological function of a transcript(s) RNA interference can be used. This is discussed in more detail in Chapter 5.

1.2 Research aims and hypotheses

This study is based on two main hypotheses, namely

- that immature *Anopheles funestus* males that have not undergone genitalia rotation will not be able to mate and will therefore not participate in swarming activities, and
- that a gene expression change occurs in the male *An. funestus* which leads to the rotation of the genitalia.

Several working hypotheses arose during the study to support the investigation. These working hypotheses are discussed in the relevant sections.

The aims of this study are therefore firstly to determine whether immature *An. funestus* males participate in swarming activities, and secondly to discover transcripts that are putatively involved in the sexual maturation of the male *An. funestus* mosquito.

The study will help increase fundamental knowledge of male *An. funestus* biology, specifically the mechanisms involved in the process of genitalia rotation that can help further our understanding of this important malaria vector.

1.3 Objectives

The aims of this study can be separated into four main objectives:

- Molecular species identification and physiological characterization of wild *Anopheles funestus* swarms in Zambia.
- Transcriptomic studies of male *An. funestus* mosquitoes during sexual maturation.
- Confirmation of transcript abundance of selected transcripts in male *An. funestus* mosquitoes.
- Investigation of the potential biological function of selected transcripts in male *An. funestus* mosquitoes.

2 Molecular species identification and physiological characterization of wild *Anopheles funestus* swarms in Zambia

2.1 Introduction

2.1.1 Anopheline swarming

Anopheline species, including *Anopheles funestus*, mate by creating swarms. The beginning of a swarm occurs when one or more males start flying simultaneously and the swarm numbers rapidly increase, reaching a maximum size and density within five min (Charlwood *et al.*, 2003; Diabaté *et al.*, 2009). During this time, a female will enter the swarm and will subsequently mate with one of the swarming males. By observing the swarming behaviour, it is clear that the females choose a specific mate based on various indicators, such as, wing beat frequencies, male pheromones, and wing size (Ng'habi *et al.*, 2008; Howell and Knols, 2009). The harmonization of wing beat frequencies between males and females, appears to play a role in in mating success, amongst other indicators (Pennetier *et al.*, 2010). It was also demonstrated by Mozūraitis *et al.* (2020), that male pheromones play a large role in the attraction and mating success of anophelines.

Swarming of anophelines occurs at sunset and sunrise in relatively open areas (Charlwood *et al.*, 2003). Swarms also occur in the same location where they were first observed, indicating that swarms reform themselves each evening (Charlwood *et al.*, 2003). Swarming behaviour appears to be species-specific in anophelines, with studies to date mainly focussed on the malaria vector species, such as *An. gambiae s.s., An. coluzzii and An. arabiensis* (Harper, 1944; Marchand, 1983; Charlwood *et al.*, 2003; Diabaté *et al.*, 2009; Diabaté *et al.*, 2011; Sawadago *et al.*, 2013).

Anopheles gambiae swarm formations have been quite extensively studied (Russell and Rao, 1942; Marchand, 1983; Yuval and Bouskila, 1993; Charlwood *et al.*, 2003; Diabaté *et al.*, 2004; Diabaté *et al.*, 2006; Diabaté *et al.*, 2009; Diabaté *et al.*, 2011; Manoukis *et al.*, 2009). Diabaté *et*

al. (2006) showed that mixed mating of *An. coluzzii* and *An. gambiae s.s.* occur within swarming males. It was also concluded that *An. gambiae s.s.* prefer to swarm over horizontal contrasting markers on the ground, but there is no evidence to support that the same behaviour occurs in the *An. funestus* species (Charlwood and Jones, 1979; Charlwood *et al.*, 2003).

Harper (1944) first observed *An. funestus* swarming in East Africa within the entrance of huts and a foot (0.3m) above the ground. In contrast, *An. funestus* from Mozambique swarm well away from houses and 2 - 4m above the ground (Charlwood *et al.*, 2003). The differences in swarm height observed between these two studies might be explained by two different species being observed, but molecular species identification was not available in 1944 (Harper, 1944). Charlwood *et al.* (2003) assumed that the *An. funestus* group adults that were collected were, most likely, *An. funestus s.s.* based on past studies in Mozambique and molecular species identification was not confirmed in this study. As mentioned in Section 1.1.3, morphological identification is very difficult due to the species being morphologically similar and species can only be differentiated through molecular techniques.

There were very few studies on *An. funestus s.s.* swarming (Harper, 1944; Charlwood *et al.*, 2003) when the current study was initiated. The level of sexual maturity in the swarming males was also unknown. The physical change of the rotation of the male genitalia can be used as a proxy for identifying sexually matured males.

It is essential for genitalia rotation to take place for mating to be successful (Clements, 1999). Swarming uses 50% of all available calories according to a study by Yuval *et al.* (1994). It was hypothesized that males that have not had their genitalia rotated would preserve their energy by not participating in swarming activities until they are physically able to find a mate (Yuval *et al.*, 1994). Their energy reserves might be better spent on achieving sexual maturity by rotating their genitalia (Yuval *et al.*, 1994). Although several studies have reported on swarming in various Dipterans, at the time of this study, no published literature confirmed that immature *An. funestus* males do not participate in swarming.

2.2 Rationale

In order to improve our understanding of wild *Anopheles funestus* swarms, molecular characterization (species identity), as well as the stage of male genitalia rotation of this species need to be confirmed and recorded. In addition to this, the population substructure within the wild *An. funestus* swarms needs to be investigated, using molecular markers. Due to the abundance of wild *An. funestus* populations in Zambia, these populations were considered ideal for this study.

My aim was to characterize *An. funestus* swarms at the molecular level with regards to species composition, genitalia rotation and population substructure in Zambia.

2.3 Material and methods

2.3.1 Biological material and swarm collection

The preservation of samples for transport was optimized at the National Institute for Communicable Diseases (NICD), prior to field work. *Anopheles funestus* males, from the FUMOZ colony which originated from Mozambique, that were reared in the Botha De Meillon insectary at the NICD, National Health Laboratory Services (NHLS), Sandringham, Johannesburg, South Africa, were collected. The males were analyzed, and their stage of genitalia rotation was recorded, and they were then placed into silica, 70% ethanol and RNAlater, as separate preservation mediums, for 72 h. The samples had their genitalia rotation re-analyzed after 72 h which confirmed that the preservation methods did not affect the genitalia rotation of the males and that post-collection analysis of samples of genitalia rotation and molecular assays could still be performed on the field samples.

The wild biological material was collected every day from 18th – 25th of May 2016, in Nchelenge, northern Zambia (S09°19’28.6” E028°47’06.9”). The swarms were caught using sweep nets during the dusk period (18:00-18:30pm), which is the best time to capture a maximum number of adult anophelines (Charlwood *et al.,* 2003; Dabiré *et al.,* 2014). The swarms were caught approximately 50 – 100m from the Kenan stream, near a dwelling, in a village (Figure 2.1). Most of the dwellings in the village were mud huts with thatched roofing. The capture of these swarms was done without any artificial markers, whereby a contrasting artificial marker is placed on the ground in order to stimulate swarm formation (Charlwood *et al.,* 2003). Capture of the swarms took place on six consecutive nights and were labeled Swarm #1 - 6, with Swarm #6 captured on the final day. Swarms were preserved in 70% ethanol.

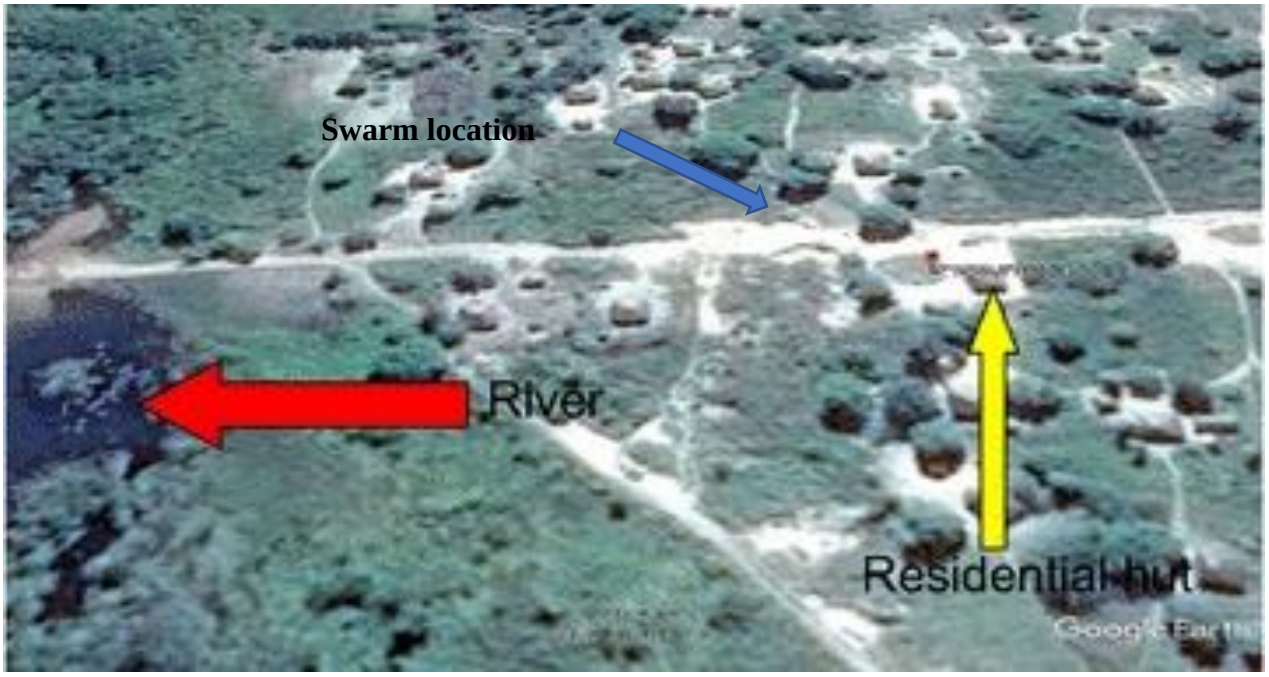


Figure 2.1 Swarming location
The image shows the location of the swarming site. https://earth.google.com/web/search/S09%C2%B019%E2%80%9928.6%E2%80%9D+E028%C2%B047%E2%80%9906.9%E2%80%9D [Accessed 27 January 2021]. Red arrow shows the river (which could be a putative breeding ground), yellow arrow shows the nearest residential hut from where the swarm was located (5- 10m away from the swarming site), the blue arrow shows the location of the swarming site.

2.3.2 Species identification

The dichotomous keys were used to morphologically identify all of the samples collected, into either the *An. funestus* group or the *An. gambiae* complex prior to molecular assays (Gillies and Coetzee, 1987; Coetzee, 2020).

2.3.2.1 Anopheles funestus group PCR identification

The mosquitoes collected from the swarms had their DNA extracted from a leg or wing by using the DNA extraction kit prepGEM®Insect (CAT No.: PIN0050). Standard multiplex PCR was used to identify members of the *An. funestus* group to species level (Koekemoer *et al.*, 2002; Cohuet *et al.*, 2003). This PCR was performed using a 12.5 µL reaction mixture containing: 1.25 µL of 10× PCR buffer (500 mM KCl, 100 mM Tris-HCL, pH 8.33), 1.5 mM $MgCl_2$, 3.3 pmol/primer of each primer, 200 µM of each dNTP, and 0.5 units taq DNA polymerase. This multiplex PCR uses unique primers that target the Inter Transcriber Spacer 2 (ITS2) region of the six species. The sequences of the primers are presented in Table 2.1. Positive controls of the different species, namely, *An. funestus s.s.*, *An. vaneedeni*, *An. rivulorum*, *An. parensis*, and *An.*

leesoni were included. The amount of template that was added from each sample as well as the positive controls to each reaction was 0.5 µL of the extracted DNA (5-25ng). A "no template" control, where the template DNA was replaced with nuclease-free water, as well as the negative DNA extraction control (DNA extraction mix with 0.5 µL of nuclease free-water), was included. The cycling conditions of the PCR were set up as follows: initial denaturation @ 94°C for 2 min, followed by 30 cycles of 94°C @ 30 s; 45°C @ 30 s, and 72°C @ 40 s. A final extension was included at 72°C @ 5 min. The PCR products were then electrophoresed on a 2.5% Tris-Acetate-EDTA agarose gel stained with 0.5 µg/mL of ethidium bromide. Specimens that were damaged to the extent that they could not be morphologically identified, were also analyzed via *An. funestus* multiplex PCR to examine the possibility that they were part of the *An. funestus* group.

Table 2.1 Table of primers used in the *An. funestus* multiplex PCR

Primer	**Sequence**	**Expected fragment size (bp)**	**Tm (°C)**
UV (Universal)	TGT GAA CTG CAG GAC ACA T	N/A	55.3
FUN (*An. funestus*)	GCA TCG ATG GGT TAA TCA TG	505	52.4
VAN (*An. vaneedeni*)	TGT CGA CTT GGT AGC CGA AC	587	58.0
PAR (*An. parensis*)	TGC GGT CCC AAG CTA GGT TC	252	60.5
LEES (*An. leesoni*)	TAC ACG GGC GCC ATG TAG TT	146	60.2
RIV (*An. rivulorum*)	CAA GCC GTT CGA CCC TGA TT	411	58.8
RIV-LIKE (*An. rivulorum-like)*	CCG CCT CCC GTG GAG TGG GGG	313	60.7

Primer information taken from Koekemoer *et al.*, 2002; Cohuet *et al.*, 2003; Garros *et al*., 2004.

2.3.2.2 Anopheles gambiae complex PCR identification

Samples from the *An. funestus* swarm that were too damaged and could therefore not be identified morphologically, were analyzed using the *An. gambiae* multiplex PCR to examine the possibility that they were morphologically part of the *An. gambiae* complex (Scott *et al.*, 1993). The DNA was extracted from a leg or wing as with the *An. funestus* group identification PCR in section 2.3.2.1. This PCR was performed in a 12.5 µL reaction mixture containing: 1.25 µL of 10× PCR buffer (500 mM KCl, 100 mM Tris-HCL, pH 8.33), 25 mM $MgCl_2$, 3.3 pmol/primer of each primer, 2.5 mM of each dNTP, and 5 U/µL taq DNA polymerase. This multiplex PCR uses unique primers that target the intergenic spacer (IGS) regions of the ribosomal DNA (rDNA) of the four species. The sequences of the primers are presented in Table 2.2. Positive controls of the different species, namely, *An. arabiensis, An. gambiae, An. merus,* and *An. quadriannulatus* were included. The amount of template that was added from each sample as well as the positive controls to each reaction was 0.5 µL DNA (5-25ng). A "no template" control, where the template DNA was replaced with nuclease-free water was included, as well as the negative DNA extraction control. The cycling conditions of the PCR were set up as follows: initial denaturation @ 94°C for 2 m, followed by 30 cycles of 94°C @ 30 s; 50°C @ 30 s, and 72°C @ 30 s. A final extension was included at 72°C @ min. The PCR products were electrophoresed on a 2.5% Tris-Acetate-EDTA agarose gel stained with 0.5 µg/mL of ethidium bromide (Sigma-Aldrich).

Table 2.2 Table of primers used in the *An. gambiae* multiplex PCR

Primer	Sequence	Expected fragment size (bp)	Tm (°C)
(UN) Universal	GTG TGC CCC TTC CTC GAT GT	N/A	58.3
(G) "*An. gambiae"*	CTG GTT TGG TCG GCA CGT TT	390	59.3
(M) *An. merus*	TGA CCA ACC CAC TCC CTT GA	464	57.2
(A) *An. arabiensis*	AAG TGT CCT TCT CCA TCC TA	315	47.4
(Q) *An. quadriannulatus*	CAG ACC AAG ATG GTT AGT AT	153	47.7

Primer information taken from Scott *et al*., 1993.

2.3.2.3 Anopheles funestus genotyping

A Restriction Fragment Length Polymorphism (RFLP) assay was used for the identification of molecular type on the *An. funestus s.s.* samples (M-, W-, or MW-type) (Koekemoer *et al.*,1998; Garros *et al.*, 2004). The intragenomic variation within the *An. funestus* populations can be seen from the *Hpa*II digestion patterns for Domain 3 (D3) of the 28S ribosomal DNA. This assay was performed by using PCR to amplify a fragment of the variable D3 domain in the 28S gene, the PCR product was digested with the *Hpa*II enzyme. The resulting product produces specific sized fragments, which corresponds to either the M-form, W-form, or the MW-form of *An. funestus s.s.*

This D3 PCR was performed on samples identified as *An. funestus s.s.* in a 25 µL reaction mixture containing: 2.5 µL of 10× PCR buffer (500 mM KCl, 100 mM Tris-HCL, pH 8.33), 1.5 µL of 25 mM $MgCl_2$, 3.3 pmol/primer of the forward and reverse of the D3 primer, 2.5 µL of 2.5 mM of each dNTP, and 5 U/µL taq DNA polymerase. From the extracted DNA 0.5 µL of template is used in the PCR mixture. The sequences of the primers are presented in table 2.3. Positive controls of the M-, W- and MW-type species were included. A "no template" control, where the template DNA was replaced with nuclease-free water was included, as well as the negative DNA extraction control. The cycling conditions of the PCR were set up as follows: 30 cycles of 94°C @ 30 s; 49°C @ 30 s, and 72°C @ 30 s. A final extension was included at 72°C @ 10 min. Subsequently, 10 µL of the PCR product was electrophoresed on a 2.5% Tris-Acetate-EDTA agarose gel stained with 0.5 µg/mL of ethidium bromide in order to confirm that the D3 region amplicon fragment to make sure that the D3 region fragment (400 bp) was present. The rest of the PCR product (15 µL) was then digested in a 20 µL reaction mixture containing 1.5 µL *Hpa*II enzyme with 2.2 µL of 10× Buffer A (60 mM Tris-HCL, 60 mM $MgCl_2$, 60 mM NaCl, 10 mM DTT, pH 7.5) and 1.3 µL of nuclease-free water. The *Hpa*II enzyme recognizes the four-base sequence CCGG and cleaves it, resulting in uniquely sized fragments that correspond to the molecular form of the *An. funestus* species (Figure 2.2). An example of the expected banding pattern can be seen in Figure 2.2, which was taken from Garros *et al.* (2004). The mixture was digested at 37°C for two hours and the *HpaII* digested PCR product was electrophoresed on a 3% TBE gel containing 4 µL/400 ml of ethidium bromide.

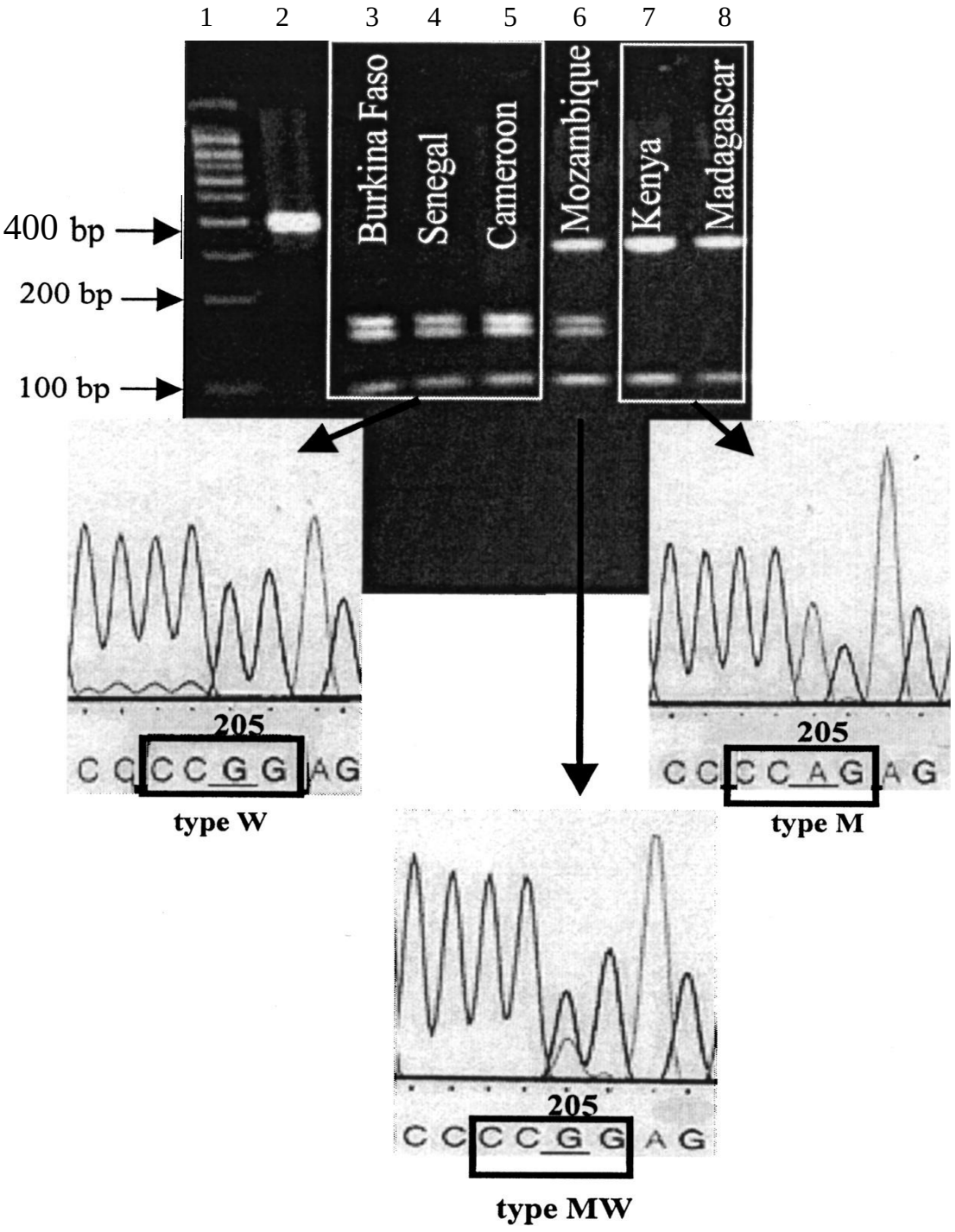


Figure 2.2 Restriction fragment length polymorphism banding pattern

The restriction endonuclease sites are represented in the boxes. Lane 1: Molecular Weight Marker; Lane 2: Undigested fragment (negative control for endonuclease and positive control for D3 amplification); Lane 3 – 5: W-type species from Burkina Faso, Senegal and Cameroon, respectively; Lane 6: MW- type species from Mozambique; Lane 7 and 8: M- type species from Kenya and Madagascar. The cloning and sequencing of the D3 fragment show the nucleotide mutation at position 205 (underlined). bp = base pairs. The presence of both M- and W-type species, gives rise to 4 amplicons on the gel. And the double peak at 205bp on the MW-type chromatogram indicates the presence of both M- and W-types being present (taken directly from Garros *et al.*, 2004).

Table 2.3 Table of primers used in the D3 amplification for RFLP of the *An. funestus* species

Primer	Sequence	Expected fragment size (bp)	Tm (°C)
D3a (Forward)	GACCCGTCTTGAAACACGGA	400	61.8°C
D3b (Reverse)	TCGGAAGGAACCAGCTACTA		59.4°C

Primer information taken from Garros *et al.* 2014.

2.3.2.4 Differentiation between An. gambiae s.s. and An. coluzzii using RFLP identification assay.

Restriction Fragment Length Polymorphism (RFLP) analysis using *Hha*I FastDigest (Thermo Scientific FastDigest, cat no: (#FD1854) was performed on the *An. funestus* swarm samples that were identified as "*An. gambiae"* to differentiate between *An. gambiae s.s.* and *An. coluzzii* (Favia *et al.,* 1997; Coetzee *et al.,* 2013). This PCR-RFLP based method identifies the different species based on the presence of two nucleotide substitutions (G→C in the S-form (*An. gambiae s.s.*) and A→T in the M-form (*An. coluzzii*)). The *An. gambiae* complex PCR was performed using the samples that were morphologically identified as *An. gambiae* group and the PCR products were digested with 2 µL of 10X FastDigest Green Buffer, 1 µL of FastDigest *Hha*I enzyme and 17 µL of nuclease-free water. A negative control "no DNA" template as well as *An. gambiae s.s.* and *An. coluzzii* positive controls were included. The digestion was performed at 37°C for 10 min. Subsequently, the *HhaI* digested PCR products were electrophoresed on a 2.5% TAE agarose gel containing 0.5 µg/mL of ethidium bromide, which generated a species-specific banding pattern on the gel. If the banding pattern consisted of one fragment at the 367 bp mark, the sample was identified as *An. gambiae s.s.* and if there are two fragments at the 257 bp and the 110 bp mark, it was identified as *An. coluzzii*.

2.3.3 Clade identification

The clade characterization of *An. funestus* species was performed according to the protocol described by Choi *et al.* (2014). The technique uses the Applied Biosystems™ TaqMan® real-time PCR assay by detecting the single nucleotide polymorphism (SNP) of clades I and II in *An. funestus s.s.* The real-time hydrolysis probes were designed to amplify the *An. funestus* s.s.

cytochrome c oxidase subunit 1 (COI) to detect either clades I or II in the *An. funestus* sample (Choi *et al.*, 2013). The primers and probes with their respective fluorescent tag are presented in Table 2.4. Each probe carried a 3′ nonfluorescent quencher and a minor grove binding (MGB) protein. The assay was performed with three positive clade I and three positive clade II controls with a “no DNA" negative control, where the template DNA was replaced with nuclease-free water. The 20 µL PCR reaction was set up in a 96 well qPCR plate. Each well contained 10 µL of IQSupermix, 1.6 µL of 10 µM forward primer, 1.6 µL of 10 µM reverse primer, 0.8 µL of 10 µM of clade I probe, 0.8 µL of 10 µM of clade II probe, 4.2 µL of nuclease-free water and 1 µL of template DNA. The cycling conditions were as follows: 95°C for 10 min, followed by 45 cycles of 10 secs @ 95°C, 45 secs @ 63°C, plate read. The 2′-chloro-7′phenyl-1,4-dichloro-6-carboxy-fluorescein (VIC) fluorescent tag fluoresces yellow with excitation at the 530 nm wavelength and emission is at the 555nm wavelength and tags the clade II probe. The 6-Carboxyfluorescein (FAM) fluorescent tag fluoresces green with excitation at the 492 nm wavelength and emission is at the 517 nm wavelength and tags the clade I probe. A substantial increase in 6-FAM fluorescence indicates the presence of clade I whilst a substantial increase in VIC fluorescence indicates the presence of clade II.

Table 2.4 Table of primers and probes used in the *An. funestus* clade characterization assay

Primer	**Sequence (5′ - 3 ′)**	**Fluorescent Tag**
Forward	GCA GGA ACA GGA TGA ACA GT	N/A
Reverse	GAA ATT CCT GCT AAA TGT AAT GAA A	N/A
Probe	Sequence (5′ - 3 ′)	Minor Groove Binding Fluorescent Tag
Clade I	TCA GGA ATT GCT CAT GCT	6-FAM
Clade II	TCA GGA ATT GCC CAT GCT	VIC

Primers taken from Choi *et al.* 2013.

2.3.4 Genitalia rotation observations

Each male mosquito was placed ventrally on a clean slide and viewed under a stereoscope (50x magnification). The stage of each individual male mosquito was recorded according to the criteria described in Dahan and Koekemoer (2014) (described in Figure 1.7 (Section 1.1.4.4)).

2.3.5 Spermathecae dissection

To determine the insemination rate of the *An. funestus* females collected in the swarms, all of the females had their spermathecae dissected and checked for the presence of sperm in the spermathecae. While viewing the female on her ventral side, under a dissecting microscope, the spermathecae can be seen through the cuticle near the terminalia inside segment 8. The spermatheca has a golf-ball like appearance, it is a spherical, fenestrated dark-brown organ. The spermathecae can be exposed along with the malpighian tubules, ovaries, and digestive tract by gently removing the terminalia and segment 9 in phosphate buffered saline (PBS). The spermathecae was then removed and viewed under 40x magnification. An example of the insemination status of the spermathecae is seen in Figure 2.3 (Biodefense and Emerging Infections Research Resources Repository). Spermatheca were dissected and scored as either mated or unmated as illustrated in Figure 2.3



Figure 2.3 Spermathecae dissection

(A) Image indicating the spermathecae under the dissection microscope; (B) a fully inseminated spermathecae full of sperm which can be seen within the spermathecae; (C) an image of an un-inseminated and empty spermathecae. Taken from Biodefense and Emerging Infections Research Resources Repository (taken directly from beiRESOURCES [Accessed 26 January 2021).

2.4 **Results**

2.4.1 Biological material preservation and swarm collection

Biological material preservation:

The different preservation methods of *An. funestus* males and the effects on males' genitalia rotation was evaluated prior to collecting field males from the wild *An. funestus* swarm. Stage 3 *An. funestus* males from laboratory colonized samples were stored in three different preservation media (silica, 70% ethanol and RNAlater) for 72 h and their genitalia rotation was observed. Results showed that regardless of the preservation method, the genitalia rotation of the males (before and after preservation) were the same (n=30). It was, therefore, decided to preserve wild field samples in 70% ethanol, due to ease of use and cost benefits.

Swarm collection:

A total of six swarms were collected from a single swarming site. The swarms occurred 2 – 4 m above the ground and were 5 to 10 min from a residential hut (Figure 2.4). The swarms reformed at the same location each night and swarmed during the same period. The earliest swarming activity that was observed was at 18:05 and the latest was at 18:27. After 18:27 swarming activity was not possible to observe due to insufficient ambient light. It is, however, possible that swarming could still have been taking place after that time. The average number of mosquitoes caught per swarm was between 20 to 30 mosquitoes. This is only an estimation based on the number of adults collected.

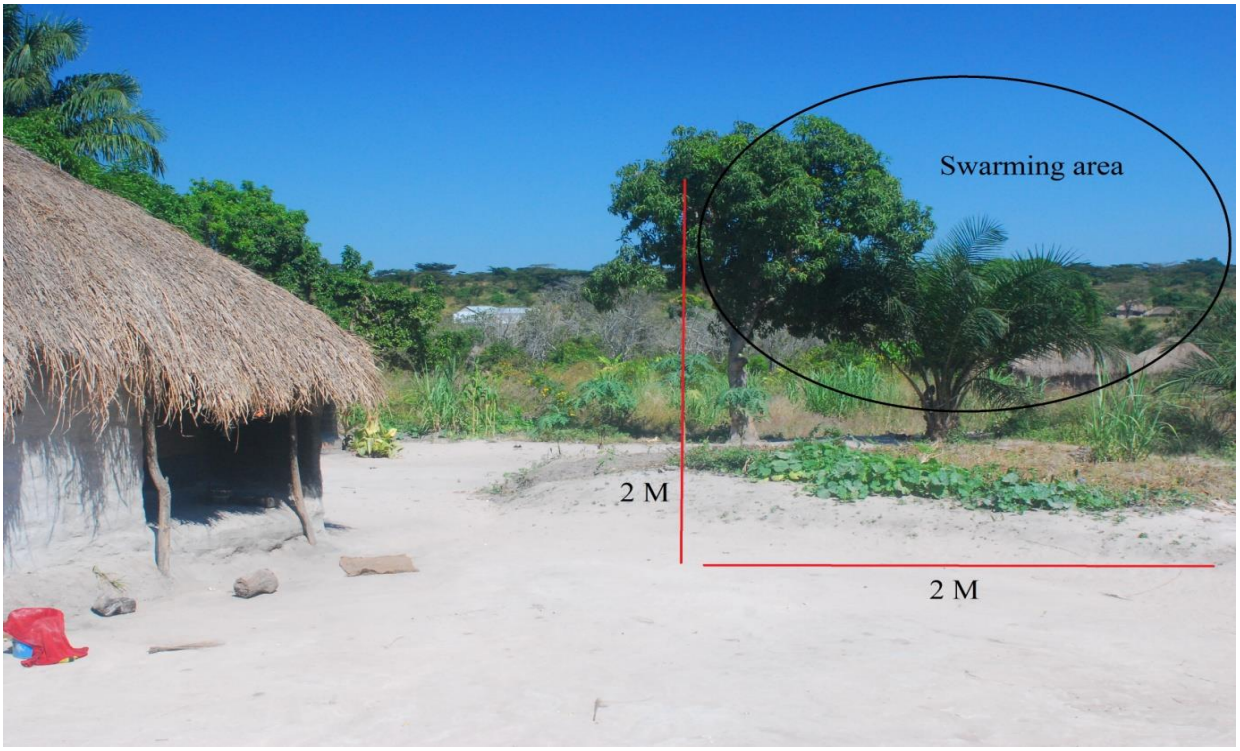


Figure 2.4 Photo of swarming area

The swarming area was in area of 2 m by 2 m, above a white/greyish, sandy area which had a defined break into a green vegetated patch (inside the black oval). The red vertical line indicates the height of the swarming area and the horizontal line indicates the width of the swarming area.

2.4.2 Species identification

A total of 89 adult mosquitoes were collected from the swarms and amongst these, 80 were males and 9 were females (Table 2.5). Samples were separated according to their morphology into either the *An. funestus* group (n=82; 75♂ and 7♀) or unknown (if morphological identification was not clear; n= 7; 5♂ and 2♀) (Table 2.5). Of the seven unknown samples, three were molecularly identified as *An. gambiae s.s.* (Table 2.6). The remaining 4 samples (2♂ and 2♀) could not be identified to a species level (Figure 2.5 and Table 2.6). All the samples morphologically identified as *An. gambiae* complex were collected from one specific swarm (swarm #6) and were males. PCR followed by *Hha*I digest confirmed that they were *An. gambiae s.s.* (previously classified as S-molecular form) (Figure 2.5).

Five of the six swarms contained females (n=9), with the majority of the swarms (#1, #2, #3, #5 and #6) containing only one female, while swarm #4 contained four females (Table 2.5). There were 7 females that belonged to the *An. funestus* group, according to morphological analysis, and one female that was too damaged to identify but was still processed with the other samples. This

sample, however, failed to amplify in the *An. funestus* group or *An. gambiae* group PCR assays and remained unknown (Table 2.5 * and Figure 2.5).

Two males were morphologically identified as belonging to the *An. funestus* group but were molecularly identified as *An. leesoni* (Table 2.5 and Figure 2.5).

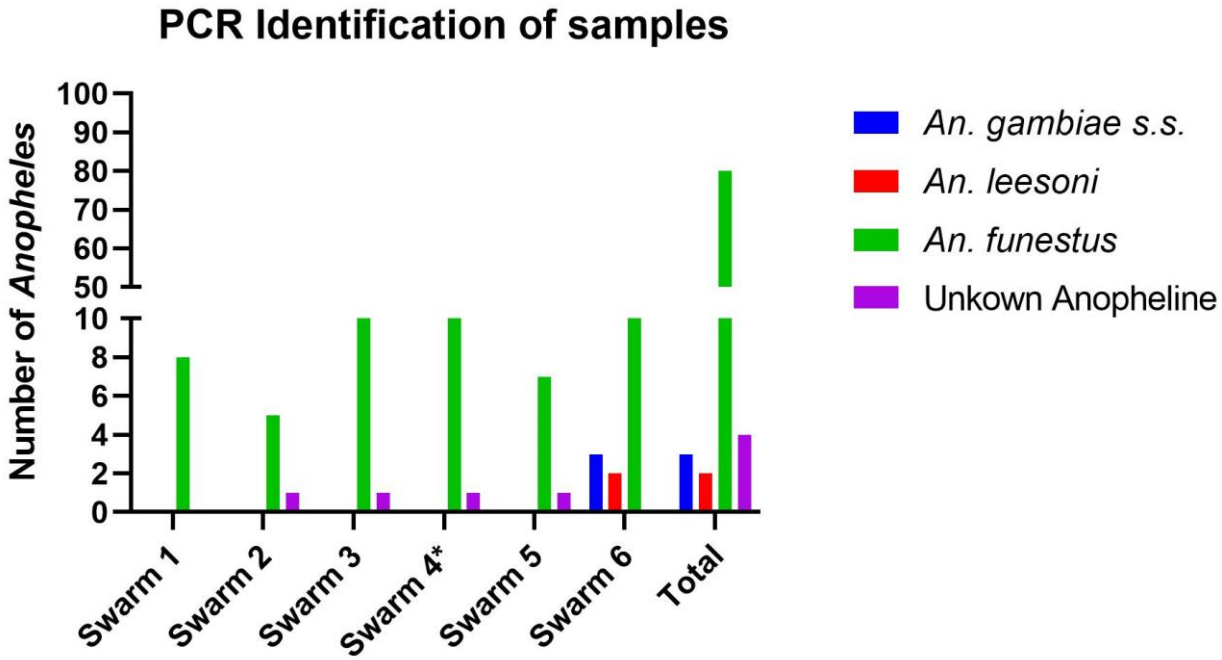


Figure 2.5 Molecular species identification of six wild *An. funestus* swarms

The colored bars represent the different species that were identified, x-axis shows which swarm the species came from, the y-axis shows the number of anophelines identified. * Swarm 4's unknown anopheline was a female that was too damaged to identify morphologically and failed to amplify in any of the species identification PCRs. An E.g., of the PCR results can be found in Appendix 8.1.

2.4.3 Species population identification and clade characterization

A total of 82 mosquitoes from the *An. funestus* group were collected, consisting of 75 males and 7 females. Through the use of species-specific PCR, there were two males that were identified as *An. leesoni* (2.6%) from swarm #6, while there were 73 *An. funestus s.s.* samples identified from all of the swarms (Table 2.5** and Figure 2.5). The male and female *An. funestus s.s.* samples were further genotyped and belong to the MW-RFLP type (n=80; 73♂ and 7♀) (table 2.5). Clade analysis identified both clade I (n=68; 64♂ and 4♀) and clade II (n=12; 9♂ and 3♀). Both clade

types were found in most of the swarms, except for swarm #2 (n=4) and #4 (n=19) which were only represented by clade I (Table 2.5).

Table 2.5 Samples collected from Nchelenge swarms

Swarm #	Total samples collected (M:F:T)	Unknown morphology (M:F:T)	*An. funestus* group (M:F:T)	*An. funestus s.s.* (M:F:T)	Mated F only	RFLP MW-type (M:F:T)	Clade I (M:F:T)	Clade II (M:F:T)
1	7:1:8	0:0:0	7:1:8	7:1:8	1	7:1:8	5:0:5	2:1:3
2	5:1:6	1:0:1	4:1:5	4:1:5	0	4:1:5	4:1:5	0:0:0
3	25:1:26	1:0:1	24:1:25	24:1:25	1	24:1:25	20:0:20	4:1:5
4	19:4:23*	0:1:1	19:3:22	19:3:22	2	19:3:22	19:3:22	0:0:0
5	7:1:8	0:1:1	7:0:7	7:0:7	0	7:0:7	6:0:6	1:0:1
6	17:1:18	3:0:3***	14:1:15**	12:1:13	0	12:1:13	10:0:10	2:1:3
Total	80:9:89	5:2:7	75:7:82	73:7:80	4	73:7:80	64:4:68	9:3:12

M=male; F=female; T=Total
*One female sample from this swarm was damaged beyond morphological identification and was molecularly identified as an Unknown anopheline.
**Two males were morphologically identified as *An. funestus* group but were molecularly identified as *An. leesoni.*
***These 3 morphologically unknown male samples were later molecularly identified as *An. gambiae s.s.*

2.4.4 Genitalia rotation observations

The stages of genitalia rotation were analysed for all the swarming males regardless of species as seen in figure 1.7 (Table 2.6). The majority (89%) of *Anopheles funestus s.s.* males was classified as Stage 4, and only two samples (2.8%) were Stage 3. One *An. leesoni* male showed Stage 3 rotation while the second sample was fully rotated (Stage 4). Two *An. gambiae* samples were fully rotated (Stage 4), while one sample could not be analysed due to sample damage. Males from the "unknown anophelines" (n=4) were also recorded as Stage 4. The observations were verified and observed by Dr. Yael Dahan-Moss, who published the paper: Dahan, Y.L. and Koekemoer, L.L., 2014. Analysis of the genitalia rotation in the male Anopheles funestus (Diptera: Culicidae). Acta Tropica, 132, pp. S20-S25, which characterises the different stages of genitalia rotation

Table 2.6 The frequency of genitalia rotation of swarming males

	Morphologically identified as *An. funestus* group		**Morphologically Unknown**	
	Anopheles funestus s.s.	***Anopheles leesoni***	***Anopheles gambiae***	**Unknown anophelines**
Stage 3	2 (2.8%)	1 (50%)	-	-
Stage 4	65 (89%)	1 (50%)	2 (66%)	2 (100%)
Too badly damaged to evaluate rotation	6 (8.2%)	-	1 (33%)	-
Total	73	2	3	2

2.4.5 Spermathecae dissection

Only 4 of the 8 females were in suitable condition to be used for spermathecae dissections. The other 4 females were too badly damaged, due to the transportation of the samples, to have their spermathecae accurately assessed. All 4 (100%) of the females that were dissected had mated. Swarms 1 and 3 contained one mated female each, whilst swarm 4 contained two mated females (Table 2.5).

2.5 Discussion

Wild swarms of mosquitoes of the *An. funestus* group were collected in Nchelenge, Zambia. The adult males were characterized based on their genitalia rotation stage, as well as their genotype and species. This was the first study of its kind and was subsequently published (Zawada *et al*., 2018).

The observations made of the swarming activity in Nchelenge was consistent with Charlwood's *et al.* (2003) findings, which confirmed that *An. funestus* swarmed 2 – 4 m above the ground, which was later confirmed by Kaindoa *et al.* (2019). However, Harper's (1944) observed that *An. funestus* swarms occurred inside the threshold or entrance of a dwelling, and that swarming occurred a foot (30 cm) above the ground. There is a possibility that the swarm that Harper (1944) observed was of another member of the *An. funestus* group, since identification to a species level was not possible because the molecular techniques were not available at that time. For example, *An. funestus s.s., An. parensis* and *An. vaneedeni* are morphologically identical (Gillies and Coetzee, 1987) and can only be identified through a standard multiplex PCR. Another possibility is that the swarming behaviour and position of *An. funestus* might change depending on environmental conditions. This study did not observe any swarming on the thresholds or entrances of huts in Zambia, although the field visit was too short to draw definitive conclusions. This investigation includes the observations that were made regarding the height of the swarms as well as their relative location to homes. Since swarming was only observed and identified in one specific location, this information is of limited value. More studies and extended field work will provide valuable further insights.

The swarms in this study were observed close to a stream as was also found by Charlwood *et al.* (2003). Charlwood *et al.* (2003), like others (Diabaté *et al.*, 2006; Dabiré *et al.*, 2014), observed that the *An. gambiae* complex members swarm over areas of contrast on the ground, but *An. funestus* appears to avoid these contrasting ground markers (both light and dark). My observations in Nchelenge indicated that the swarming occurred at the transition point between bare ground (light colour) and low-lying vegetation (dark colour), which appears to act as a horizontal contrasting marker. This kind of transition point for swarming was reiterated by Kaindoa *et al*. (2019). This is in contrast to the Charlwood *et al.* (2003) study, where *An.*

gambiae complex males tend to swarm over contrasting markers (Charlwood *et al.,* 2003) whereas *An. funestus* appears to avoid ground markers (both light and dark).

Although swarm adults were depleted during collection, this did not appear to prevent a swarm forming in the same place on consecutive nights. Although the reason for this is difficult to establish conclusively, it is most likely that newly emerged males that have subsequently matured and were therefore ready for mating, replenished the depleted swarm. Whether this will also occur in swarms in other parts of Africa is still to be determined. It is also unknown if *An. funestus* swarm in fewer larger swarms or in multiple small swarms. Swarm numbers that were reported by Charlwood *et al.* (2003) were also relatively low and had an average of less than 50 adults/swarm. Harper (1944) estimated that there were 300 – 500 mosquitoes that were present in the swarm that he reported on. Subsequent to this study, Kaindoa *et al.* (2019), reported very small swarm sizes of *An. funestus s.s.* (10 – 15 adults/swarm), but this was mostly attributed to low vector densities at the time of the study.

This study (also published in Zawada *et al.*, 2018), performed in Zambia, and another by Kaindoa *et al.* (2019) in Tanzania, found that some form of mixed swarming does occur within *An. funestus* swarms. This corresponds to mixed swarms being reported by Charlwood *et al.* (2003) in Mozambique where he collected one *An. gambiae s.s.* sample in a wild *An. funestus* swarm. Although the overall abundance of other species in *An. funestus* swarms are low, it most likely indicates that *An. funestus* swarming occurs in close proximity to other species, e.g., the *An. gambiae* complex. Zawada *et al.* (2018) caught three *An. gambiae s.s.* and two *An. leesoni* in a confirmed *An. funestus s.s.* swarm. Recently, Dahan-Moss *et al.* (2020) showed that morphologically mis-identified members of the *An. gambiae* complex can be identified as *An. leesoni* in the *An. funestus* group PCR assay. The *An. leesoni* samples reported here were not tested in the *An. gambiae* complex PCR to confirm their identity as the samples were no longer available in 2020. However, as morphological analysis placed these samples in the *An. funestus* group, it is likely that these samples are *An. leesoni.* The *An. leesoni* species is also known to occur in Zambia (Gillies and DeMeillon, 1968). Kaindoa *et al.* (2019) caught between 2 and 7 *An. arabiensis* males within a confirmed and targeted *An. funestus s.s.* swarm in Tanzania.

The Kaindoa *et al.* (2019) study agreed with the results of this study. *Anopheles funestus* swarms occur 2-4 m off the ground and over horizontal contrasting markers (Zawada *et al.*, 2018;

Kaindoa *et al.,* 2019). The Kaindoa *et al.* (2019) study also confirmed the findings of this study that mixed swarming can occur within *An. funestus* swarms (Zawada *et al.,* 2018; Kaindoa *et al.,* 2019). Kaindoa *et al.* (2019) found a mean size swarm size of 6.6 to 10.8. This does not suggest that there is any evidence that larger swarms attract mixed species. However, a more in-depth study needs to be performed in order to answer the question of whether larger swarms may attract other species.

It is interesting to note that the mixed swarms were only recorded after 4 successive nights of collecting swarms from the same location. A possible explanation for this may be that the swarm collections had decreased the local population of *An. funestus* and this dropped the swarming competitiveness of the *An. funestus* swarm allowing for another species, other than *An. funestus*, to enter the swarm.

Mixed species swarms' collections are not unusual and have been reported by others. Mixed swarms are characterized by having a low frequency of the mixed species within the swarm. For example, Sawadogo *et al.* (2014) reported that in Burkina Faso mixed swarms occurred between *An. gambiae s.s.* (≤ 10%) and *An. coluzzii* (≥ 90%), but this was not the case in Mali (Sawadogo *et al.,* 2014). Sawadogo *et al.* (2014) showed that *An. gambiae s.s.* swarmed earlier (3:35 ± 0:68 min:sec) and stopped earlier (4:51 ± 1:21) compared to the *An. coluzzii* species, and that even if the same marker was used, the swarming times seldom overlapped. Environmental factors can result in a few swarming *An. gambiae* males to from time to time overlap with those of *An. coluzzii* swarming and *vice versa*, resulting in the low frequency of mixed swarms. Diabaté *et al.* (2006) showed that the frequency of hybridisation between *An. gambiae s.s.* and *An. coluzzii* is still lower than the frequency of mixed swarming. They therefore concluded that mate recognition is more important than swarm segregation, which means that the occurrence of mixed swarms does not imply that cross-mating and hybridization occurs (Diabaté *et al.,* 2006).

The small number of females collected makes it difficult to draw solid conclusions of cross-species mating (either RFLP or clade mixed mating). This study was not designed to investigate mating pairs and the sperm from mated females was not analysed as part of this study. Possible cross-species mating, and mating pairs should, however, be investigated in future.

On subsequent species identification, the *An. funestus* samples were also genotyped to characterize their molecular typing. RFLP-typing confirmed that the only genotype present in Nchelenge (500 km north of Ndola) was the MW-type. Previously Koekemoer *et al*. (2006) also identified the Z-, MW- and Y type from Ndola in Zambia, (close to the Democratic Republic of the Congo) and the MW-type from Chipata (bordering Malawi). The sample size was, however, small with only seven from Ndola and one mosquito from Chipata analysed (Koekemoer *et al.,* 2006).

In addition to the RFLP-typing, the samples were also genotyped using a mitochondrial marker to differentiate between clade I and II. Results revealed that both clades I and II occurred in all of the six swarms. The ratio of clade I (85%): clade II (15%) is similar as to that recorded in the wild adult population 80% clade I and 20% clade II (Choi *et al.,* 2015). The small sample size of the swarms (n=80) might explain the shift towards clade I. The small sample size (n=80) is due to the difficulty in catching live specimens and the time constraints of being in the field. To date there has been no evidence to suggest that mitochondrial subpopulations are resulting in population isolation. Michel *et al.* (2005) postulated that clade II originates from historical gene flow between previously isolated populations or even related species. In addition to this, the fact that adults of both clade types were found in the same swarm implies that no specific clade swarming characteristics were identified during this study.

This study clearly showed that the majority of the males swarming are sexually mature (fully rotated claspers), further indicating that immature males do not have the tendency to swarm. This corresponds to a similar study performed by Kaindoa *et al.* (2019) in Tanzania. The underlying molecular and biological processes that mitigate the sexual maturation of the male are still unknown. Transcriptomic studies provide valuable insight into these processes which should be investigated as they could possibly reveal novel control targets for vector control.

3 Transcriptomic studies of male *An. funestus* during sexual maturation

3.1 Introduction

Anopheles funestus swarming only takes place once the males have sexually matured and the genitalia have rotated to Stage 4 (as discussed in Chapter 2). Functional genomics could be valuable to assist the understanding of the processes of male sexual maturation. There are several techniques available to study the relationship between genotype and phenotype. Gene expression analysis measures the relative abundance of mRNA between two biological conditions. There are several molecular techniques available for the study of quantitative differential gene expression. The gold standard before Next-Generation sequencing (NGS) was the microarray technology (McGettigan, 2013).

In contrast to hybridization techniques such as the microarray, there are other techniques such as the Sanger sequencing of cDNA or Expressed Sequence Tag (EST) libraries (Boguski *et al.,* 1994; Gerhard *et al.,* 2004) and Tag-based methods such as the Serial Analysis of Gene Expression (SAGE) (Velculescu *et al.,* 1995; Harbers and Carninci, 2005). RNA-sequencing (RNA-seq) has since superseded these technologies and is the current gold standard for studying gene expression in transcriptomic studies (McGettigan, 2013).

As with all techniques there are disadvantages to next generation sequencing techniques. One disadvantage is the cost of these sequencing technologies; another disadvantage is that some sequences might not map onto the reference genome (Wang *et al.,* 2009). RNA-seq does, however, directly determine the mRNA sequence of the organism. RNA-seq therefore has clear advantages over other pre-existing technologies. RNA-seq provides the depth of sequence coverage of a transcript that should be proportional to its expression level makes it an extremely powerful means of measuring gene-expression (Wang *et al.,* 2009).

To my knowledge there has only been this one study (Crawford *et al.,* 2010) that performed RNA-seq using *An. funestus*, most NGS techniques have been performed on the *An. gambiae* species (Crawford *et al.,* 2010; Bonizzoni *et al.,* 2012; Bryant *et al.,* 2019). These studies were all performed on female anophelines and no attempt has been made to use this NGS technique to study the male *An. funestus* sexual development (Crawford *et al.,* 2010; Bonizzoni *et al.,* 2012; Bryant *et al.,* 2019).

RNA-seq uses total RNA, which is converted into cDNA fragments with adaptors attached to one or both ends. These adaptors contain the amplification element and the primary sequencing site. These fragments are then sequenced to obtain short sequences from both directions (paired-end sequencing) in a high-throughput manner. These 30 - 400 bp long sequences are then aligned to a reference genome or assembled *de novo* to produce a genome-scale transcription map. The map reflects the transcriptional structure and the level of expression of each gene (Wang *et al.,* 2009). Once targets are identified through RNA-seq, traditional Real-time quantitative PCR (qPCR) is used to confirm the abundance of a specific transcript(s) (Heid, 1996). This is generally followed up with a "loss of function" technique (i.e., RNAi) to verify that the target is associated or plays a role in that.

To my knowledge there have been no studies hypothesizing genes or transcripts that are differentially expressed and involved with the genitalia rotation process.

3.2 **Rationale**

In the previous study, Chapter 2, it was discovered that only mature males (with Stage 3 and 4 rotated genitalia) swarm and participate in the copulating activity within a swarm. The rotation of the genitalia was used as a physiological marker to differentiate between immature (Stage 0) and sexually mature (Stage 4) adults. It was hypothesized that the genitalia rotation causes a behavioral change, which allows the males to start swarming. It is possible that a gene expression change occurs which leads to the rotation of the male *An. funestus* genitalia. It has been noted that there is a delay in transcription levels at the point where the phenotype is observed. There is a time delay from when the transcription levels are the highest and when the phenotype is transcribed. It was, therefore, decided to use Stage 3 adults to elucidate transcription levels to give the highest representation of transcription levels for sexually mature adults. It was hypothesized that differentially expressed transcripts involved in developmental mechanics are associated with sexual maturation of the *An. funestus* males. The transcripts that are expected to be differentially expressed between different ages of the *An. funestus* male will be putatively annotated to chitin (e.g., CPRR 1 and CPRR 2 families), muscle (myosin, actin, troponin, flightin and tetraspanin) genes, immunity genes (phenol oxidases, lysozyme, TNF-receptor and CYP450) and developmental genes (farnesyltransferase, allototropin and 20-hydroxy-ecdysone signal transduction), that are possibly involved in sexual maturation.

The aim of this investigation was to compare and discover differentially expressed genes of *An. funestus* between Stage 0 and Stage 3 of the genitalia rotation process. This study will help increase the fundamental knowledge of the male *An. funestus* biology specifically involved in the process of genitalia rotation that can help further our biological understanding of this important malaria vector.

3.3 Materials and methods

3.3.1 Biological material, RNA extraction and preparation

Anopheles funestus males from the Botha De Meillon insectary at the NICD, NHLS, Sandringham, Johannesburg, South Africa, were used for the purpose of this study. The *An. funestus* (FUMOZ) colony, which originates from southern Mozambique in 2000 (Hunt *et al.*, 2005), was used. The samples were collected in the same phase of the insectary light : dark cycle, to prevent circadian rhythm from biasing the gene expression results.

Ten male *An. funestus* mosquitoes, equivalent to 100mg of biological material, prior to rotation (Stage 0), were homogenized and RNA was extracted using Life Technologies TRIzol®Reagent (CAT. No.: 15596). The homogenized mosquitoes in TRIzol were incubated at room temperature for min. Chloroform (200 µL) was added to the TRIzol/mosquito mixture, which was then shaken by hand and incubated at room temperature for min, followed by centrifugation as 12 000 × *g* for 10 min at 4°C. The upper-most layer was aliquoted into a new 1.5 mL microcentrifuge tube and 500 µL of isopropanol was added and incubated for a further 5 min at room temperature. Following incubation at room temperature, the RNA was precipitated by centrifugation at 12 000 × *g* for 10 min at 4°C. The aqueous solution was removed leaving an RNA pellet, which was washed with 70% absolute ethanol. The ethanol was removed and air dried for 5 min before being resuspended in 50 µL of diethyl pyrocarbonate (DEPC) water. No DNase treatment was performed at this stage, as poly-A selection in the RNA-sequencing step will remove all contaminant DNA.

The RNA concentration was measured in ng/µl using a NanoDrop 2000 Spectrophotometer. The RNA was evaluated for the presence of proteins, phenols, or other contaminants by considering its absorbance at or near the 280 nm wavelength (A280/260 ratio). The quality or integrity of the RNA was evaluated by electrophoresing the RNA on an agarose gel (Noriega and Wells, 1993; Jing *et al.*, 2006), where the definition of the RNA fragment will determine its quality (Noriega and Wells, 1993). This showed the presence of the “hidden break” by presenting one clear band characteristic of insect RNA (Winnebeck *et al.*, 2010). The resulting RNA was quality and quantity checked by using Agilent Technologies 2100 bioanalyzer (or 2200 TapeStation). The use of degraded RNA can result in low yield or over representation of the 5’-ends of the RNA molecule.

The original quality control criteria were that the concentration (ng/µL) > 10 ng/µL, volume (µL) > 10 µL and the total amount (µg) > 0.1 µg.

The RNA extraction procedure was repeated with 10 *An. funestus* males which had their genitalia rotated at stage 3. Three biological replications were performed for both time points. These three replications for Stage 0 and Stage 3 were labelled and named according to their stage and biological replicate number, i.e., Stage 0 replicate 1 was named St. 0.1 and Stage 3 replicate 1 was named St. 3.1, and so on.

3.3.2 RNA-sequencing and data production

The RNA was couriered to Macrogen Inc. (Seoul, South Korea) for RNA-seq using the Illumina HiSeq2000/2500. Macrogen Inc. performed Poly-A selection and then randomly fragmented the purified RNA for short read sequencing. Selection of 3′ Poly-A tails ensures that only the RNA that is mature and are coding sequences, is selected. This was done by mixing RNA with Poly-T oligomers that were covalently attached to a substrate. This means that the Poly-A selection ignores non-coding RNA and, therefore, does not introduce a 3′ bias. The fragmented RNA was reverse transcribed into cDNA and adaptors were ligated on both ends of the cDNA fragments to allow for PCR amplification. The adaptor sequences that were ligated are as follows: read 1 adaptor: 5′ AGA TCG GAA GAG CAC ACG TCT GAA CTC CAG TCA 3′; read 2: 5′ AGA TCG GAA GAG CGT CGT GTA GGG AAA GAG TGT 3′. Fragments with insert sizes between 200 - 400 bp were selected for paired-end sequencing, both ends of the cDNA were sequenced by read length.

The data analysis was performed on the sequences to set up expression profiles and functional annotation reports. The raw read sequences were quality controlled by FastQC v0.10.0, which is a program that ensures data integrity before analysis on the raw sequences, where the overall read quality, total bases, total reads, GC (%) and basic statistics were calculated. The GC-content is the percentage of nitrogenous bases in the RNA reads, which gives the proportion of guanine and cytosine bases compared to thymine and uracil bases within the reads. The phred quality score is a measure of the identification of the nucleobases generated by NGS Sequencing. A phred quality score of 20 means that there is a 99% accuracy and reads over 20 were accepted as good quality,

which gives the Q20 (%) score. A phred quality score of 30 (Q30) means that there is a 99.9% accuracy and reads over 30 were accepted as good quality. Low quality, adaptor sequences, contaminant DNA or PCR duplicates were removed in the analysis to reduce bias.

Trimmomatic 0.32 was used to trim the paired end reads by removing adapter sequences and bases with a base quality lower than 3 from the ends. The sliding window method was used to trim bases off reads that did not qualify for a window size of 4 and a mean quality of 15. Any reads that had a read length shorter than 36 bases were not included to produce clean data.

3.3.3 Reference mapping and assembly

TopHat version 2.0.13, bowtie2 2.2.3, which is a bioinformatic analysis tool used for the fast and high through-put of shotgun cDNA sequencing reads generated by transcriptome technologies, was used. The trimmed reads were then mapped to the reference genome VectorBase (VectorBase AfunF3.1). This was done by using the mapping results to provide provisional exon location and exon junctions.

The aligned reads that contain paired-end information were used to assemble the transcript by using Cufflinks version 2.2.1. Cufflinks is a sequence assembly program that connects reads from the mapping results using the TopHat aligner which can then predict the expression level of the assembled transcriptomes. This provided the known transcripts, novel transcripts, and alternative splicing transcripts.

Known genes and transcripts were assembled with StringTie based on the reference genome *An. funestus* (AfunF3.1) available on VectorBase (www.vectorbase.org) (Neafsey *et al.*, 2015; Conesa *et al.*, 2016). After assembly, the abundance of gene/transcript was calculated in the read count and normalized to the value as FPKM (Fragments Per Kilobase of transcript per Million mapped reads) for each sample.

These expression profiles were based on the transcript length and depth of coverage which was be represented as a normalization and the value unit was in FPKM (Fragments per Kilo base of transcript per Million mapped reads). The raw signals were scattered along a wide range and most signals were concentrated on the low signal value, to facilitate the log2 transformation, 1

was added to the raw FPKM signal. This reduced the range of signals and produced an even distribution of data. Statistical hypothesis testing was used to filter out transcripts that express differentially.

The raw sequencing data from the assembly was uploaded to the SRA under the accession number SRX7649914 (https://www.ncbi.nlm.nih.gov/sra/SRX7649914).

3.3.4 Differentially expressed gene analysis

The differential gene expression of *An. funestus* FUMOZ Stage 3 vs. *An. funestus* FUMOZ Stage 0 was compared. The differential expression between these two biological sample types was done in triplicate (i.e., 3.1 vs. 0.1, 3.2 vs. 0.2 and 3.3 vs. 0.3).

The differentially expressed genes (DEGs) had functional annotation and gene-set enrichment analysis performed on them using Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) databases done by Macrogen Inc. These databases were used as the *An. funestus* genome has not been fully annotated and these databases can be used to assign putative annotations to the differentially expressed transcripts.

Blast2Go (B2G) was used as an annotation finder for all the transcripts that Macrogen Inc. could not find any annotations for. Blast2Go is a bioinformatics platform for high-quality functional annotation and analysis of genomic datasets. The Entrez identification numbers or gene I.D.'s were found by searching for the transcript I.D. on the National Center for Biotechnology Information website (NCBI). These Entrez I.D.'s were listed and were processed through B2G. Blast2Go ran the program blastx-fast on all the transcripts using the NCBI nucleotide database. There was no taxonomy filter to provide the most widespread results, the blast expectation value (E-value) was set at 1.0E-3. An InterProScan was performed, which translates the FASTA sequences into amino acid sequences which then searches for those protein sequences in the following databases: BlastProDom, FPrintScan, HMMPIR, HMMPfam, HMMSmart, HMMTigr, ProfileScan, HAMAP, SuperFamily, HMMPanther, Gene3D, CDD, SFLD and MobiDBLite. The Blast hits of each sequence were then mapped with Gene Ontology Terms. B2G performs different mapping steps to link all Blast hits to the functional information stored in the Gene Ontology (GO) database (Conesa *et al.*, 2005). B2G uses different public resources provided by

the NCBI, PIR and GO to link the different protein I.D.'s to the information stored in the Gene Ontology database, which contains several million functionally annotated gene products for hundreds of different species (Conesa *et al.*, 2005). An evidence code is associated to these annotations which provides information about the quality of the functional annotation. B2G was then used to run annotation searches, with an annotation Cut-Off of 55 to keep the annotations specific, a GO weight of "5" was used to make sure that the unique annotations are kept. The Blast parameters used an E-value of 1.0E-6 to ensure that only high identity hits are used for annotation, an HSP-Hit Coverage Cut-Off of 0 and a Hit Filter of 500. These bio-informatic steps provided further annotation on the RNA-seq results. These annotation results were then added to the overall annotation results.

The expression profiles and functional annotation reports identified the transcripts that were differentially expressed between the stages. Through gene ontology, transcripts were selected that were most likely to be associated with the sexual maturation process, based on known biological function from other organisms. Transcripts that were unannotated but were significantly differentially expressed between the two time points, which implies that they may have a possible significant role during the sexual maturation process, were also selected. All data is available upon reasonable request from Prof. Lizette Koekemoer (Lizette.koekemoer@wits.ac.za) or Mr. Jacek Zawada (Jacek.zawada@wits.ac.za).

3.4 **Results**

3.4.1 Assessment of RNA quality and degradation

The RNA integrity needs to be assessed to ensure that the RNA is intact and not degraded for the results from the expression data to be trustworthy. Figure 3.1 shows the total RNA extracted from *An. funestus* (FUMOZ) at their different stages (Stage 0 and Stage 3) which was electrophoresed on a native RNA gel. As can be seen in Figure 3.1 there is a strong 18S band and a faint 5S band showing that all of the RNA was intact in all of the samples. The presence of one clearly defined fragment shows that there is the "hidden break" characteristic of insect RNA (Noriega and Wells, 1993).

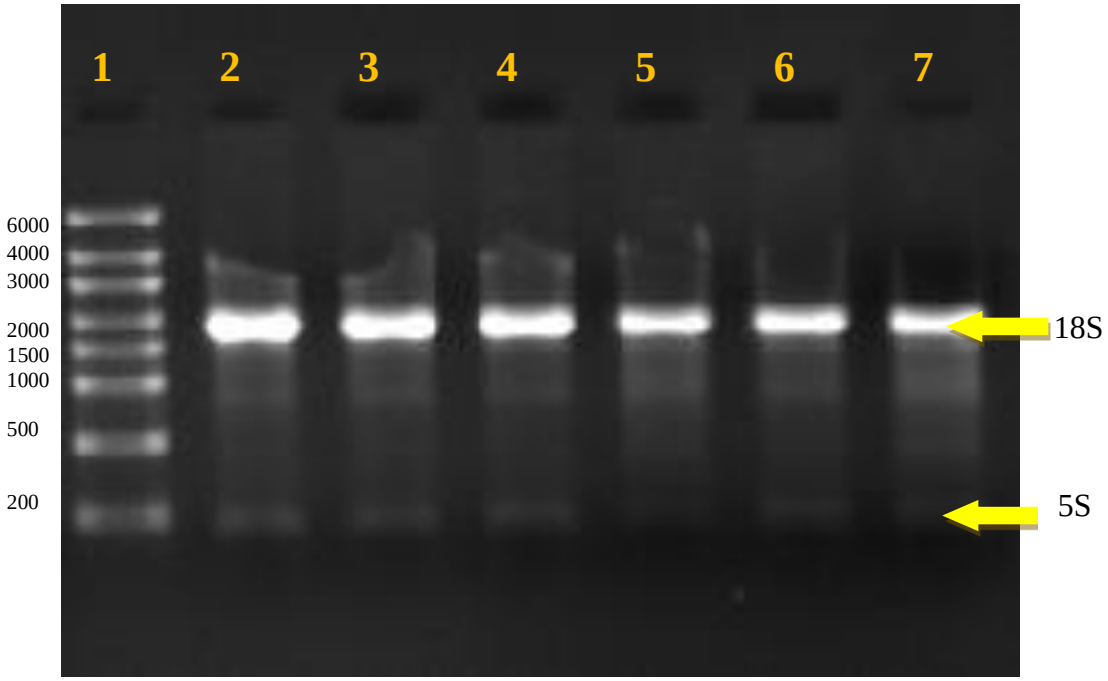


Figure 3.1 Native RNA gel of total RNA

Native 2% TBE RNA agarose gel. Lane 1: RNA Molecular Weight Marker (RiboRuler High Range RNA Ladder; CAT no.: SM1821). Lanes 2-4: Stage 3, biological replicate 1 (3.1); (3.2) and (3.3) 1 µg of total RNA, respectively. Lanes 5-7: Stage 0, biological replicate 1 (0.1); (0.2) and (0.3) 1 µg of total RNA, respectively.

The Nanodrop 2000 results of these samples can be seen in Table 3.1. All the samples from both Stage 0 and Stage 3 had a nucleic acid concentration ranging between 263.7 ng/µL – 735.9 ng/µL. All the samples had a 260/280 ratio above 2, which shows that there was minimal contamination of proteins, or any other contaminants that absorb at the 280 nm wavelength. The 260/230 ratio was deemed acceptable and indicated that there was a low amount of phenol contamination, or any other contaminants that absorb at the 230 nm wavelength.

Table 3.1 Nanodrop 2000 results

Sample*	Nucleic Acid Concent. (ng/µL)	A260/280	A260/230	Sample Type
St. 3.1	518.1	2,11	2.15	RNA
St. 3.2	446	2,15	1,68	RNA
St. 3.3	447	2,15	1,2	RNA
St. 0.1	263,7	2,15	1,46	RNA
St. 0.2	559,2	2,13	1,68	RNA
St. 0.3	735,9	2,17	1,94	RNA

*St. n.n = Stage n. Biological replicate number n

This RNA was electrophoresed on an Agilent Technologies 2100 Bioanalyzer (or 2200 TapeStation) and the results can be seen in Table 3.2. The RNA Integrity Number (RIN) is an automated approach of showing the quality of RNA, 10 being of the highest quality and 1 being the lowest (Shroeder *et al.,* 2006). As can be seen in Table 3.2, all of the samples showed RIN values ≥ 9.6, indicating that the RNA was of extremely high quality. In Figure 3.2 the results of sample 0.1 (Stage 0, biological replicate 1) can be seen. There is a clear and defined peak at the 1,500-nucleotide mark, which shows the ribosomal 18S fragment. At the 315-nucleotide mark, the ribosomal 28S peak can be seen. The Bioanalyzer graphs of the samples that are not shown in Figure 3.2 can be found in Appendix 8.2.

Table 3.2 Agilent technologies 2100 bioanalyzer results

#	Sample Name*	Concent. (ng/µL)	Add Volume (µL)	Final Volume (µL)	Total Amount (µg)	RIN
1	St. 0.1	87.641	50	50	4.382	10
2	St. 0.2	189.787	50	50	9.489	9.9
3	St. 0.3	213.751	50	50	10.688	9.8
4	St. 3.1	170.150	50	50	8.508	9.6
5	St. 3.2	127.284	50	50	6.364	9.9
6	St. 3.3	128.753	50	50	6.438	9.6

*St. n.n = Stage n. Biological replicate number n

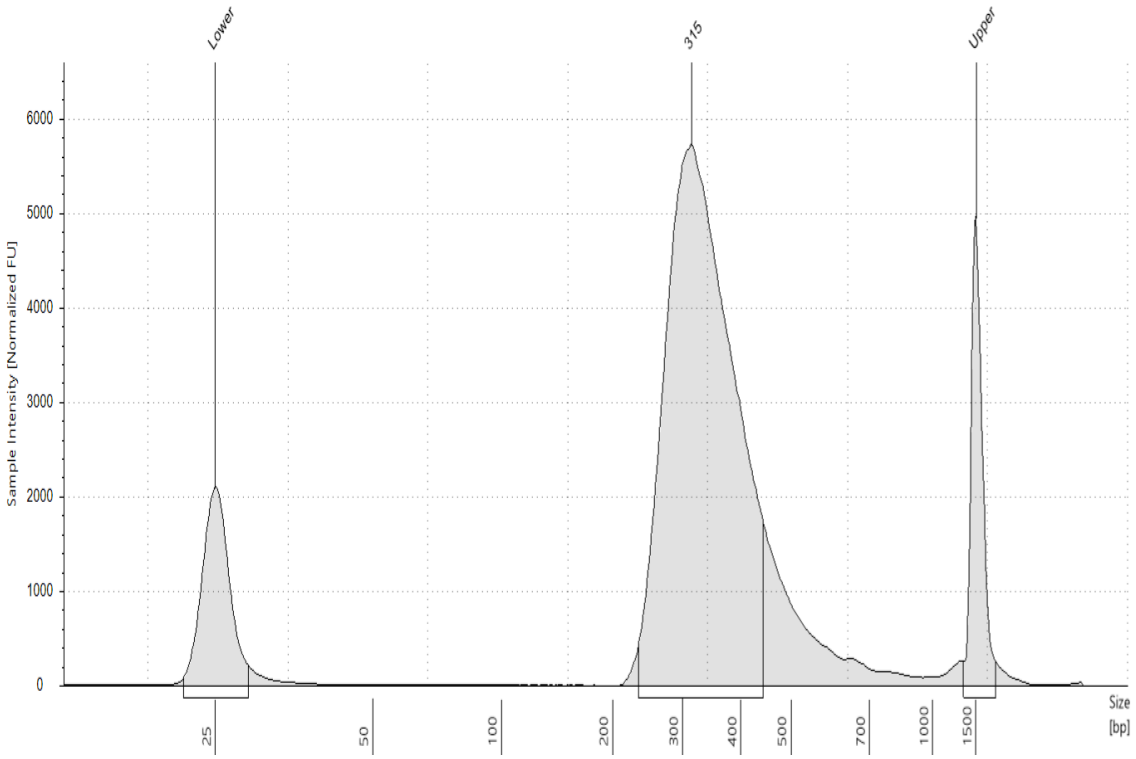


Figure 3.2 Agilent Technologies 2100 Bioanalyzer graph of St. 0.1, *An. funestus* total RNA

The "upper peak" indicates the 18S ribosomal fragment. The "315" peak (315 bp mark) indicates the presence of the 28S ribosomal fragment. The lower peak indicates the presence of the 5S ribosomal fragment at the 25 bp mark.

3.4.2 RNA-Sequencing and data production

The RNA-seq data was trimmed of their adaptor sequences as well as low quality bases. Trimming of the adaptor sequences was performed by matching the first 6-8 nucleotides of the adaptor, and then trimming the adaptor away.

After adaptor removal, sequence filtering was undertaken in order to remove other undesirable characteristics of the sequence data. If there were any unknown bases in the sequence, it was discarded, giving fewer total read bases in the trimmed data (Table 3.3). Table 3.3 shows the raw data statistics which shows the total read bases, total reads, GC percent content as well as their Q20 and Q30 scores, before trimming and sequence filtering. The Q20 (%) and the Q30 (%) in Table 3.3, shows the percentage of bases that had a Q20 score or above, and the percentage of bases with a Q30 score or above. The low-quality bases are trimmed according to theses Q20 and Q30 scores as can be seen in Table 3.3.

Raw data statistics, in Table 3.3, showed that the average number of bases sequenced between the three biological replicates of *An. funestus* males with genitalia rotated at Stage 0 and Stage 3 were 10.03 GigaBases (Gb) and 9,71 Gb, respectively. Both sample sets displayed similar GC content (44.92% and 42.72% respectively) with 95.87% of the *An. funestus* males with genitalia rotated at Stage 0 and 95.92% of the *An. funestus* males with genitalia rotated at Stage 3. The phred quality score is ≥30 (with an accuracy of each nucleotide call being 99.9%), which is the measure of the quality of the identification of nucleobases from sequencing. After removal of artifacts such as low-quality reads, adapter sequences, PCR duplicates and contaminant DNA, the average number of bases sequenced between the three biological replicates of *An. funestus* males with the genitalia rotated to Stage 0 and Stage 3 samples were 9.86 Gb and 9,57 Gb respectively (Table 3.3).

Table 3.3 Table of raw data statistics of the RNA-sequencing data of each replicate

Sample I.D.*	**Total read bases (Gb)****		**Total reads**		**GC (%)**		**Q20 (%)**		**Q30 (%)**	
	U	**T**	**U**	**T**	**U**	**T**	**U**	**T**	**U**	**T**
St. 0.1	8.736	8.611	86,499,406	85,690,086	45.03	45.03	98.78	99.09	96.14	96.59
St. 0.2	10.235	10.050	101,343,778	100,020,088	43.81	43.81	98.57	98.99	95.66	96.25
St. 0.3	11.110	10.931	110,000,778	108,804,004	45.93	45.93	98.65	99.00	95.82	96.33
St. 3.1	9.906	9.753	98,084,094	97,088,078	42.40	42.38	98.67	99.00	95.78	96.25
St. 3.2	7.756	7.635	76,792,800	76,021,002	41.91	41.89	98.74	99.07	96.00	96.46
St. 3.3	11.482	11.314	113,686,858	112,595,782	43.85	43.84	98.74	99.06	95.98	96.44

*St. n.n = Stage n. Biological replicate number n
U = Untrimmed data
T = Trimmed data
The total number of bases reads, GC (%), Q20(%) Q30(%), were calculated for the 6 samples. **Total read bases = Total reads X Read Length.
Q20(%): is the ratio of reads that have a phred quality score of over 20, Q30(%): is the ratio of reads that have a phred quality score of over 30.

Trimmomatic 0.32 removed the adapter sequences with a base quality lower than 3 from the ends. By using the sliding window method, bases of reads that do not qualify for window size 4 and mean quality 15 were trimmed. Reads with lengths shorter than 36 base pairs were dropped to produce clean data (Table 3.3).

3.4.3 Reference mapping and assembly

To map the cDNA fragments obtained from the RNA-seq, VectorBase (VectorBase AfunF3.1) was used as a Reference genome (Neafsey *et al.*, 2015). However, the *An. funestus* genome has not been fully annotated and new, multiple unannotated transcripts were obtained (Neafsey *et al.*, 2015; Conesa *et al.*, 2016). Table 3.4 shows the statistics obtained from HISAT2, which is known to handle spliced read mapping through the Bowtie2 aligner of the mapped and unmapped reads.

Table 3.4 Table of mapping data statistics

Sample I.D.*	# of processed reads	# of mapped reads (%)	# unmapped reads (%)
St. 0.1	85,690,086	61,082,915 (71.28%)	24,607,171 (28.72%)
St. 0.2	100,020,088	68,309,709 (68.3%)	31,710,379 (31.7%)
St. 0.3	108,804,004	80,476,568 (73.96%)	28,327,436 (26.04%)
St. 3.1	97,088,078	58,139,557 (59.88%)	38,948,521 (40.12%)
St. 3.2	76,021,002	44,459,508 (58.48%)	31,561,494 (41.52%)
St. 3.3	112,595,782	75,457,220 (67.02%)	37,138,562 (32.98%)

*St. n.n = Stage n. Biological replicate number n
processed reads: number of cleaned reads after trimming
of mapped reads: number of reads mapped to reference
# of unmapped reads: number of reads that failed to align to the reference genome.

Overall, the *An. funestus* male cohort with the genitalia rotated to Stage 0 had 71% mapped reads and 29% unmapped reads, while *An. funestus* males with the genitalia rotated to Stage 3 had 62% mapped and 38% unmapped reads (Table 3.4).

The density plots in Figure 3.3 show the expression distribution before and after adding 1 to the raw FPKM signal for the log2 transformation and quantile normalization.

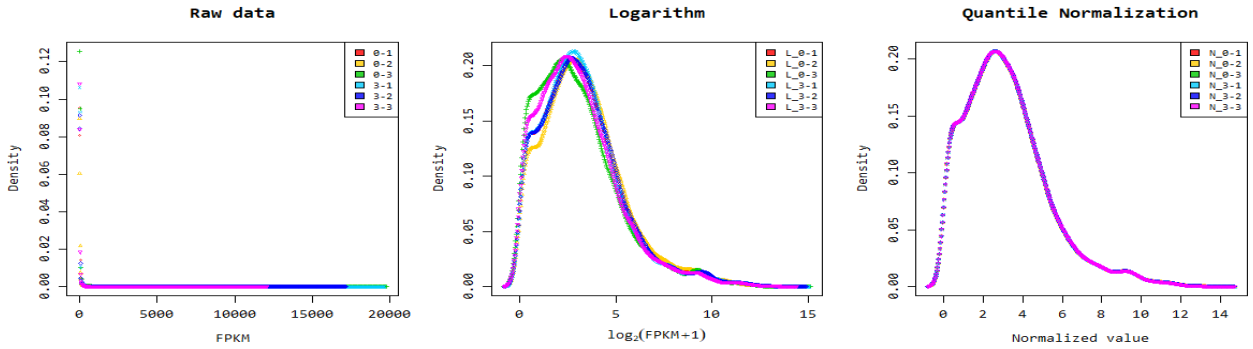


Figure 3.3 Expression density plot per sample

N_0-1 = Stage 0 - Replicate 1; N_3-1 =Stage 3 - Replicate 1.

The similarity between samples was obtained by using Pearson's coefficient of the log2 (FPKM +1) value as can be seen in the correlation matrix in Figure 3.4. The range was: -1≤ r ≤ 1, the closer the value is to 1, the more similar the samples were.

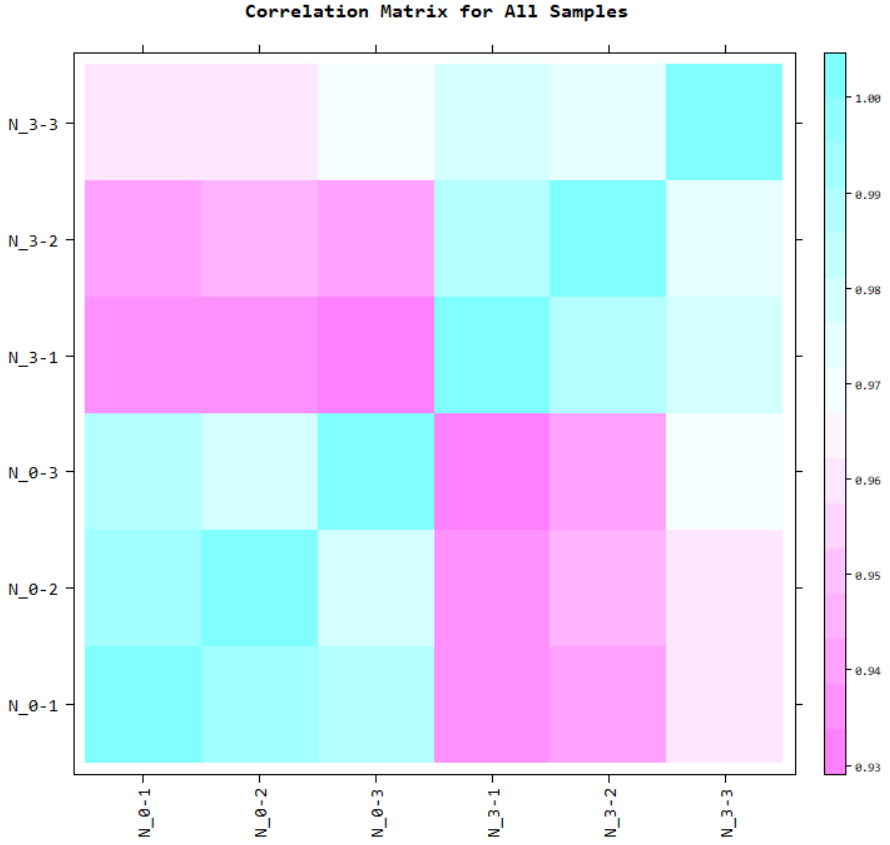


Figure 3.4 Correlation matrix of all the samples

N_0-1 = Stage 0 - Replicate 1; N_3-1 = Stage 3 - Replicate 1. The range of blue to purple hue indicates the relative difference of that sample to another. The closer the correlation is to 1.00 indicated by the solid blue hue which indicates the closeness of correlation. A correlation of 1.00 means that there is no difference between those two samples (N_0-1 vs. N_0-1). The closer the hue is to purple the less correlation there is and indicates a large difference between the two samples (for e.g.: N_3-1 vs. N_0-1).

Using each sample's normalized value, the similarity between the samples is graphically shown in a 2D plot to show the variability of the total data. This allows identification of any outlier samples, or similar expression patterns between sample groups. The multidimensional scaling (MDS) plot in Figure 3.5 shows that the replicates of each Stage are similar to one another, indicating that there is not a large difference in expression between replicates.

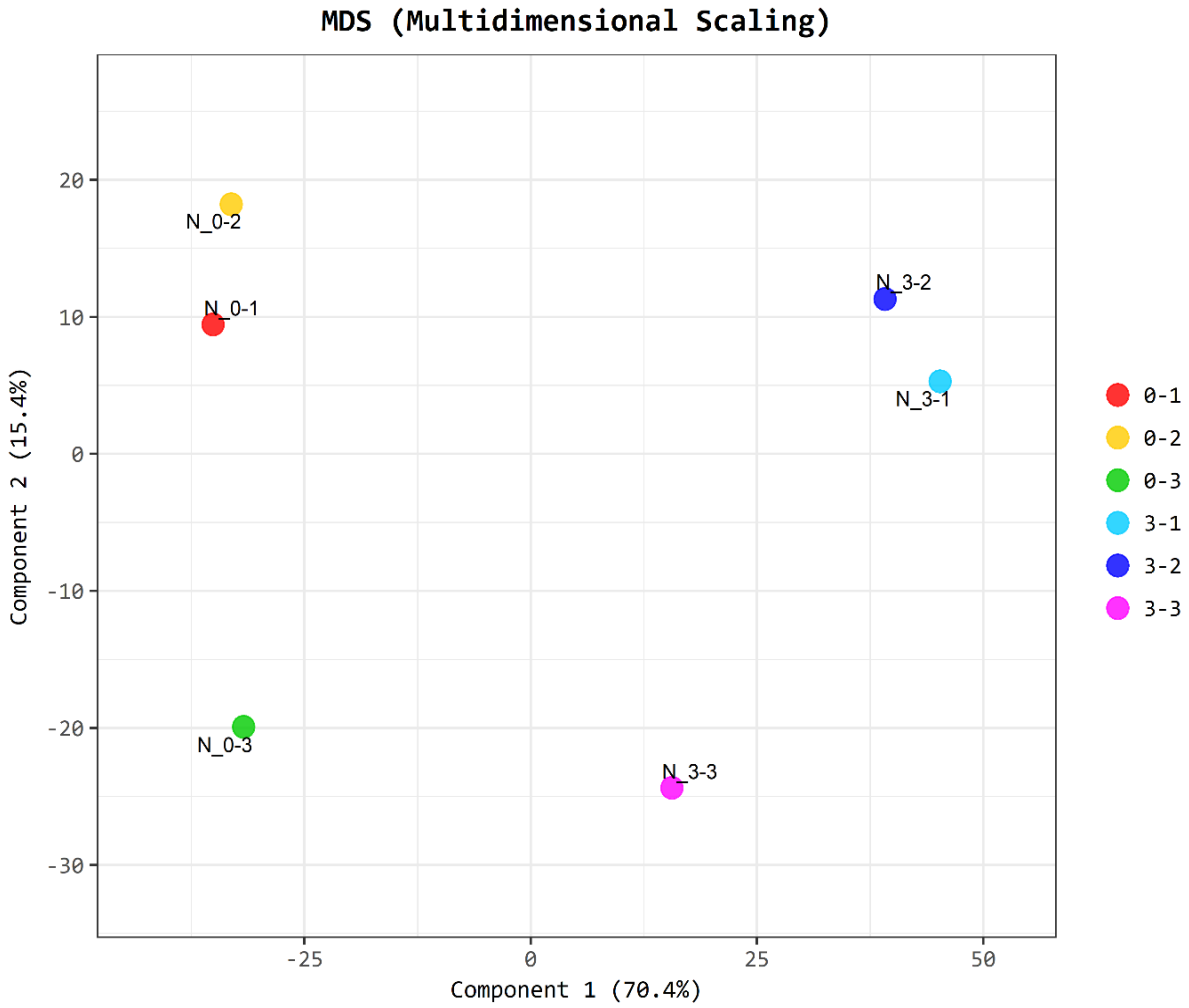


Figure 3.5 Multidimensional Scaling analysis plot

N_0-1= Stage 0 - Replicate 1; N_3-1 = Stage 3 - Replicate 1. Component 1 shows the differential expression between Stage 0 and Stage 3 with each replicate. Whereas component 2 shows the number of genes that were significantly differentially expressed.

The third biological replicate (*An. funestus* males with the genitalia rotated to Stage 3, biological replicate 3 (named N_3-3)) and *An. funestus* males with the genitalia rotated to Stage 0, biological replicate 3 (named N_0-3) showed more variation in transcript abundance in comparison with biological replicate 1 and 2 (Figure 3.5). This variance in transcript abundance

still showed a high difference in expression between the two biological stages, even though transcript abundance was dissimilar to the other two biological replicates. A large variance (in terms of Pearson's coefficient) in expression between the *An. funestus* males with the genitalia rotated to Stage 3 compared to the *An. funestus* males with the genitalia rotated to Stage 0 is clearly shown (Figure 3.5).

3.4.4 Differentially expressed gene analysis

A total of 11,663 transcripts were statistically analysed to determine differentially expressed transcripts between *An. funestus* males with the genitalia rotated to Stage 0 and Stage 3. A heat map of the hierarchical clustering analysis indicated the similarity between clusters of transcript abundance, from *An. funestus* males with the genitalia rotated to Stage 0 and Stage 3, after normalisation (using z-score) (Figure 3.6). The heat map in Figure 3.6 demonstrates that the three biological replicates of *An. funestus* males with the genitalia rotated to Stage 0 cluster together and similarly so for *An. funestus* males with the genitalia rotated to Stage 3. Overall, transcript abundance between the three biological replicates within a group (*An. funestus* males with the genitalia rotated to Stage 0 and Stage 3) are similar.

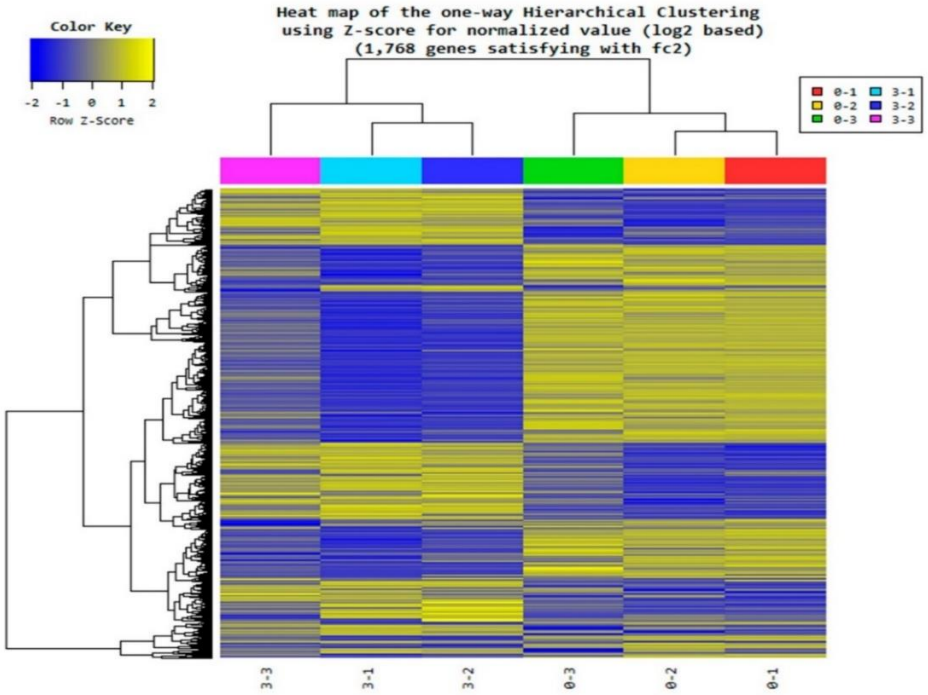


Figure 3.6 Heatmap of one-way hierarchical clustering analysis

0-1 = Stage 0 - Replicate 1; 3-1 = Stage 3 - Replicate 1. Using the Euclidean method and complete linkage, cluster of similarity of genes and samples by expression level (normalized value).

Differentially expressed gene analysis for each biological replicate revealed that a total of 1,768 transcripts were overexpressed, with a $2 \leq |FC| \leq -2$, in one or more of the biological replicates between *An. funestus* males with the genitalia rotated to Stage 3 compared to Stage 0. Many of these 1,768 transcripts are duplicates of one another as more than one replicate could be over expressed for the same transcript. There were 510, 490, 267 genes that were up regulated for all the respective replicates (biological replicate 1, 2 and 3) with a $|FC| \geq 2$ in *An. funestus* males with the genitalia rotated to Stage 3 compared to *An. funestus* males with the genitalia rotated to Stage 0. Similarly, there were 851, 683, 508 genes that were down regulated for all the respective replicates with a $|FC| \leq -2$ in *An. funestus* males with the genitalia rotated to Stage 3 compared to *An. funestus* males with the genitalia rotated to Stage 0. This showed that there is a high degree of difference in expression between the two stages (Figure 3.7). As expected, biological replicate 3 showed slightly fewer differentially expressed transcripts than the other 2 biological replicates, due to low transcript abundance (mentioned in Section 3.4.3).

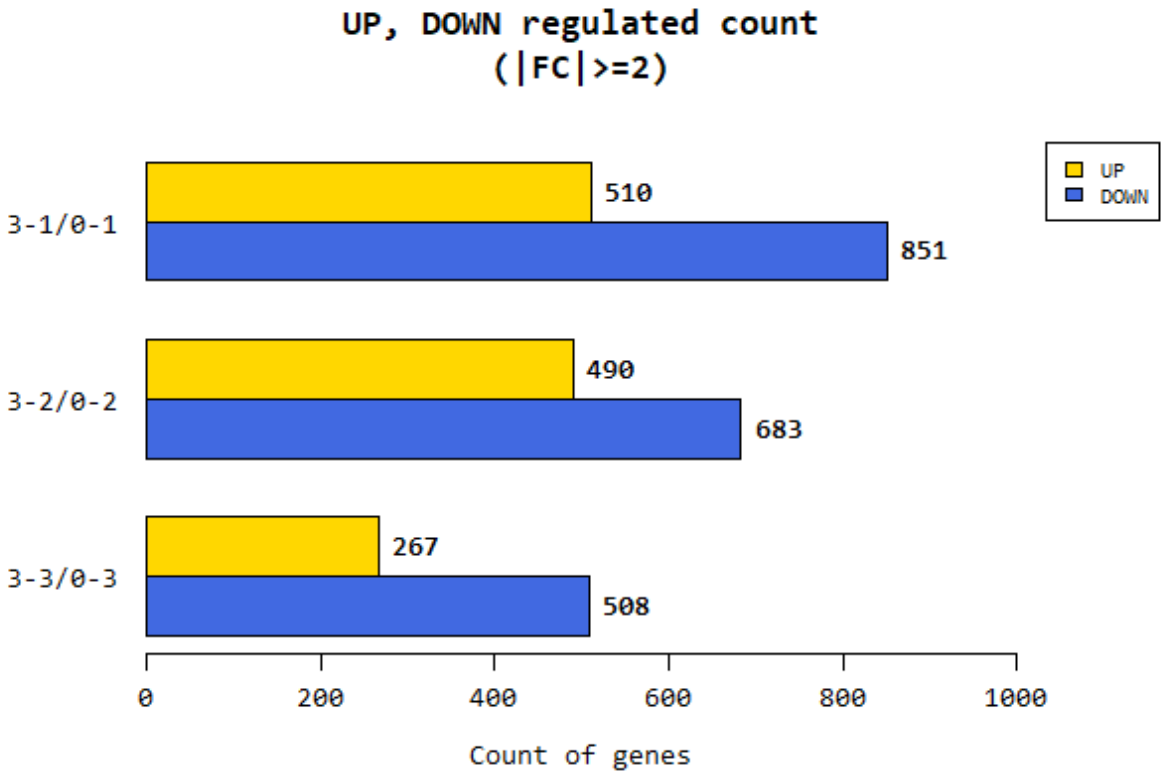


Figure 3.7 Significant count by $2 \leq |FC| \leq -2$ of Stage 3 compared to Stage 0

3-1 = Stage 3 - Replicate 1; 0-1 = Stage 0 - Replicate 1. Yellow bars indicate transcripts that were up regulated in Stage 3 compared to Stage 0 and blue bars indicate transcripts that were down regulated in Stage 3 compared to Stage 0.

Scatter plots have been set up to show the expression levels between comparison pairs (Figure 3.8). The scatter plots visualize the differentially expressed transcripts in Stage 3 compared to the control group (Stage 0) of each of the three biological replicates. As can be seen in figure 3.8, there are hundreds of transcripts that are significantly differentially expressed by a fold change greater or less than 2, depicted as red dots. The greyed-out dots represent all of the transcripts that were mapped to the reference gene that were not differentially expressed in Stage 3 compared to Stage 0.

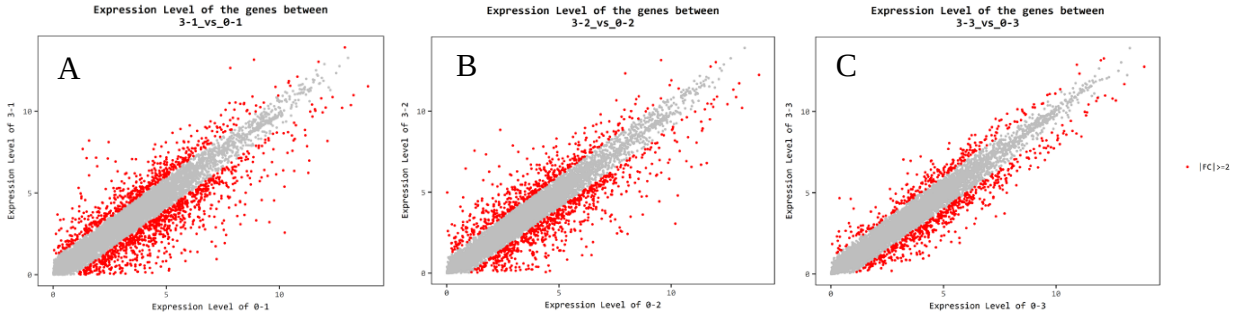


Figure 3.8 Expression level of genes

3-1 = Stage 3 - Replicate 1; 0-1 = Stage 0 - Replicate 1. A) X-axis is the expression level of 0-1 and the Y-axis is the expression level 0-3. B) X-axis is the expression level of 0-2 and the Y-axis is the expression level 3-2. C) X-axis is the expression level of 0-3 and the Y-axis is the expression level 3-3. The scatter plot indicates all of the transcripts that have a $2 \leq |FC| \leq -2$ (significant level of expression) in red (either under or over expressed) and all of the transcripts have a $2 \geq |FC| \geq -2$ in grey (not significant).

Expression volume is defined as the geometric mean of two group's expression level. A volume plot was drawn to confirm that the genes that showed higher expression difference compared to the control, according to expression volume (Figure 3.9). This means that even if the fold change might be different by 2-fold, the genes with a higher volume might be more credible than the genes with the highest fold change. Figure 3.9 shows that most of the transcripts that had a $2 \leq |FC| \leq -2$ also had a relatively high volume. Indicating that they might be differentially expressed even though they are not highly differentially expressed but does satisfy the differentially expressed fold change of $2 \leq |FC| \leq -2$.

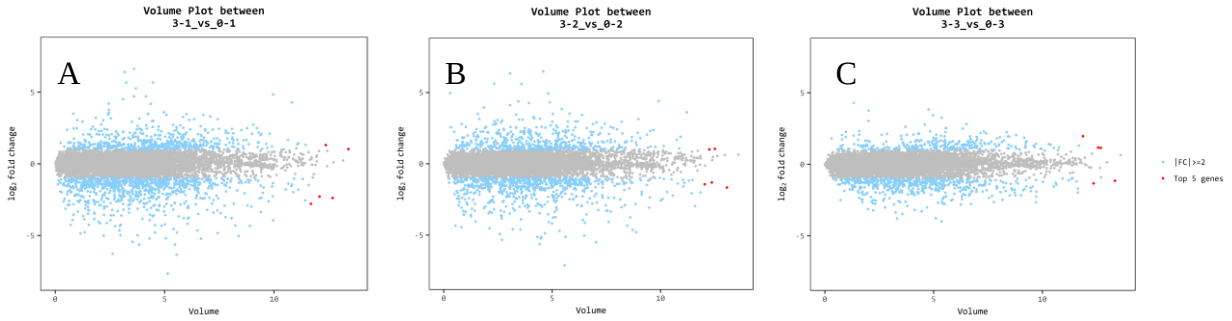


Figure 3.9 Volume plot of expression

3-1 = Stage 3 - Replicate 1; 0-1 = Stage 0 - Replicate 1. A) X-axis is the expression level of 0-1 and the Y-axis is the expression level 3-1. B) X-axis is the expression level of 0-2 and the Y-axis is the expression level 3-2. C) X-axis is the expression level of 0-3 and the Y-axis is the expression level 3-3. Volume of each transcript (x-axis) relative to the $\log_2$ fold change (y-axis). Red dot = The top 5 genes by volume which satisfies $2 \leq |FC| \leq -2$.

The average fold change ($2 \leq |FC| \leq -2$) over all three biological replicates, identified 1,038 differentially expressed genes (656 transcripts were down-regulated; 382 transcripts were up-regulated) between *An. funestus* males with the genitalia rotated to Stage 3 compared to Stage 0 (Appendix 8.4). A total of 10,629 transcript were not significantly expressed ($2 \leq |FC| \leq -2$) in either of the *An. funestus* males with the genitalia rotated to Stage 3 compared to Stage 0.

There was a total of 35 transcripts that were significantly up or down regulated in Stage 3 compared to Stage 0 ($2 \leq |FC| \leq -2$) that were annotated to various cuticular families (Table 3.5). Of the 35 annotations that up or down regulated, 32 were down regulated, in Stage 3 compared to Stage 0, by a $|FC| \leq -2$ and only 3 were up regulated by a $|FC| \geq 2$, in Stage 3 compared to Stage 0 (Table 3.5). There are 10 transcripts were down regulated in Stage 3 compared to Stage 0 ($|FC| \leq -2$) that were annotated to the cuticular RR-1 family. There were two transcripts that were up regulated ($|FC| \geq 2$) in Stage 3 compared to Stage 0, that were part of cuticular RR-1 family. One of the transcripts, in the cuticular RR-1 family, that was down regulated is the AFUN020892 (CPRR1) transcript, which is later used for validation in the qPCR study (Chapter 4). There are 10 transcripts that were down regulated in Stage 3 compared to Stage 0 ($|FC| \leq -2$) that was part of the cuticular RR-2 family. Three transcripts from fifty-one amino acid family, two from the CPLCG family, and one transcript each, from the CPFL,CPCFC and CPLCA families, were also down regulated by $|FC| \leq -2$ in Stage 3 compared to Stage 0 (Table 3.5). There were four transcripts that have not been classified into a family that were down regulated by $|FC| \leq -2$ in Stage 3 compared to Stage 0 (Table 3.5).

Table 3.5 Number of cuticular transcripts that were differentially expressed and their families

Cuticular family	Number of transcripts down regulated in Stage 3 compared to Stage 0	Number of transcripts up regulated in Stage 3 compared to Stage 0	Number of transcripts that were differentially expressed in Stage 3 compared to Stage 0
Cuticular RR-1 family	10*	2	12
Cuticular RR-2 family	10	0	10
fifty-one aa family	3	0	3
CPLCG family	2	1	3
CPFL family	1	0	1
CPCFC family	1	0	1
CPLCA family	1	0	1
Unclassified family	4	0	4
Total	32	3	35

*Includes AFUN020892 (CPRR1) transcript, later used for validation in chapter 4.

There were six transcripts that were selected for transcript expression validation (in Chapter 4) that had a significant differential expression between Stage 3 compared to Stage 0 according to their fold-change across at least two replicates. The six transcripts, shown in Figure 3.10 and Table 3.6, were selected for their significant and stable differential expression between Stage 3 compared to Stage 0 according to their fold-change across all three replicates from the RNA-seq data.

The CPRR1 (AFUN020892) transcript was selected for further analysis because of the working hypothesis that there is a softening of the endocuticle for the genitalia rotation to take place. Without this endocuticle softening, the muscle might not be able to rotate the 8th abdominal section, or the cuticle will break under rotation. The results are indicative of endocuticle softening as there is a down-regulation in Stage 3 compared to Stage 0 of a fold-change of -11.39.

The UNN1 (AFUN000965), UNN2 (AFUN000144) and UNN 3 (AFUN005543) transcripts were selected as they were unannotated (UNN) and consistently differentially expressed across all three of the replicates between Stage 3 compared to Stage 0 (Figure 3.10, Table 3.6). The UNN1 (AFUN000965) transcript was up-regulated in Stage 3 compared Stage 0 by an average fold-change of 2.403, across the three biological replicates. The UNN2 (AFUN000144) transcript was down-regulated in Stage 3 compared to Stage 0 by an average fold-change of -2.50, across the

three biological replicates. Whereas UNN3 was up-regulated in Stage 3 compared to Stage 0 by an average fold-change of 3.43, across the three biological replicates.

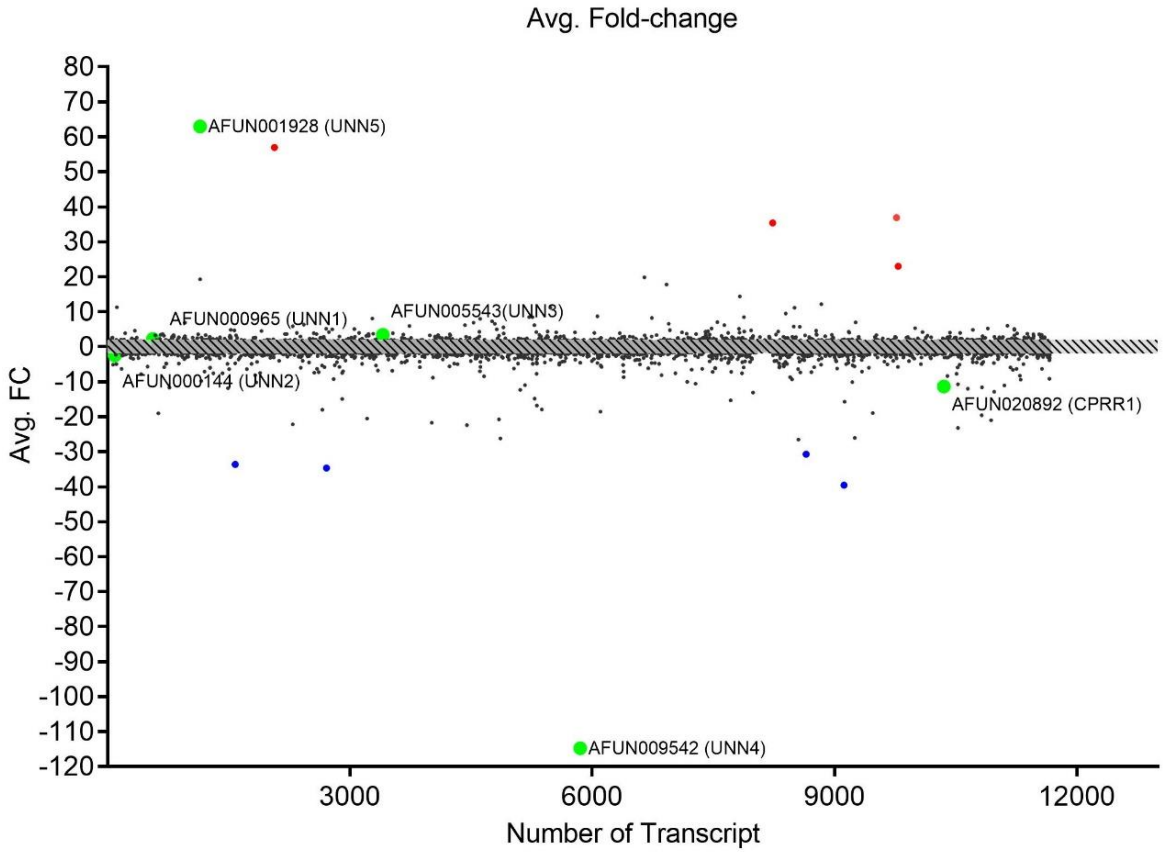


Figure 3.10 Scatter plot of Avg. fold-change of differentially expressed genes

Y-axis depicts the average log2 fold-change, x-axis depicts the number of transcripts. Blue dots represent the top five under expressed transcripts in Stage 3 compared to Stage 0, red dots depict the top five over expressed transcripts in Stage 3 compared to Stage 0. The green dots represent the transcripts that were selected for confirmation in Chapter 4. The grey bar indicates transcripts that did not exceed the fold-change cut of $2 \leq |FC| \leq -2$.

Only four transcripts that had annotations from the top 5 up regulated and top 5 down regulated transcripts, as well as from the four transcripts selected for validation (Table 3.5). Three transcripts were annotated and were down regulated, in Stage 3 compared to Stage 0, by a $|FC| \leq -2$ (AFUN019559; AFUN002627; AFUN020892 (CPRR1)) (Table 3.6). Only one annotated transcript was up regulated, in stage 3 compared to Stage 0, by a $|FC| \geq 2$ (AFUN020303) (Table 3.6). The AFUN019559 was annotated to polyubiquitin-B [Source: Projected from *Anopheles gambiae* (AGAP001970) VB Community Annotation] and was down regulated in Stage 3 compared to Stage 0 by a $|FC| \text{ of} \leq -39.52$. The AFUN002627 transcript was annotated to Flavin-containing monooxygenase FMO GS-OX-like [Source: Projected from *Anopheles gambiae* (AGAP010399) VB Community Annotation] and was down regulated in Stage 3 compared to

Stage 0 by a |FC| of ≤ -33.61. The AFUN0020892 (CPPR1) transcript was annotated to the cuticular protein RR-1 family [Source: Projected from *Anopheles gambiae* (AGAP005999) VB Community Annotation] and was down regulated in Stage 3 compared to Stage 0 by a |FC| of ≤ -11.386. The transcript AFUN020303 was up regulated by a |FC| of 23.00 in Stage 3 compared to Stage 0 and was annotated to the heme peroxidase [Source: Projected from *Anopheles gambiae* (AGAP003714) VB Community Annotation] transcript.

The UNN4 (AFUN009542) and UNN5 (AFUN001928) were selected as they had the lowest and the highest average |FC|, across the three biological replicates, respectively (Figure 3.10, Table 3.6). The UNN4 (AFUN009452) transcript was down regulated in Stage 3 compared to Stage 0 by an average |FC| of -114.72, across all three biological replicates. The UNN5 (AFUN001928) transcript was up regulated in Stage 3 compared to Stage 0 by an average |FC| of 62.94, across all three biological replicates.

Table 3.6 Selected differentially expressed transcripts from the RNA-Seq data between Stage 0 and Stage 3 of the male *An. funestus* mosquito

Gene I.D.	Transcript Name	Annotation	3.1/0.1	3.2/0.2	3.3/0.3	Average Fold change (log2)
AFUN009542	UNN4	N/A	-200.84	-139.81	-3.48	-114.72
AFUN019559	-	Polyubiquitin-B*	-80,76	-34,79	-3,02	-39,52
AFUN004371	-	N/A	-49,90	-49,53	-4,49	-34,64
AFUN002627	-	Flavin-containing monooxygenase FMO GS-OX-like **	-77,28	-18,15	-5,41	-33,61
AFUN018979	-	N/A	-49,97	-35,11	-7,01	-30,70
AFUN020892	CPRR1	Cuticular protein RR-1 family***	-17.72	-12.65	-3.37	-11.39
AFUN000144	UNN2	N/A	-2.11	-3.21	-2.16	-2.50
AFUN000965	UNN1	N/A	2.36	2.68	2.16	2.403
AFUN005543	UNN3	N/A	4.22	2.17	3.87	3.425
AFUN020303	-	Heme peroxidase ****	38,53	20,90	9,57	23,00
AFUN018474	-	N/A	50,94	48,19	7,02	35,38
AFUN020285	-	N/A	50,67	49,28	10,76	36,90
AFUN003379	-	N/A	84,46	80,84	5,74	57,01
AFUN001928	UNN5	N/A	98.81	88.84	1.15	62.94

3.1 = Stage 3 - Replicate 1; 0.1 = Stage 0 - Replicate 1.
Green = Samples that were selected for qPCR validation in chapter 4 (including the highest down and up regulated transcripts i.e., AFUN009542 and AFUN001928).
Blue = Top 5 Down regulated transcripts
Red = Top 5 up regulated transcripts
*[Source: Projected from *Anopheles gambiae* (AGAP001970) VB Community Annotation]
**[Source: Projected from *Anopheles gambiae* (AGAP010399) VB Community Annotation]
***[Source: Projected from *Anopheles gambiae* (AGAP005999) VB Community Annotation]
****[Source: Projected from *Anopheles gambiae* (AGAP003714) VB Community Annotation]

3.5 Discussion

The adult male mosquitoes' main biological role is currently assumed to be that of mating (Clements, 1999). Disturbance of the mating process directly impacts the population reproductive success. The lack of knowledge on the molecular mechanisms involved in male sexual maturation is, however, preventing research expanding into this field. The use of functional genomics was investigated to initiate studies in understanding male sexual maturation in a major African malaria vector, *Anopheles funestus*.

RNA-Sequencing was used to determine the transcriptome of male *An. funestus* mosquitoes which have their genitalia rotated (Stage 3) compared to that of males who have not had their genitalia rotated (Stage 0). Quantifying the transcripts and comparing the difference in expression of the transcripts of Stage 3 compared to Stage 0 determined which transcripts are differentially expressed between the two developmental points. Transcripts that were differentially expressed could be involved with the process of genitalia rotation or another process involved during the sexual maturation of male *An. funestus* mosquitoes.

There were 8.89% transcripts of the total transcripts that were sequenced that were differentially expressed in Stage 3 compared Stage 0 with a $2 \leq |FC| \leq -2$ (1,038 transcripts, 382 transcripts were up regulated, and 656 transcripts were down-regulated with a $2 \leq |FC| \leq -2$) . This indicates that there are many transcripts that are possibly involved in the development of the male anopheline between 0 h (Stage 0) and 16 h (Stage 3).

Of the 11,663 transcripts that were sequenced and aligned to the reference genome, 5,803 (49.75%) were unannotated and 5,860 (50.25%) were annotated, regardless of differential expression. The transcripts that were down regulated in Stage 3 compared to Stage 0 was 656, 365 of these were unannotated compared to the 291 that were annotated. Of the transcripts that were up regulated in Stage 3 compared to Stage 0 (382), 194 were unannotated compared to the 188 annotated transcripts. This indicates that between the developmental time points of 0 h (Stage 0) and 16 h (Stage 3) of the male *An. funestus* mosquito, that 55.64% and 50.79% of the down and up regulated transcripts between these two developmental time points were unannotated, respectively. This shows that the *An. funestus* genome still needs to be fully annotated if we are to understand this important malaria vector.

Many of these differentially expressed transcripts, either annotated or unannotated, will also be differentially expressed in the female between the two time points (0 h compared to 16 h). These transcripts will include transcripts that are involved in the development of the adult mosquito, regardless of sex, such as cuticle development and metabolic processes. The transcripts that are differentially expressed, between Stage 3 as compared to Stage 0, and are shared with the female transcriptome will need to be determined in order to differentiate transcripts that are involved with the male *An. funestus* sexual development only.

There were three unannotated transcripts that had highly and consistent differential gene expression, over the 3 biological replicates, between Stage 0 and Stage 3 of the male *An. funestus* mosquito. These were UNN1 (AFUN000965), UNN2 (AFUN000144) and UNN3 (AFUN005543). Due to their consistency of differential gene expression between Stage 0 and Stage 3, over the three biological replicates, it was assumed that they play a functional role in the sexual maturity of the male anopheline mosquito. However, their biological importance or role in sexual maturation, such as their involvement in the male genitalia rotation, antennal fibrillae erection, accessory gland maturation and spermatogenesis will have to be investigated (Roth, 1948; Clements, 1999). The transcripts UNN4 (AFUN009542) and UNN5 (AFUN001928) had the lowest and the highest average differential expression respectively, across the three replicates, in Stage 3 as compared to Stage 0. The UNN4 (AFUN009542) transcript was down regulated by an average of -114.72 |FC| in Stage 3 as compared to Stage 0 across all three replicates, whilst the UNN5 (AFUN001928) transcript was up regulated by an average of 62.94 |FC| in Stage 3 when compared to Stage 0, across all three replicates. Due to the transcripts being unannotated, it is difficult to predict the putative role of these transcripts. The differential expression of the UNN4 (AFUN009542) and the UNN5 (AFUN001928) transcripts should be confirmed with real-time quantitative PCR in order to ensure that these transcripts were indeed differentially expressed.

Thousands of transcripts from the RNA-seq data (11,663) were expected to be differentially expressed. However, some of the non-differentially expressed transcripts could be involved in female development and maturation and are thus not specifically expressed in the male which had their genitalia rotated at Stage 3 (which were 16 h old) compared to males which had no genitalia rotation (0 h old). Other processes that these non-differentially transcripts could be involved in are embryo development, larval development, pupal development, emergence, wing development and immunity development, amongst others. The 1,038 transcripts that were differentially

expressed ($2 \leq |FC| \leq -2$), are assumed to be involved in the development of the male *An. funestus* development between the time points of 0 h (Stage 0) and 16 h (Stage 3). Although not all of these transcripts are involved with the sexual maturation of the male, there is a strong likelihood that some of these transcripts are involved with the sexual maturation of the male *An. funestus* mosquito. The RNA-Seq results showed that there is a large difference in expression between Stage 0 and Stage 3. There was an average of 422.3 genes that were up regulated with a $|FC| \geq 2$ and 680.6 genes that were down regulated with a $|FC| \leq -2$.

Of the top five up and down regulated transcripts in Stage 3 compared to Stage 0, there were only three that had annotations. The AFUN019559 was down regulated by a |FC| of – 39.52 in Stage 3 compared to Stage 0 and was annotated to polyubiquitin-B. Polyubiquitin plays a functional role in mediating the degradation by the 26S proteasome of the intracellular ATP-dependant protein (Hershko, 1991; Rechsteiner, 1991; Callis *et al.*, 1995; Viestra, 1996; Varshavsky, 2011; Mehrtash, 2019). It has been proposed that ubiquitin conjugation plays a role in spermatogenesis (Mezquita *et al*., 1997).

Transcript AFUN002627 was also down regulated in Stage 3 compared to Stage 0 by a |FC| of -33.61 and was annotated to flavin-containing monooxygenase FMO GS-OX-like. According to Uniprot (https://www.uniprot.org/uniprot/Q9FWW6) flavin-containing monooxygenase FMO GS-OX-like is an enzyme that catalyses the conversion of methylthioalkyl glucosinolates of any chain length into methylsulfinylalkyl glucosinolates. The function in *An. funestus*, however, needs to be clarified.

There was only one transcript (AFUN020303) that was annotated and in the top five of the up regulated transcripts in Stage 3 compared to Stage 0. This AFUN020303 transcript was up regulated by |FC| of 23.00 and was annotated to heme peroxidase. Heme peroxidases catalyze the hydrogen peroxide-dependent oxidation of a variety of different substrates. It was found that heme peroxidase plays a role in the development of *Plasmodium* in the midgut of the *Anopheles* mosquito (Kajla *et al.*, 2017). This could mean that the AFUN020303 transcript is a transcript that is expressed in both females and males as the males do not get infected by the *Plasmodium* parasite. Although, the heme peroxidase is widely distributed in biological systems and the AFUN020303 transcript is possibly expressing the heme peroxidase for general oxidation of

substance and most likely is not involved in the development of the male *An. funestus* mosquito.

There were 87 cuticle annotations that were found within the dataset, 32 of these were down regulated with a $|FC| \leq -2$, 3 transcripts were over expressed with a $|FC|$ of ≥ 2. There were 52 cuticular transcripts that had a $|FC|$ of between ≤ 2 and ≥ -2 and were not considered differentially expressed. There were 12 transcripts, out of the cuticular transcripts, that had annotated to the cuticular RR-1 family that were differentially expressed. Ten of these transcripts were down regulated in Stage 3 compared to Stage 0. One of these transcripts was the AFUN020892 (CPRR1) transcript. There were ten transcripts that were annotated to the cuticular protein family RR-2 (AFUN002143, AFUN019930, AFUN007919, AFUN021427, AFUN021421, AFUN21428, AFUN005982, AFUN004125, AFUN020921 and AFUN00798), they were all down regulated in Stage 3 compared to Stage 0. It was expected to find many cuticular transcripts that were differentially expressed. However, it was thought that these transcripts would be up regulated in Stage 3 as compared to Stage 0, in order for the development of the arthropods' exoskeleton.

The annotation of the AFUN020892 transcript showed that it is from the cuticular protein RR-1 family (endocuticle). The CPRR1 gene family belongs to the largest family of chitin genes in insects (Willis, 2010 and Yang *et al.*, 2017). The AFUN020892 transcript is 444 bp long which translates into a 109 amino acid peptide. The transcript transcribes the cuticular RR protein which belongs to the cuticular protein that has its RR consensus attached to the 68-aa conserved sequence.

It has been hypothesized that there is an up regulation in cuticular transcripts post ecdysis to ensure a strengthening and hardening of cuticles as arthropods age (Wang *et al.*, 2009). The dataset only showed 3 transcripts that were annotated to cuticular proteins and up regulated in stage 3 as compared to Stage 0. The AFUN020892 annotation was amongst 87 cuticular annotations in the dataset, however, focus was kept on the AFUN020892 transcript because it had an average $|FC|$ of -11.386, which was one of the highest consistent $|FC|$ differences, as well as its contradiction to the hypothesis by Yang *et al*. (2017) of up regulation of cuticular transcripts post ecdysis.

Due to the cuticular RR-1 genes, which are involved in the expression of the endocuticle, down regulation could show that there is softening of the endocuticle of the 7th and 8th abdominal segments to allow for genitalia rotation in the later stages of the inversion process. This annotation of the AFUN020892 transcript and the down regulation regarding its function indicates that the transcript could be involved with the genitalia rotation process.

There is a possibility that at Stage 3 the *An. funestus* male is fully developed in regard to its cuticle development. This could be the reason why there is a down regulation of the CPRR1 (AFUN020892) transcript in Stage 3 as compared to Stage 0. However, at Stage 3 the exoskeleton of the male anopheline mosquito is still quite soft and sensitive to insecticides relative to 36-hour old male anophelines (considered fully matured) (Mahmood and Reisen, 1982; Chambers and Klowden, 2001). This could be attributed to the mosquitoes in Stage 3 not having fully developed yet and it would be expected to see an increase in cuticular transcript expression if the mosquito is still aging.

To further understand the role of CPRR1, further validation and testing needs to be performed to elucidate the exact role and function of CPRR1 (AFUN020892) concerning the sexual maturity of the male *An. funestus* mosquito.

This study has characterized some of the expression patterns of the sexual development of the male *An. funestus* mosquito.

4 Confirmation of transcript abundance of selected genes

4.1 Introduction

The RNA-seq procedure, performed in the previous study, was an investigative tool for this study, where differentially expressed transcripts were isolated that could putatively be involved in the *Anopheles funestus* males' sexual development. This requires the confirmation that the genes of interest, such as the CPRR1 (AFUN020892), UNN1 (AFUN000965), UNN2 (AFUN000144), UNN3 (AFUN005543), UNN4 (AFUN009542) and UNN5 (AFUN001982) are differentially expressed at the different genitalia rotation stages. This is done by using quantitative real-time quantitative PCR (qPCR). There are several reasons why one would confirm the RNA-seq data, namely, technical reproducibility (are these transcripts truly differentially expressed in these samples using different assays), biological reproducibility (are these transcripts generally differentially expressed in other samples to confirm differential expression in the biological system), and significance (do these transcriptional changes represent phenotypic difference) (Van Bakel and Hughes, 2009). Confirming the transcripts' differential expression through the use of an independent assay provides technical reproducibility. If a separate cohort of samples were used for the independent assay, this will provide biological reproducibility as well (Van Bakel and Hughes, 2009). It has been found that there is an excellent correlation between qPCR results and RNA-seq results (Asmann *et al.,* 2009; Griffith *et al.,* 2010; Wu *et al.,* 2014; Shi and He, 2014).

Quantitative Real-Time PCR (qPCR) is a technique used to monitor the amplification of a targeted DNA or cDNA in real-time (Higuchi *et al.,* 1992). By the use of a fluorophore the amplification of the targeted DNA (or cDNA) molecule can be measured after each cycle by detecting the amount of fluorescence. This means that the rate of amplification of that specific DNA molecule can be measured after each cycle. If cDNA was targeted, then the mRNA copy number of several samples can be measured. The mRNA copy number of each sample will be

proportional to that of the expression of the gene that encoded for that particular mRNA. This means that by comparing the mRNA copy number of a single gene between two different samples (i.e., two developmental stages) the differential expression can be determined of that gene between the two samples (or two developmental stages) (Higuchi *et al.,* 1992). Reference genes (RG) are used to normalize any differences between the amount and quality of the RNA between the different samples (Schefe *et al.,* 2006).

The reference gene (RG) is generally a house-keeping gene that has a stable expression throughout the two experimental conditions. To ensure that the change in target genes expression is due to biological effect (i.e., developmental stage) and not the difference in the cDNA quantity, RGs are used (Schefe *et al.,* 2006). It is necessary to analyze the RG in parallel with the samples in order to obtain an accurate ratio between the RG's and the samples. A range of potential RG's needs to be evaluated to verify the most stable RG's between the two experimental conditions (Vandesompele *et al.,* 2002; Bustin *et al.,* 2009; Chervoneva *et al.,* 2010; Derveaux *et al.,* 2010; Hellemans and Vandesompele, 2014; Shakeel *et al.,* 2018).

Real-time quantitative PCR, like conventional PCR, has an exponential phase, whereby the amount of PCR product approximately doubles with each reaction cycle. The reaction components (e.g., dNTPs) are limiting and will lead to the exponential phase reaching a plateau, and there will be no further doubling of the PCR product. There are two different chemistries of qPCR: 1) fluorescent dyes that are non-specific are used to intercalate with the dsDNA, and 2) probes that are sequence specific and labeled with a fluorescent reporter, which only permits detection after hybridization of the probe to its complimentary sequence. As can be seen in Figure 4.1, during the initial phase of the PCR, the fluorescence is too low to be detected as the intercalated product is below the threshold of detection. Once the product reaches a threshold, the number of amplification cycles whereby the fluorescence can be detected, is referred to as the quantification cycle or the C_q value. The C_q is determined by the amount of starting material.

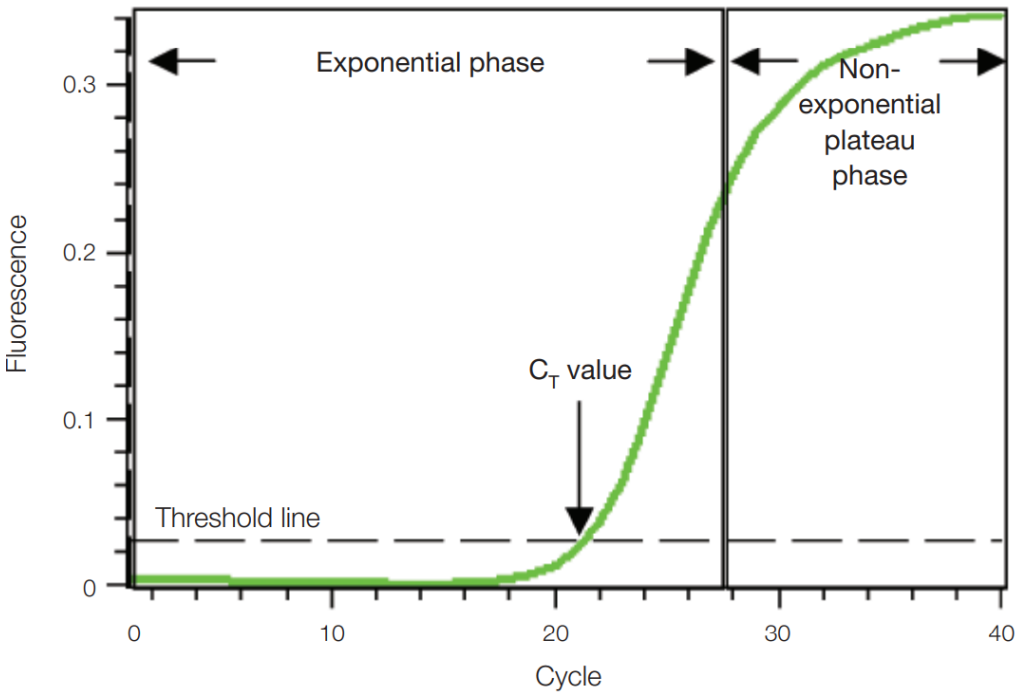


Figure 4.1 Theoretical amplification curve of a single PCR product

The C_q value is indicative of the starting material allowing for the quantification of a particular transcript (Taken directly from BioRad Real-Time PCR applications guide).

There are many different approaches for data analysis in order to determine relative quantification, such as, the Pfaffl method (Pfaffl, 2001), and the Livak method ($2^{-\Delta\Delta Ct}$) (Livak and Schmittgen, 2001). With the Livak method the amplification efficiency of the reference and target genes needs to be almost the same (as close as to 100% as possible), so the Pfaffl method is more commonly used. The Pfaffl method uses the formula, below, to calculate relative expression ratio between the target genes and the reference gene.

$$\text{Ratio} = (E_{target})^{\Delta Cq,\ target\ (calibrator - test)} / (E_{ref})^{\Delta Cq,\ ref\ (calibrator - test)}$$

Where:

E = Efficiency (the slope of the standard curve is calculated)

E_{target} and E_{ref} = are the target and reference genes amplification efficiencies, respectively.

ΔC_q target (calibrator – test) = is the C_q of the target gene in the control sample minus C_q of the target gene in the test sample.

ΔC_q ref (calibrator – test) = is the C_q of the reference gene in the control sample minus C_q of the reference gene in the test sample (Taken directly from BioRad Real-Time PCR applications guide).

Despite the fact that RNA-seq technology has become the cornerstone for gene expression analysis, qPCR is still the recommended technique for validation and confirmation (Shi *et al.*, 2014). Issues such as RNA extraction, cDNA synthesis, RNA quantification, the different qPCR chemistries, reference gene suitability and stability and data analysis are extremely important in order to maintain reproducibility of results (Bustin *et al.*, 2009). A set of guidelines has been published, Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE), in order to maintain a minimum set of requirements for qPCR data publication (Bustin *et al.*, 2009; Nolan *et al.*, 2011; Bustin, 2017; Bustin and Huggett, 2017; Bustin and Nolan, 2017).

The use of qPCR has been used in several other studies on anophelines in order to validate the RNA-seq and make sure that the direction of the fold change is conserved (Bonizzoni *et al.*, 2012; Bonizzoni *et al.*, 2015; De Marco *et al.*, 2017; Mongkol *et al.*, 2018; Bryant *et al.*, 2019). These studies were all able to strongly correlate their RNA-seq data with their qPCR results. This shows that qPCR can be effectively used to validate the RNA-seq data. However, to my knowledge this is the first study that has validated *An. funestus* RNA-seq data with the use of qPCR, at the time of this study.

4.2 Rationale

In Chapter 3, the CPRR1 (AFUN020892), UNN1 (AFUN000965), UNN2 (AFUN000144), UNN3 (AFUN005543), UNN4 (AFUN009542) and the UNN5 (FUN001928) transcripts were selected due to consistent differential expression between Stage 0 and Stage 3 in the male *Anopheles funestus* mosquito through RNA-sequencing. The CPRR1 (AFUN020892) transcript was putatively chosen for its possible role in the male maturation process. The three unannotated transcripts, UNN1 (AFUN000965), UNN2 (AFUN000144), UNN3 (AFUN005543), were chosen because of their consistency of differential expression across the three biological replicates. The two unannotated transcripts, UNN4 (AFUN009542) and UNN5 (FUN001928), were chosen as they had the lowest and the highest average differential expression between the two stages, respectively.

Quantitative PCR were carried out on these six transcripts and standard curves were produced. This will confirm whether the results of the RNA-Sequencing data, which showed the differential gene expression of specific genes is correct.

4.3 Material and methods

4.3.1 Biological material, RNA extraction and preparation

Anopheles funestus (FUMOZ) males from the insectary (Botha DeMeillon insectary at the NICD, NHLS, Sandringham, Johannesburg, South Africa) were used.

Total RNA was extracted from 10 Stage 0 *An. funestus* males and analysed as described in the RNA-sequencing objective above (Chapter 3.3.1). The RNA extraction procedure was repeated with 10 Stage 3 *An. funestus* mosquitoes. Three biological repeats were performed for both time points. The RNA was electrophoresed on a native agarose gel in order to determine the integrity of the RNA, as described in Chapter 3 (see Section 3.3.1). The RNA was measured with a NanoDrop 2000 Spectrophotometer as described in Chapter 3 (see Section 3.3.1) in order to determine the RNA concentration.

4.3.2 Primer design and reference gene selection

Primers were designed by using the Invitrogens free online primer design tool, OligoPerfect™ Designer based on the sequence information of *An. funestus* transcripts. The transcripts selected for validation were the AFUN020892 (a CPRR-1 endocuticle family transcript) and five unannotated differentially expressed transcripts (UNN1 [AFUN000965], UNN2 [AFUN000144], UNN3 [AFUN005543], UNN4 [AFUN009542] and UNN5 [AFUN001928]) (www.VectorBase.org unannotated accessed on 3 June 2020). The primer sequences for these transcripts can be found in table 4.3. The relative expression of the selected transcripts was normalized with a selected reference gene (RG) using the Bio-Rad CFX Maestro™ Software version 4.0.2325.0418 (Vandesompele *et al.*, 2002). This relative expression was used to confirm the results of the sequencing results. The reference gene was selected from five possible house-keeping genes (Nardini *et al.*, 2012).

Primers were designed so that the amplicon size would be between 75 - 200 bp in size, all secondary structures such as hairpin loops were avoided, the primers' GC content was between 50 – 60 % and their melting temperatures were between 50°C - 65°C. WhiteSci's qPCR tool (https://eu.idtdna.com/calc/analyzer,) was used to make sure that there weren't any hairpin bends or self-dimers of the primers of the selected transcripts.

Four candidate RGs can be found in Table 4.1. A number of potential genes were analysed for the suitability for use as reference gene. NormFinder was used for RG selection from the candidate reference genes (Andersen *et al.*, 2004; Vandesompele *et al.*, 2002).

Table 4.1 Table of candidate reference genes

Primer	Sequence (5′ - 3′)	Primer Tm	Genbank Accession no./ EST Database No.	Amplicon Length	Stability value
RPL19 F	GAAACACCAACTCCCGACA	62.9°C	DQ910355.1	223bp	0.007
RPL19 R	TCAACAGGCGACGCAACAC				
RPS26 F	GATAAGGCAATCAAGAAGTTCG	60.4°C	CD577850.1/AF-NORA-contig_270	160bp	0.014
RPS26 R	TACGGACAACCTTCGAGTGG				
GapDH F	GACTGCCACTCGTCCATC	61.6°C	EZ966147.1	139bp	0.021
GapDH R	CCTTGGTCTGCATGTACTTG				
RPS7 F	TTACTGCTGTGTACGATGCC	63.1°C	EF450776.1	134bp	0.092
RPS7 R	GATGGTGGTCTGCTGGTT				

4.3.3 cDNA synthesis

cDNA was synthesized using the Qiagen QuantiTect® Reverse Transcription Kit (CAT no.: 205311) following the manufacturer's instructions. 1 µg of extracted RNA was incubated for 2 m at 42°C with 2 µL of genomic Wipeout Buffer and made up to 14 µL with RNase-free water. The following was added after incubation: 1 µL of Quantiscript Reverse Transcriptase; 4 µL of Reverse Transcriptase Buffer; and 1 µL of Reverse Transcriptase Primer Mix. The mixture was then incubated for 1min at 42°C, and then subsequently incubated for 3 m at 95°C in order to inactivate the Quantiscript Reverse Transcriptase. The cDNA was measured using the NanoDrop 2000 Spectrophotometer as described in Chapter 3 (see Section 3.3.1) in order to determine the cDNA concentration.

Aliquots of 100 ng/µL of cDNA were made from the synthesized stock cDNA using the formula: $C_1 \times V_1 = C_2 \times V_2$.

4.3.4 Relative-quantitative Real-time PCR

The Bio-Rad CFX96^{TM} Real-Time PCR Detection System (Bio-Rad Laboratories Inc.) was used to perform the real-time qPCR experiments. The following genes of interest were targeted for qPCR analysis: CPRR1 (AFUN020892), UNN1 (AFUN000965), UNN2 (AFUN000144), UNN3 (AFUN005543), UNN4 (AFUN009542) and UNN5 (AFUN001928). These transcripts were differentially expressed according to the RNA-seq results in Section 3.4.4. If the transcripts were significantly differentially expressed with a p-value of ≤ 0.05, consistent with the RNA-seq results (regulation was down or up regulated in the same direction as the RNA-seq data) then they were considered confirmed and used for further study in Chapter 5 (Shi and He, 2014).

UNN1 (AFUN000965), UNN3 (AFUN005543) and UNN5 (AFUN001928) were selected from the RNA-seq results as they were both unannotated and up regulated consistently throughout the three replicates. CPRR1 (AFUN020892) was chosen as it was annotated to the cuticular RR-1 protein and was consistently down regulated in Stage 3 as compared to Stage 0 in the RNA-sequencing data. UNN2 (AFUN000144) and UNN4 (AFUN009542) was also selected as it was consistently down regulated in Stage 3 compared to Stage 0 in the RNA-seq data.

The 25 µL reactions were set up using: 12.5 µL IQ SYBR Green Supermix (CAT: 170-8882; Bio-Rad Laboratories Inc.); 1 µL Forward primer (10 µM); 1 µL Reverse primer (10 µM); 1 µL cDNA and 9.5 µL nuclease-free water. The primers and amplified PCR product were sequenced (by Macrogen Inc.) to ensure that the correct gene was amplified.

The protocol was identical for all primer sets (94°C for 2 min; 39 X (94°C for 30 secs, annealing temperature for 30 secs, 72°C for 30 secs) 72°C 10 min.), except for the annealing temperature, which was optimized during primer selection. CPPR1 primers used an annealing temperature of 63.3°C and the UNN primers all had an annealing temperature of 61.4°C. The samples were amplified using a thermal cycler in a 96 well plate. The samples were performed with three biological replicates as well as three technical replicates, to produce biological and technical reproducibility. RPL19 and RPS26 were used as reference genes to normalize the data. The significance level (α) was set 0.05. The results were analysed using the Bio-Rad CFX Maestro software (Bio-Rad, 2006).

4.4 **Results**

4.4.1 Assessment of RNA quality

The RNA that was extracted had its integrity assessed as in the previous section (see Section 3.3.1). The RNA was of good quality and again showed the "hidden break", as was expected from insect RNA.

The Nanodrop 2000 results of these samples can be seen in Table 4.2. All the samples from both Stage 0 and Stage 3 had a nucleic acid concentration ranging between 263.7 ng/µL – 735.9 ng/µL. All the samples had a 260/280 ratio above 2, which shows that there was no contamination of proteins, phenols, or any other contaminants.

Table 4.2 Nanodrop 2000 results RNA concentration

Sample*	**Nucleic Acid Concent. (ng/µL)**	**A260/280**	**A260/230**	**Sample Type**
St. 3.1	518.1	2,11	2.15	RNA
St. 3.2	446	2,15	1,68	RNA
St. 3.3	447	2,15	1,2	RNA
St. 0.1	263,7	2,15	1,46	RNA
St. 0.2	559,2	2,13	1,68	RNA
St. 0.3	735,9	2,17	1,94	RNA

*St. n.n = Stage n. Biological replicate number n

4.4.2 Primer design and reference gene selection

Table 4.1 shows the candidate reference genes evaluated and the primer sequences (Nardini *et al.*, 2012). The stability values were calculated using NormFinder (*.xla, MS Excel 2003) v0.953. RPL19 had a stability value of 0.007 and RPS26 had a stability value of 0.014, which were the two reference genes with the most stable reference genes over the two different time points and were subsequently used as reference genes for normalization of the qPCR data. Table 4.3 show the primers that were designed for the six transcripts that were to be validated.

Table 4.3 Primer sequences for the CPRR1, UNN1, UNN2, UNN3, UNN4 and UNN5 genes

Primer	Sequence	Tm (°C)
CPRR1 AFUN020892; Forward	TCTCAAACCACGACATCTGC	55.3°C
Reverse	GTTGACCGTGACCAGAGGAT	56.4°C
UNN1 AFUN000965; Forward	CGTACATCTGGGGTAACATAGTGG	57.0°C
Reverse	GCACTGTAACCTATCTCGTCCTG	57.1°C
UNN2 AFUN000144; Forward	ATGTCCTCGGATGAGAGTGG	55.6°C
Reverse	GTCCCTGTTACGGTTGCATT	55.5°C
UNN3 AFUN005543; Forward	AGGGGTTACCGTGATCAC	53.3°C
Reverse	GCACACGAGTTATCCACATG	54.1°C
UNN4 AFUN009542; Forward	AGTGCCATCGAAGCGATCCA	61.96
Reverse	GTGGTTCCAGCGGCAAGATG	61.93
UNN5 AFUN001928; Forward	ACCCGTACGACGTTCTTCCC	61.87
Reverse	CTCCTCCATACATGGCCGCA	62.04

4.4.3 cDNA synthesis

The RNA was reverse transcribed into cDNA and was quality checked using the nanodrop. All samples had a 260/280 ration of between 1.7 - 1.9 and a 260/230 ratio of between 2 - 2.2. Sample St. 0.1 had a 260/230 ratio below 2, however the absorbance at 230 is indicative of guanidium thiocyanate contamination from the RNA extraction. Guanidium thiocyanate is present in most commercial cell lysis buffers and considered not to have an effect on qPCR (Stahlberg and Bengtsson, 2014). There is no evidence that guanidium thiocyanate has an effect on qPCR if in a low enough concentration (≤50 mM, as found in commercial products) (Stahlberg and Bengtsson, 2014). This ratio shows that there was no phenol or protein contaminants. Aliquots of 100 ng/µL cDNA for each sample were diluted from the stock cDNA (Table 4.4), by using the formula $C_1 \times V_1 = C_2 \times V_2$, as can be seen in Appendix 8.3.

Table 4.4 Stock cDNA concentrations

Sample*	Nucleic Acid Concent. (ng/µL)	A260/280	A260/230	Sample Type
St. 3.1	1 530,6	1,86	2,28	DNA
St. 3.2	1 607,9	1,86	2,16	DNA
St. 3.3	1 548,8	1,85	2,12	DNA
St. 0.1	648.8	1,71	1,58	DNA
St. 0.2	1 626,9	1,85	2,19	DNA
St. 0.3	1 452,2	1,86	2,23	DNA

*St. n.n = Stage n. Biological replicate number n

4.4.4 Relative-quantitative real-time PCR

The normalized expression of the qPCR results (figure 4.2; table 4.5), show that the results of the RNA-seq from Chapter 3 (Section 3.4.4, Table 3.6) are reproducible with qPCR. CPRR1 (AFUN020892), UNN2 (AFUN000144) and UNN4 (AFUN009542) were all down regulated in Stage 3 compared to Stage 0 by a log2 fold change of -4.339, -1.872 and -15.610 RNE (Relative normalized expression), respectively. Whereas UNN1 (AFUN000965), UNN3 (AFUN005543) and UNN5(AFUN001928) were all up regulated in Stage 3 compared to Stage 0 by a log2 fold change of 1.864, 3.755 and 2.461 Relative normalised expression (RNE), respectively (Figure 4.2; Table 4.5). CPRR1 (AFUN020892), UNN1 (AFUN000965), UNN2 (AFUN000144), UNN3 (AFUN005543), UNN4 (AFUN009542) and UNN5 (AFUN001928) all had a differential RNE that did not exceed the P-value threshold of 0.05, indicating that they were all six statistically significant in their differential expression. Figure 4.2 and Table 4.5 shows that all six transcripts were confirmed significantly differentially expressed (Shi and He, 2014).

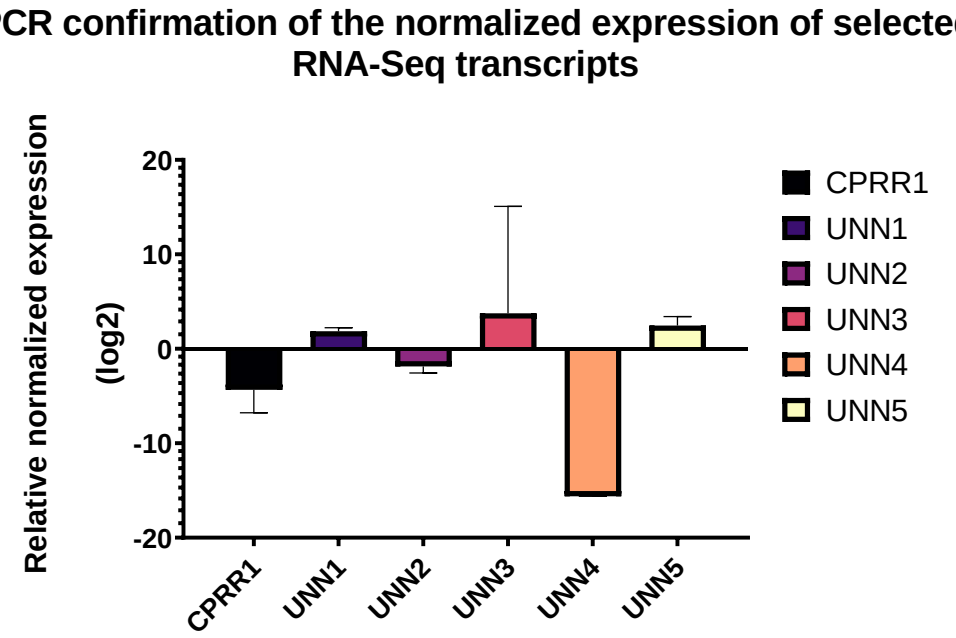


Figure 4.2 Validation of the normalized expression of RNA-Seq transcripts

The bars indicate the expression of the transcripts of Stage 3 compared to Stage 0. The transcripts were relative to the control group (Stage 0), which is set to zero. The y-axis depicts the relative normalized expression in log2. The results were normalized to the reference genes RPL19 and RPS26. CPRR1 (AFUN020892), UNN2 (AFUN 000144) and UNN4 (AFUN009542) of Stage 3 is down regulated compared Stage 0 of their corresponding transcript. UNN1 (AFUN000965), UNN3 (AFUN005543) and UNN5 (AFUN001982) of Stage 3 were upregulated compared to Stage 0 of their corresponding transcript. The UNN3 (AFUN005543) transcript had a higher variation when compared to the other transcripts and was most likely due to the transcripts small size. All samples P- value ≤ 0.05.

Table 4.5 qPCR normalized expression results

Target*	**Relative Normalized Expression of qPCR (log2)**	**P-Value of qPCR RNE**	**Relative Normalized Expression of RNA-seq data (log2)**
Control (St. 0)	N/A	N/A	N/A
CPRR1 (St. 3)	**-4.339**	**0,007234**	**-11.386**
UNN1 (St. 3)	**1.864**	**0,000011**	**2.403**
UNN2 (St. 3)	**-1.872**	**0,007364**	**-2.500**
UNN3 (St. 3)	**3.755**	**0,000017**	**3.425**
UNN4 (St. 3)	**-15.610**	**0.025277**	**-114.715**
UNN5 (St. 3)	**2.461**	**0.031855**	**62.939**

*St. n.n = Stage n. Biological replicate number n
Green = Up-regulation in expression
Blue = Down-regulation in expression
RNE = *Relative normalized expression*
The transcripts were relative to the control group (Stage 0), which is set to one. The results were normalized to the reference genes RPL19 and RPS26. CPRR1 (AFUN020892), UNN2 (AFUN 000144) and UNN4 (AFUN009542) of Stage 3 are down regulated compared to Stage 0 of their corresponding transcript. UNN1 (AFUN000965), UNN3 (AFUN005543) and UNN5 (AFUN001982) of Stage 3 were up regulated compared to Stage 0 of their corresponding transcript. All samples P- value ≤ 0.05.

4.5 Discussion

In the previous study (Section 3.5) some of the expression patterns of the sexual development of the male *An. funestus* mosquito were characterized.

The confirmed differentially expressed regulation of the CPRR1 (AFUN020892), UNN1 (AFUN000965), UNN2 (AFUN000144), UNN3 (AFUN005543), UNN4 (AFUN009542) and UNN5 (AFUN001928) transcripts in Stage 3 as compared to Stage 0 provides confirmation that the RNA-sequencing results are differentially expressed.

The RNA that was used for the qPCR and the RNA-sequencing, was collected from separate cohorts, which provides biological reproducibility and confidence. Both the qPCR and RNA-seq results were done in triplicate, providing technical reproducibility. The qPCR assay contained 3 biological repeats and 3 technical repeats, per time-point, according to the MIQE guidelines (Bustin *et al.*, 2009; Nolan *et al.*, 2011; Bustin, 2017; Bustin and Huggett, 2017; Bustin and Nolan, 2017).

An increase in selected cuticle transcript expression during the developmental phase (Stage 0, 0 h post ecdysis to Stage 4, 24 h post ecdysis) of the adult male anopheline was expected. This expectation was based on the hypothesis that the male *Anopheline* during development needs to harden and strengthen its exoskeleton so that it can sustain any physical and chemical forces it will experience when fully developed. The CPRR1 (AFUN020892) transcript was confirmed to have been up regulated in Stage 0 and decreased in expression by Stage 3. The confirmed down regulation of the CPRR1 (AFUN020892) transcript in Stage 3 as compared to Stage 0 is in support of a softening of the endocuticle during the process of sexual maturation, as mentioned in section 3.5. The hypothesis of the possible softening of the endocuticle could allow for the terminalia of the genitalia to fully rotate to sexual maturity and is possibly the reason why the CPRR1 (AFUN020892) transcript is up regulated in Stage 0 (no genitalia rotation) with a decreased expression in Stage 3 (rotation of the genitalia).

There is the possibility that at Stage 3, the male *An. funestus* mosquito has fully developed in regard to its cuticle development. This means that it is possible that at Stage 3 the expression of the cuticle transcript CPRR1 (AFUN020892) turns off as there is no further need to develop the

cuticles. However, at Stage 3, the male *An. funestus* mosquito seems to be darker in colour (an indication that cuticle has not hardened or developed) and seems to have a softer exoskeleton relative to its 36-hour old counterparts. The confirmation of the down regulation of the CPRR1 (AFUN020892) transcript is in support of the hypothesis of the softening of the endocuticle in order for genitalia rotation to take place. Although no studies could be found that indicate specifically that down regulation of the CPRR1 gene in mosquitoes leads to the softening of the endocuticle, a study on termites found that thicker endocuticles are caused by an up regulation of the CPRR1 transcripts (Rasheed *et al.*, 2019). However, the three up regulated transcripts in Stage 3 as compared to Stage 0 could be involved in the hardening of the exoskeleton and could be responsible for cuticle development during the first 36 h post ecdysis. Further research needs to be undertaken on the biological function of the CPRR1 transcripts, including the CPRR1 (AFUN020892) transcript in male *An. funestus* before any firm conclusions can be drawn.

There were several genes that were found to be differentially expressed between Stage 0 and Stage 3, as was to be expected. Three unannotated transcripts that had consistent differential expression between Stage 0 and Stage 3, UNN1 (AFUN000965), UNN2 (AFUN000144) and UNN3 (AFUN005543) were confirmed. The lowest differentially expressed transcript, UNN4 (AFUN009542), and the highest differentially expressed transcript, UNN5 (AFUN001928), from the RNA-seq data was also confirmed. Due to the consistency in differential gene expression between Stage 0 and Stage 3 it was assumed that they play a functional role in the sexual maturation of the male mosquito, however as with CPRR1 (AFUN020892), their biological importance or role in sexual maturation, such as their involvement in male genitalia rotation, antennal fibrillae erection, accessory gland maturation and spermatogenesis will still have to be investigated in future studies (Roth, 1948; Clements, 1992). It would be useful to confirm the expression levels of unannotated genes to assign function in future studies in order to further understand the underlying genomic events that lead to the male maturation.

5 Investigation of the potential biological function of selected transcripts

5.1 Introduction

RNA interference (RNAi) was discovered by researchers working on the nematode *Caenorhabditis elegans* (Fire *et al.,* 1998). They observed that the introduction of either sense or antisense RNA caused the homologous genes expression to be suppressed. Fire *et al.,* 1998, showed that, approximately a 10-fold increase, in specific silencing effect could be achieved, if both sense and antisense RNA were introduced into the organism, when compared to either of the single RNA strands alone. This study indicated that the presence of contaminating double stranded RNA (dsRNA) caused the silencing effect seen in earlier experiments (Fire *et al.,* 1998). The RNAi is a post-transcriptional process, which requires that the RNAi molecule to be homologous to the exons of a target gene, after transcription of the mRNA. There have been many gene-silencing experiments on many different organisms, including plants (Vierstra, 1996; Waterhouse *et al.,* 1998; Chuang and Meyerowitz, 2000), mice (Wianny and Zernicka-Goetz, 2000), *Drosophila* (Kennerdell and Carthew, 2000), and mosquitoes (Hoa *et al.,* 2003; Catteruccia and Levashina, 2009).

The transcription of specific genes with complimentary nucleotide sequences are degraded by the RNA interference (RNAi) pathway. By degrading the mRNA after transcription, causes the translation of that specific mRNA to be interrupted (Agrawal *et al.,* 2003). It has been hypothesized that RNA silencing evolved as a defence against “invasive” nucleic acids (Yoder *et al.,* 1997; Jones *et al.,* 1999; Jensen *et al.,* 1999; Li *et al.,* 2002). Interference can be brought about through different RNA molecules which includes, small interfering RNA (siRNA), sometimes known as short interfering RNA or silencing RNA. These siRNA’s have a well-defined structure: a short (between 20 – 24 bp) dsRNA with 5' phosphorylated ends and 3' hydroxylated ends with two nucleotides that overhang the sides. An endoribonuclease enzyme, called Dicer, produces small hairpin RNAs from long dsRNAs, as can be seen in Figure 5.1. The

siRNA mechanism for interference of gene translation is: 1) Dicer (an endoribonuclease) cleaves long dsRNA to form short interfering RNA or siRNA. 2) After delivery of the siRNA into the organisms cells, the RISC complex is formed when proteins, including one of the Argonaute family, is incorporated with the siRNA. 3) The siRNA is unwound when incorporated into the RISC complex, to form single stranded RNA. 4) The thermodynamically less stable strand remains part of the RISC complex, due to its base pairing at the 5´ end. 5) messenger RNA is scanned by the RISC complex for mRNA that is homologous to the single stranded siRNA that is incorporated to the RISC complex. 6) The target mRNA is bound to the single stranded siRNA (part of the RISC complex) and mRNA cleavage is induced. 7) The mRNA is now cleaved, and the cell recognizes the cleaved mRNA fragments as abnormal. This causes degradation of the mRNA and in turn no translation of the mRNA into amino acids and then proteins (Agrawal *et al.,* 2003). In effect silencing the gene that encodes mRNA.

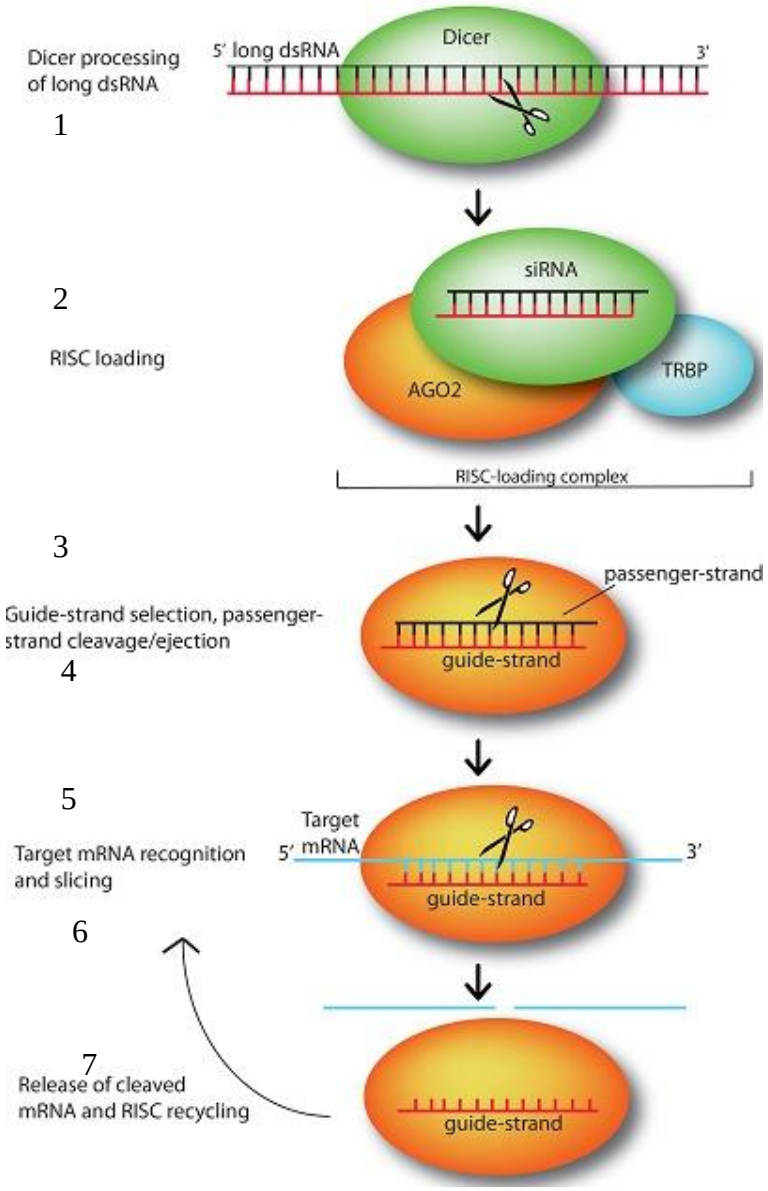


Figure 5.1 RNAi pathway

1) The endoribonuclease Dicer attaches and cleaves dsRNA into siRNA, 2) which is taken into the cell and is incorporated into other proteins such as of the Argonaute family to form the RNA-Induced Silencing Complex (RISC). 3) The siRNA unwinds to form single stranded siRNA, and 4) the thermodynamically most stable single stranded siRNA attaches to the RISC complex. 5) The single stranded siRNA can scan for complementary mRNA, 6) and binds to it, 7) allowing the RISC complex to cleave to the mRNA, which the cell will recognize as abnormal and will dispose of the cleaved mRNA (taken directly from Dana *et al.*, 2017).

The siRNAs have a 100% complementarity and thus have a very tight target specificity and work by cleaving the mRNA before translation. Small interfering RNAs (siRNAs) can also be transfected into cells, which means that synthetic siRNAs can be used as a gene knock-down technique. siRNAs are often used for validating gene function and drug targeting (Arts *et al.*, 2003; Dorsett and Tuschl, 2004; Xie *et al.*, 2006).

One of the most challenging parts of using the RNAi technique is that because it consists of RNA, it is easily degraded and destroyed in natural environments due to RNases (Zhang *et al.,* 2010). The consequence of this degradation is that the delivery of any RNAi, dsRNA, or siRNA, is key to successful knock-down (Zhang *et al.,* 2010). Directly exposing the organism through feeding with RNAi is normally the easiest/simplest delivery method, although there is a higher likelihood of degradation of the RNA due to the longer exposure in an uncontrolled environment. There are other delivery methods, like nano-injection, soaking etc. Feeding, however, allows for a high amount of siRNA degradation, so it is postulated that the dsRNA/siRNA should be attached to nanoparticles in order to preserve the integrity of RNA through-out the delivery period (Zhang *et al.,* 2010).

It has been proposed that the efficacy of siRNA can be increased by conjugating the siRNA to nanoparticles, which prevents degradation of the siRNA. Chitosan is the most commonly used nanoparticle for drug delivery due to its biodegradability and non-toxic effect (Nagpal *et al.,* 2010). Chitosan is a linear polysaccharide composed of randomly distributed *N*-acetyl-D-glucosamine (acetylated unit) and β-(1→4)-linked D-glucosamine (deacetylated unit) and (Figure 5.2, A). It is made by treating the chitin shells of shrimp and other crustaceans with an alkaline substance, like sodium hydroxide. Chitosan has been used before for production of nanomaterials and improved drug delivery (Agnihotri *et al.,* 2004; Nagpal *et al.,* 2010; Tiyaboonchai, 2013; Li *et al.,* 2018). Chitosan and siRNA nanoparticles have been synthesized, where the electrostatic forces between the amino groups of the chitosan and negative charges of the phosphate group on the backbone of the siRNA bind together (Figure 5.2, B) (Zhang *et al.,* 2010).

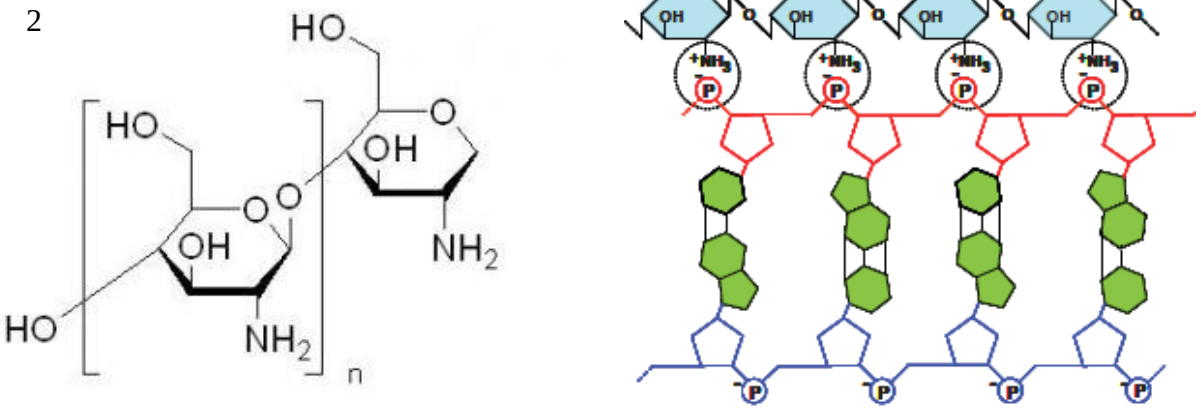


Figure 5.2 Chitosan nanoparticle and siRNA interaction

A: Molecular structure of chitosan. **B:** Chitosan and siRNA nanoparticles were formed through electrostatic binding between the amino groups of the chitosan and negative charges of the phosphate group on the backbone of the siRNA (Taken directly from Zhang *et al.*, 2010)

In the previous studies (Chapters 3 and 4), it was discovered and confirmed that six transcripts were differentially expressed between sexually mature (Stage 3) and sexually immature (Stage 0) male *An. funestus* mosquitoes. These six transcripts were selected as they were possibly associated with the sexual maturation process. The UNN1 (AFUN000965) transcript, which was up-regulated in Stage 3 as compared to Stage 0, was selected for knock-down, to attempt to characterize the transcript's functional annotation.

Knocking down a transcript through the use of RNAi and observing the corresponding phenotypic change(s), it is possible to elucidate that specific transcripts biological function or role. The UNN1 (AFUN000965) transcript has a possible role in the male maturation process, as the previous studies have associated an up regulation of the UNN1 (AFUN000965) transcript in Stage 3 as compared to Stage 0 of the development in male *An. funestus* mosquitoes (Chapters 3 and 4). Knocking-down the UNN1 (AFUN000965) transcript and observing the effects that the knock-down has on the sexual development of the *An. funestus* male could provide information of the functional role that the UNN1 (AFUN000965) transcript plays. It was hypothesized that the knock-down of the UNN1 (AFUN000965) transcript would lead to a phenotypic change in the male maturation development.

In order to ensure that there is no non-specific knock-down caused by the siRNA or that the presence of the siRNA does not have any effect on the selected transcripts or that the delivery of the siRNA does not cause any transcriptional changes, an endogenous positive control (RPL19) and an exogenous negative control (mβ2M) as well as an untreated negative control (Untreated) is used, respectively. Designing an exogenous negative control is done by designing siRNA for a transcript that is not found in the target organism.

As mentioned in the section on the male mosquito mating biology and development (Section 1.1.3), the organs of the male reproductive system can provide valuable data about sexual maturity and the ability to copulate (Mahmood and Reisen 1982; Huho *et al.,* 2006). Male age has been associated with testes size in *An. stephensi* in which it was found that sexually mature individuals had longer testes while immature or older mosquitoes had shorter testes (Mahmood and Reisen, 1982). The measuring of testes size, after the UNN1 (AFUN000965) transcript has been knocked down through RNAi, could determine if the UNN1 (AFUN000965) transcript plays a functional role in the sexual development of the male *An. funestus* mosquito. It was thought that if UNN1 (AFUN000965) played a role in the sexual development of the male *An. funestus* mosquito and was knocked down, there would be a phenotypic change in the testes size.

To my knowledge this is the first study where the RNAi technique has been performed on the *An. funestus* mosquito. At the time of this study, all previous studies were performed on *An. gambiae* and *An. stephensi* (Brown *et al.*, 2003; Hoa *et al.*, 2003; Catteruccia and Levashina, 2009; Pondeville *et al.,* 2013; Puglise *et al.*, 2016; Negri *et al.,* 2019; Letinic *et al.*, 2020).

5.2 Rationale

In the previous study (Chapter 4), six genes were selected for their possible role in the male maturation process of the male *An. funestus* mosquito. The rationale of this part of the study was to determine if there was a sexual development phenotypic role associated with the UNN1 (AFUN000965) transcript in *An. funestus* males. As well as to optimise and evaluate the use of siRNA conjugated to chitosan nanoparticles in order to determine if this is a suitable method for *An. funestus* transcript silencing. To my knowledge, this is the first study to attempt this technique on the *An. funestus s.s.*

Small interfering RNA (siRNA) was designed to silence the Ribosomal protein L19 (RPL19) as a positive endogenous control for the UNN1 (AFUN000965) transcript. The UNN1 (AFUN000965) transcript was unannotated and through the use of RNAi the genes' function could possibly be determined.

There has been evidence of coupling siRNA to nanoparticles in order to improve the delivery and efficacy of the siRNA. Chitosan nanoparticles were conjugated to RPL19 as well as UNN1 (AFUN000965) siRNA, and fed to newly emerged male *An. funestus* mosquitoes, to determine if nanoparticles increased the efficacy of the siRNA knock-down in this main malaria vector species.

5.3 Material and methods

5.3.1 The effect of exogenous Mβ2M siRNA on transcript abundance of RPL19 and UNN1 with and without nanoparticles.

To ensure that the presence of the siRNA did not have any effect on the selected transcripts, an exogenous negative control was used. This was done by designing siRNA for a transcript that is not present in the *An. funestus* mosquito. Mouse beta-2 microglobulin (Mβ2M) was selected as the exogenous negative control (Table 5.1). In order to ensure that knock-down was not caused by the delivery of the siRNA, a negative control was added. The negative control was an untreated sample group where siRNA-free water was delivered.

5.3.1.1 Experimental set up

Anopheles funestus (FUMOZ) males from the insectary (Botha DeMeillon insectary at the NICD, NHLS, Sandringham, Johannesburg, South Africa) were used for the experiment. As males emerged from eclosion (Stage 0) they were transferred to a 2L holding cage with an appropriate feeding solution.

Four cages with fifteen newly emerged male *An. funestus* (FUMOZ) were used in each. One cohort was fed on a 10% sucrose solution only (Untreated), and another on 10% sucrose solution and nanoparticle mixture that has not been conjugated to any siRNA (Untreated-NP). The third cohort was fed with Mβ2M siRNA and 10% sucrose solution (Mβ2M), and a fourth cohort was fed on Mβ2M siRNA-NP and a 10% sucrose solution (Mβ2M-NP). All four cohorts were allowed to feed of their respective treatments for 24 h. The experiment was repeated three times to represent different biological replicates.

5.3.1.2 siRNA design

siRNA was designed by the Ambion® novel algorithm to be homologous to the Mouse beta-2 microglobulin (Mβ2M) sequences to target the Mβ2M mRNA. Mβ2M siRNA was synthesized by ThermoFisher using Ambion® in *vivo* chemical modifications to ensure minimal off-target activity. The *Silencer*® Select algorithm was used on putative Mβ2M siRNA to analyze the full coding sequence of each gene to prevent off target effects as well as to allow for 5- to 20- fold lower concentrations of siRNA than is needed. The siRNA design for targeting exogenous Mβ2M

transcripts is in Table 5.1. Each siRNA strand was individually purified by high performance liquid chromatography (HPLC), desalted, and annealed with its complimentary strand. A semi-permeable membrane was used to remove any excess salt. The Mβ2M siRNAs were then filtered through a 0.2 µm pre-sterilized filter and tested for the presence of endotoxin. The Mβ2M siRNAs were re-suspended according to Custom Ambion® *In vivo* siRNA instructions. The Mβ2M siRNA was resuspended in PBS to make 200 µM stock solution. The Mβ2M siRNA was aliquoted out into 169 µL of 20 µg/ml aliquots which was used to for knock-down or was combined with nanoparticles and then subsequently used for knock-down, experiments.

Table 5.1 Table of siRNA designed for the knock-down of Mβ2M

	Mβ2M	
Strand	**Sense**	**Antisense**
Sequence (5′ ->3′)	GCCUGUAUGCUAUCCAGAAtt	UUCUGGAUAGCAUACAGGCcg
Length	21	21
Percent G/C	43%	52%
Molecular weight	6800	6800
Molar Extinction Coefficient	205700	208100
Annealed Molecular Weight	13500 (40 µg/A260 was used to calculate concentration)	

5.3.1.3 Preparation of chitosan/siRNA nanoparticles

Chitosan-siRNA nanoparticles were conjugated as described by Zang *et al.* (2010). The chitosan-siRNA nanoparticles (NP) were generated from chitosan from shrimp shells (CAT No.: C3646-25G; ≥75% deacetylated; Sigma-Aldrich, Milwaukee, WI, USA) which was dissolved in a sodium acetate buffer (0.1 M sodium acetate – 0.1 M Acetic acid, pH 4.5) to make a 0.02% wt/vol working solution. A total of 160mg of chitosan was dissolved in 100 mL of sodium acetate buffer to make a chitosan solution.

The siRNA of Mβ2M was resuspended in 200 µM in 1250 µL of nuclease-free water, giving a 200 µM Mβ2M siRNA solution of 1.250 mL at 20 µg/mL. A 169 µL of Mβ2M siRNA solution (20 µg/mL) was added to 169 µL chitosan solution and incubated at room temperature for 30 m and vortexed quickly to make up 338 µL of Mβ2M siRNA-nanoparticle (Mβ2M siRNA-NP complexes (stored at -20°C)). All of the Mβ2M siRNA that was to be used with nanoparticles

(NP) to determine if the exogenous Mβ2M siRNA coupled with nanoparticles induces a silencing effect on the target transcripts.

5.3.1.4 siRNA delivery of nanoparticles

Four 2L cages were prepared containing fifteen newly emerged male *An. funestus* mosquitoes: The cages were labelled 1) Untreated 2) Untreated-NP 3) Mβ2M 4) Mβ2M-NP, respectively. The Mβ2M siRNA was delivered through feeding fifteen newly emerged adult males 676 µL of Mβ2M siRNA/sugar-water mixture for 24 h. A total of 338 µL of 200 µM Mβ2M siRNA was added to 338 µL of 10% sucrose solution, to make up the MβM siRNA/sugar-water mixture (Mβ2M). This was repeated with the Mβ2M-nanoparticle siRNA mixture (Mβ2M-NP). A 676 µL of 10% sucrose solution was prepared for a negative untreated control to determine if the 10% sucrose solution had an effect on gene expression (Untreated). A second negative untreated control was mixed by adding 338 µL of 10% sucrose solution and 338 µL of nanoparticle mixture that did not contain any nanoparticles that had siRNA conjugated to it (Untreated-NP). This was to ensure that the nanoparticles themselves did not cause a knock-down effect. The 10% sucrose solution consisted of RNase free water in order to minimise any RNase degradation on the siRNA.

The treatments were placed into empty RNase-free microcentrifuge tubes with the lid removed. The opening was closed with a ball of cotton wool and a small hole was made at the bottom of the microcentrifuge tube. The microcentrifuge tube was placed upside down on a 2L cage, with the cotton wool ball against the netting where the mosquitoes were able to feed from the cotton ball that was now soaked in their respective siRNA treatments. The hole at the bottom of the microcentrifuge tube, which was now at the top, allowed for the siRNA/sugar-water mixture to continually soak the cotton ball and prevent it from drying out, allowing the mosquitoes constant access to the sugar-water mixture to which contains the Mβ2M siRNA mixture. After 24 h, the *An. funestus* from each cage were sacrificed for 10 m @ -20°C and the mosquitoes in each cage were examined for phenotypic changes (i.e., genitalia rotation, antennal fibrillae erection and male accessory gland development) and stored in TRIzol. The siRNA delivery protocol, as described above, was performed three times to represent three biological replicates.

5.3.1.5 Knock-down effects on phenotype

The antennal fibrillae, genitalia rotation, accessory gland maturation and testis size of a total of 45 samples were evaluated for each treatment.

Fifteen samples of each biological replicate, Untreated, Untreated-NP, Mβ2M siRNA treated, and Mβ2M siRNA-NP treated mosquitoes, each, had their antennal fibrillae viewed under the dissecting microscope in order to determine if antennal fibrillae erection had taken place. The antennal fibrillae need to be able to erect in order to find a mate within a swarm and to be considered sexually mature (Clements *et al.,* 1992).

The fifteen male *An. funestus* mosquitoes treated with Mβ2M siRNA, Untreated and Mβ2M siRNA-NP, Untreated-NP had their genitalia rotation examined after 24 h to determine if clasper rotation was observed and the stage of rotation was recorded (described in Section 2.3.4).

After antennal fibrillae and genitalia rotation phenotypes were determined, the Mβ2M siRNA and Mβ2M siRNA-NP treated mosquito males had their accessory glands and testes examined for sexual maturity against the two negative controls (Untreated and Untreated-NP). This was achieved by placing the fifteen male mosquitoes ventrally in a drop of phosphate buffered saline (PBS) with the abdomen resting in the PBS. Under the stereoscope, with a pair of forceps, the males were grasped by the thorax and with a fine tip needle, the claspers were removed by piercing them and gently pulling them away from the thorax. The extraneous tissues were removed with dissecting needles and the reproductive organs were isolated under a cover slip (as can be seen in Figure 1.3; Section 1.1.4.2 on male accessory glands). The testes and MAGs of the Untreated-NP, Mβ2M and Mβ2M-NP were examined under a microscope and compared to the untreated control to determine if there were any phenotypic changes or abnormalities observed between the untreated negative control and the treated controls (MR4 manual). Since there were only males in the cage, mating would not have taken place, hence if there is absence of a clear or transparent area around the edge of the MAG, the sample was recorded as mature, if there were no abnormalities observed with regards to testes morphology.

After dissection, all biological material was immediately placed in TRIzol for RNA extraction and subsequent differential gene expression studies of the RPL19 transcript and the UNN1

(AFUN000965) transcripts to determine if treatment with the exogenous Mβ2M siRNA, with and without nanoparticles caused a silencing effect on the RPL19 and the UNN1 (AFUN000965) transcripts.

5.3.1.6 RNA extraction, cDNA synthesis and qPCR

Total RNA was extracted from the fifteen Untreated, Untreated-NP, Mβ2M siRNA and Mβ2M siRNA-NP treated mosquitoes using the Life Technologies TRIzol®Reagent (CAT. No.: 15596). This was performed as in Section 3.3.1.

The RNA was reverse transcribed to cDNA using QuantiTect Reverse Transcription kit (Qiagen) according to the manufacturer's instructions, to determine if there was silencing of the transcript, through qPCR analysis. This was performed as described in section 4.3.3.

The RPL19 (as an endogenous positive control) and the UNN1 (AFUN000965) transcript was analysed by qPCR which was performed on all four treatments (Untreated, Untreated-NP, Mβ2M siRNA and Mβ2M siRNA-NP). The qPCR reaction in (25 µL reactions) contained iQ^{TM} SYBR® Green Supermix (CAT No.: 170-8882) and the appropriate primers (0.3 µM) for the specific transcripts ((Tables 4.1 and 4.2) (Section 4.3.2)). RPS26 and RPS7 were the preferred reference genes after stability evaluation (section 4.3.2).

The relative normalized expression (RNE) of the RPL19 transcript in the untreated treatment group was compared to the RNE of the sample group treated with siRNA targeting the Mβ2M transcript (Untreated Vs. Mβ2M). This was to ensure the siRNA delivery, or that the siRNA did not have any off-target effects on the expression of the RPL19 transcript, acting as a negative control. This was repeated with cohorts treated with a 10% sucrose and nanoparticles solution and compared with the RNE of the RPL19 transcript in the cohort treated with Mβ2M conjugated with nanoparticles (Untreated-NP Vs. Mβ2M siRNA-NP). It was hypothesized, when comparing the negative control (Untreated) to an exogenous negative control (Mβ2M), that there will be no differential expression of the RPL19 transcript, even if conjugated to chitosan nanoparticles.

Once the RPL19 transcription levels were measured with regards to the negative exogenous control, the UNN1 (AFUN000965) transcription levels were measured with regards to the negative exogenous control. This was performed by repeating the RNE expression as above but measuring the transcription levels of the UNN1 (AFUN000965) transcript in the Untreated, Untreated-NP, Mβ2M siRNA and the Mβ2M siRNA-NP treated samples.

5.3.2 Effect of RPL19 siRNA and UNN1 siRNA on transcript abundance with and without nanoparticles

Once it was established that the Mβ2M siRNA or delivery of the siRNA itself does not cause any off-target effects on the RPL19 and UNN1 (AFUN000965) transcripts, Mβ2M was used as an exogenous negative control in order to determine if RPL19 and UNN1 transcripts were successfully knocked-down by the RPL19 and UNN1 siRNA, respectively. The conjugation of nanoparticles to the RP19 and UNN1 siRNA was also evaluated.

5.3.2.1 Experimental set up

Biological material was collected as in Section 5.3.1.1.

5.3.2.2 siRNA design

The RPL19 and UNN1 siRNA was designed as described in Section 5.3.1.2 (Table 5.2).

Table 5.2 Table of siRNA designed for the knock-down of RPL19 and UNN1 (AFUN000965) transcript

	UNN1 siRNA		**RPL19 siRNA**	
Strand	**Sense**	**Antisense**	**Sense**	**Antisense**
Sequence (5′ ->3′)	CGUCUUCGAU GGAAUCCGAtt	UCGGAUUCCAUC GAAGACGgt	GAAACAAGCAGC AGCAUGAgU	ACUCAUGCUGCU GCUUGUUUc
Length	21	21	21	21
Percent G/C	48%	52%	48%	48%
Molecular weight	6800	6800	6903.2	6616.2
Molar Extinction Coefficient	203500	208300	219000	191500
Annealed Molecular Weight	13500 (40 µ/A260 was used to calculate concentration)		13519.4 (40 µ/A260 was used to calculate concentration)	

5.3.2.3 Preparation of chitosan/siRNA nanoparticles

Preparation of the RPL19 and UNN1 siRNA nanoparticles were prepared as described in Section 5.3.1.4. Instead of Mβ2M siRNA being conjugated to nanoparticles, the RPL19 and UNN1 siRNA from Section 5.3.2.1 was used. They were labelled RPL19 siRNA-NP and UNN1 siRNA-NP, respectively.

5.3.2.4 siRNA delivery of nanoparticles

The RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP was delivered as described in section 5.3.1.4.

5.3.2.5 Knock-down effects on phenotype

The phenotypic knock-down effects of RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP were observed as in Section 5.3.1.5.

5.3.2.6 RNA extraction, cDNA synthesis and qPCR

The RNA extraction and cDNA synthesis of the 4 samples treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP was performed as in Section 5.3.1.6. The RNE levels of the RPL19 transcript and the UNN1 transcript was measured as in Section 5.3.1.6. The expression levels were, however, compared to the negative exogenous control, Mβ2M and Mβ2M-NP instead of the negative untreated and untreated-NP controls.

The effect of the nanoparticles conjugated to the siRNA was measured by comparing the difference in the log2 fold change of the transcript expression when treated with only siRNA and siRNA conjugated to nanoparticles. The difference in RNE in the log2 fold change of RPL19 siRNA and RPL19 siRNA-NP was taken as the effect of the nanoparticles.

5.4 **Results**

The results of Chapter 5.4 are broken up into two sections, section 5.4.1, and section 5.4.2. Section 5.4.1 is concerned with the exogenous control and provides evidence that the delivery of the siRNA (Mβ2M) and the nanoparticles did not have any off-target effects on the expression of the selected transcripts (RPL19 and UNN1 (AFUN000965)), whereas section 5.4.2 concerns with the knock down of the RPL19 and UNN1 (AFUN000965) transcripts when treated with RPL19 and UNN1 siRNA, respectively.

5.4.1 The effect of exogenous Mβ2M siRNA treatment on transcript abundance of RPL19 and UNN1 with and without nanoparticles

5.4.1.1 Phenotypic observations

The samples that were treated with the siRNA mixture with Mβ2M, Mβ2M-NP and untreated-NP's had their male sexual development phenotype compared to the untreated negative control after being immobilised. The male sexual development phenotypes that were compared were the genitalia rotation stage, the antennal fibrillae erection, testes size and MAG. There were no phenotypic changes observed in testes size of all the treatments when compared to the untreated negative control, when viewed under the microscope (Table 5.3). There were no phenotypical changes of male sexual maturity between the cohorts when comparing the treatments to the negative untreated control (Table 5.3). No phenotypical changes between the cohorts treated with chitosan nanoparticles and those without were observed.

Table 5.3 Effect of male sexual maturity phenotype after siRNA treatment

Treatment	Testes morphology and male accessory gland maturity	Genitalia rotation stage	Antennal fibrillae erection status	Number in cohort
Untreated	Mature	Stage 4	Erect	15
Untreated-NP	Mature	Stage 4	Erect	15
Mβ2M	Mature	Stage 4	Erect	15
Mβ2M/NP	Mature	Stage 4	Erect	15

5.4.1.2 Relative normalized expression of the RPL19 and UNN1 (AFUN000965) transcript treated with Mβ2M siRNA and Mβ2M siRNA-NP compared to the Untreated sample group

Relative normalized expression of the RPL19 and UNN1 (AFUN000965) transcripts were measured in *An. funestus* males fed on Mβ2M siRNA as well Mβ2M siRNA conjugated to nanoparticles (Mβ2M siRNA-NP). There was no statistically significant (P value ≥ 0.05) difference in the expression of the RPL19 transcript between the samples treated with Mβ2M siRNA and the untreated samples as can be seen in Figure 5.3 and Table 5.4. The delivery and the exogenous negative control did not have an effect on the expression of the RPL19 transcript. There was no significant (P value ≥ 0.05) difference in the expression of the RPL19 transcript treated with Mβ2M siRNA-NP and the untreated-NP samples (Table 5.4).

The same can be seen when the relative normalised expression (RNE) of the UNN1 (AFUN000965) transcript levels were measured for the samples treated with Mβ2M siRNA and Mβ2M siRNA-NP against the untreated and untreated-NP controls (Table 5.4).

The UNN1 (AFUN000965) transcript, however, was down regulated by a log2 fold change of -0.828 in the sample treated with Mβ2M siRNA (Figure 5.3). This was not statistically significant when compared to the exogenous Mβ2M control as the P value ≥ 0.05 (Table 5.4).

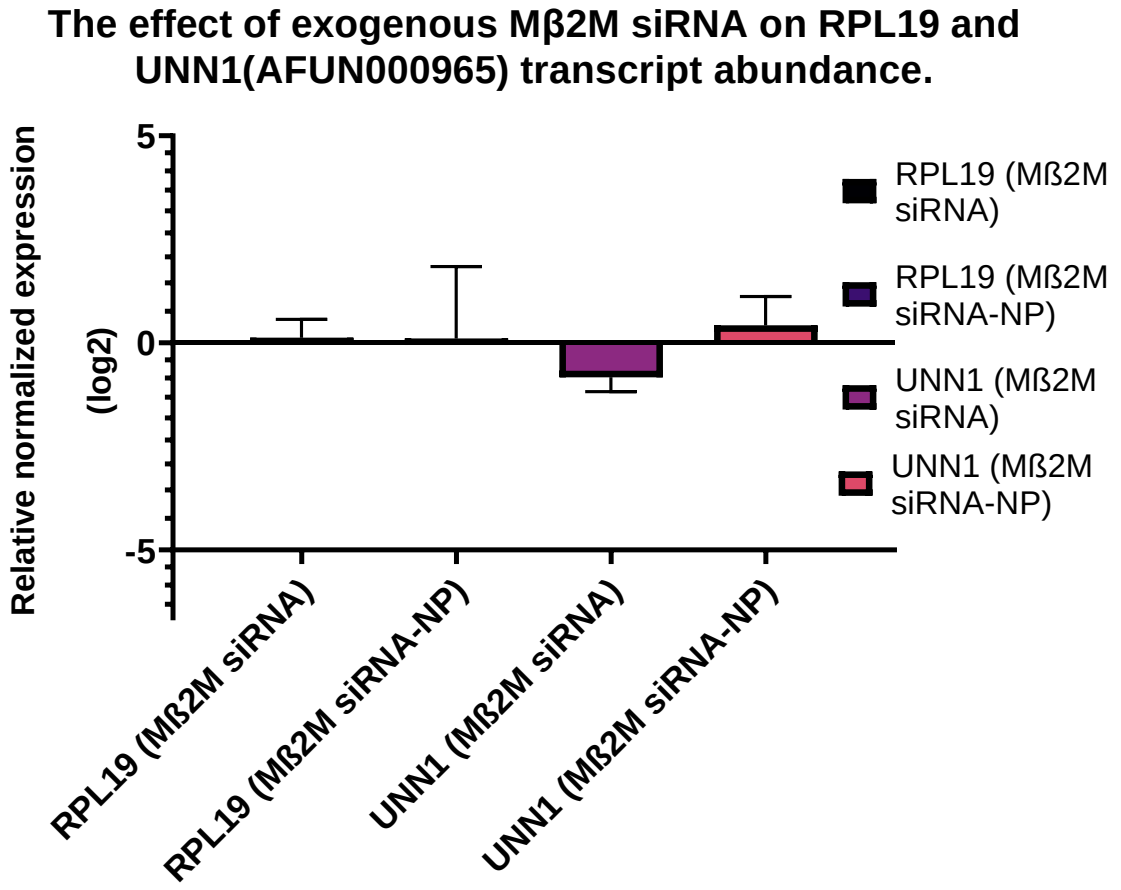


Figure 5.3 Relative Normalized Expression (RNE) in log2 fold change of the RPL19 and UNN1 transcript in *An. funestus* males treated with Mβ2M siRNA and Mβ2M siRNA-NP

Table 5.4 RNE of the RPL19 and UNN1 (AFUN000965) transcript in *An. funestus* males treated with Mβ2M siRNA and Mβ2M siRNA-NP compared to untreated males

Target	Relative Normalized Expression (log2)	P value of qPCR RNE
Untreated	N/A	N/A
Untreated-NP	N/A	N/A
RPL19 (Mβ2M siRNA)	**0.12**	**0.862489**
RPL19 (Mβ2M siRNA-NP)	**0.104**	**0.125808**
UNN1 (Mβ2M siRNA)	**-0.828**	**0.050162**
UNN1 (Mβ2M siRNA-NP)	**0.423**	**0.203895**

The transcripts were relative to the negative control groups (Untreated and Untreated-NP), which is set to one. The results were normalized to the reference genes RPS7 and RPS26. RPL19, UNN1 (AFUN 000965) transcripts had no differential expression when compared to the negative untreated and untreated-NP control groups. All samples P value ≤ 0.05 using unpaired student t-test.

Green = Up-regulation in expression
Blue = Down-regulation in expression
RNE = *Relative normalized expression*

Table 5.4 shows that the *An. funestus* males that fed on Mβ2M siRNA and Mβ2M siRNA-NP (exogenous negative control) compared to only a 10% sucrose solution (Untreated negative control) and a 10% sucrose and nanoparticle mixture (Untreated-NP negative control), had no signifcant difference in expression levels of the RPL19 transcript or the UNN1 (AFUN000965) transcript. This indicates that the siRNA or the delivery of siRNA does not induce an RNAi effect on the RPL19 or UNN1 (AFUN000965) transcript and does not influence the expression of RPL19 and UNN1 (AUN000965).

5.4.2 Effect of RPL19 siRNA and UNN1 siRNA on transcript abundance with and without nanoparticles

5.4.2.1 Phenotypic observations

After treatment, *An. funestus* cohorts fed with siRNA (Mβ2M siRNA, Mβ2M siRNA-NP, RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP) were immobilised and viewed in PBS for any phenotypic changes. The phenotypic changes that were recorded were the genitalia rotation stage, the antennal fibrillae erection, testes size and MAG maturity. There were no abnormalities observed in testes morphology when viewed under the microscope and compared to the untreated control group. Table 5.5 depicts the phenotypic changes between the samples treated with RPL19 siRNA and UNN1 siRNA when compared to the sample group treated with Mβ2M. Table 5.5 also depicts the phenotypic changes of the sample treated with RPL19 siRNA-NP and UNN1 siRNA-NP when compared to the sample group treated with Mβ2M-NP. There were no phenotypical changes of male sexual maturity between the cohorts. There were not any phenotypical changes between that of the cohorts treated with chitosan nanoparticles and those without. These observations were a limitation of the study and future studies should include the measurements of the testes size and the width of the MAG clear zone.

Table 5.5 Effect of male sexual maturity phenotype after siRNA treatment

Treatment	Testes morphology and male accessory gland maturity	Genitalia rotation stage	Antennal fibrillae erection status	Number in cohort
Untreated	Mature	Stage 4	Erect	15
Mβ2M siRNA	Mature	Stage 4	Erect	15
Mβ2M siRNA-NP	Mature	Stage 4	Erect	15
RPL19 siRNA	Mature	Stage 4	Erect	15
RPL19 siRNA-NP	Mature	Stage 4	Erect	15
UNN1 siRNA	Mature	Stage 4	Erect	15
UNN1 siRNA-NP	Mature	Stage 4	Erect	15

5.4.2.2 RNE of the RPL19 and UNN1 (AFUN000965) transcript treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP compared to the sample group treated with Mβ2M siRNA and Mβ2M siRNA-NP, respectively

After testes dissection, the biological material of the cohort RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP were placed in TRIzol, RNA extraction and cDNA synthesis were subsequently performed. This was done with Mβ2M siRNA and the Mβ2M siRNA-NP treated samples as well. The resulting cDNA was used in qPCR to determine the relative normalized expression of the RPL19 and UNN1 (AFUN000965) transcript between each of the treatments.

Of the samples that were treated with RPL19 or UNN1 siRNA, with or without nanoparticles, expressed the transcripts of RPL19 and UNN1 (AFUN000965) significantly lower than the untreated controls (P value ≤ 0.05), except for the sample that was treated with UNN1 siRNA without nanoparticles (Figure 5.4). This indicates a silencing of the RPL19 and UNN1 (AFUN000965) transcripts when treated with homologous siRNA attached to nanoparticles.

Samples treated with RPL19 siRNA had a down regulation of the RPL19 transcript of a log2 fold change of -1.003 when compared to the sample treated with the exogenous negative control (Mβ2M siRNA). Samples treated with RPL19 siRNA-NP had a down regulation of the RPL19 transcript of a log2 fold change of -1.596 when compared to the sample treated with the exogenous negative control (Mβ2M siRNA-NP). This shows that the nanoparticles increased the knock-down effect of the RPL19 siRNA significantly ($p \leq 0.277$, using unpaired student t-test) by a log2 fold change of -0.593 (Figure 5.4; Table 5.6).

Samples treated with UNN1 siRNA did not show a knock-down effect of the UNN1(AFUN000965) transcript when compared to the sample treated with the exogenous negative control (Mβ2M siRNA-NP). The expression of the UNN1(AFUN000965) transcript that was treated with UNN1 siRNA was up regulated by 0.281 log2 fold changes when compared to the exogenous negative control (Mβ2M siRNA). However, the samples treated with UNN1 siRNA-NP had a down regulation of the UNN1(AFUN000965) transcript of a log2 fold change of -1.984 when compared to the sample treated with the exogenous negative control (Mβ2M siRNA-NP). This shows that the nanoparticles increased the knock-down effect of the UNN1 siRNA significantly ($p \leq 0.277$, using unpaired student t-test) by a log2 fold change of -1.703 (Figure 5.4; Table 5.6).

As can be seen in Figure 5.4 and Table 5.6, conjugating nanoparticles to the siRNA increased the siRNA knock-down effect by 0.593 log2 fold changes in the RPL19 transcript and by 1.703 log2 fold changes in the UNN1(AFUN000965) transcript.

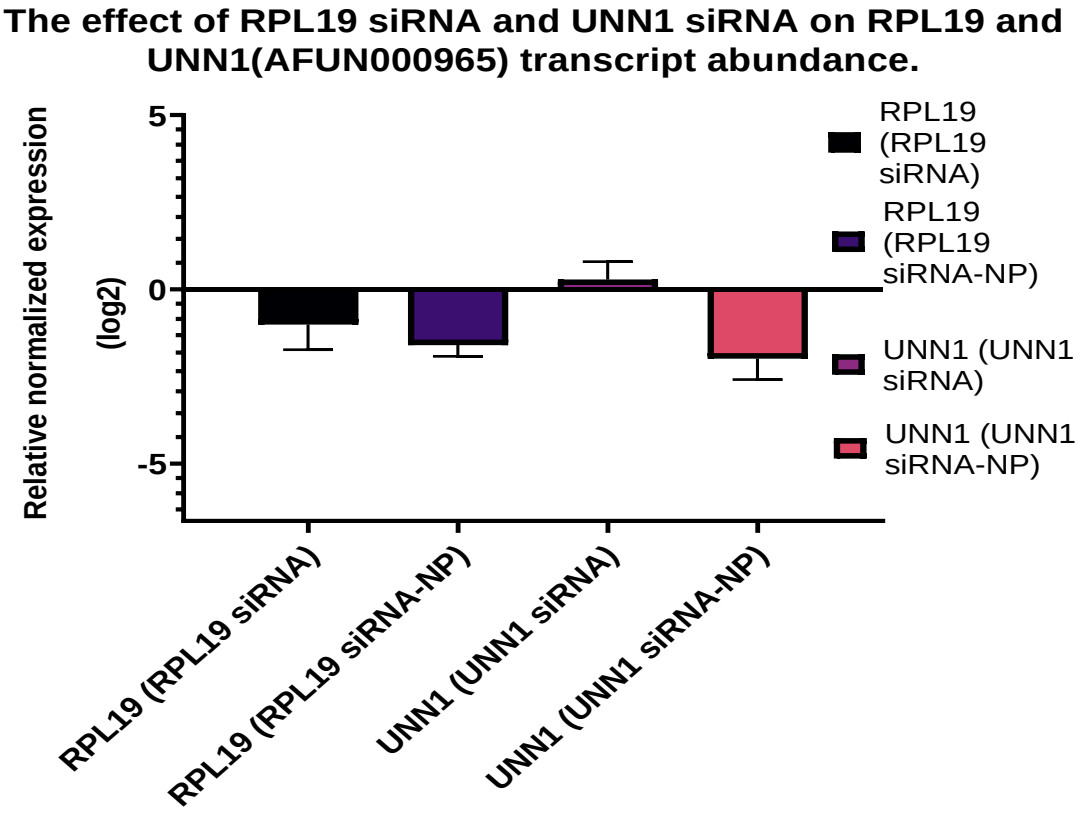


Figure 5.4 Relative Normalized Expression (RNE) of the RPL19 and UNN1 transcripts in *An. funestus* males treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP.

Table 5.6 RNE of the RPL19 and UNN1 (AFUN000965) transcript in *An. funestus* males treated with RPL19 siRNA, RPL19 siRNA-NP, UNN1 siRNA and UNN1 siRNA-NP when compared to samples treated with an exogenous negative control

Target	Relative Normalized Expression of qPCR (log2)	P value of qPCR RNE
Mβ2M siRNA	N/A	N/A
Mβ2M siRNA-NP	N/A	N/A
RPL19 (RPL19 siRNA)	**-1.003**	**0.007879**
RPL19 (RPL19 siRNA-NP)	**-1.596**	**0.015630**
UNN1 (UNN1 siRNA)	**0.281**	**0.014782**
UNN1 (UNN1 siRNA-NP)	**-1.984**	**0.026771**

The transcripts were relative to the negative exogenous control group (Mβ2M siRNA and Mβ2M siRNA-NP), which is set to one. The results were normalized to the reference genes RPS7 and RPS26. All samples P value ≤ 0.05 using unpaired student t-test.

Green = Up-regulation in expression
Blue = Down-regulation in expression
RNE = *Relative normalized expression*

5.5 Discussion

The RPL19 and UNN1 siRNA treatment of *An. funestus* males successfully silenced the RPL19 and UNN1 (AFUN000965) transcripts expression in the males. Observations of the male accessory organ development, genitalia rotation and erection of the antennal fibrillae were, however, normal when compared to the untreated negative control, indicating that there were no phenotypic changes observed either behaviourally or physiologically, regarding male sexual maturity, after the 24-hour period. These observations, i.e., that the expected changes did not occur, are a limitation of the study and it is proposed that future studies include the measurements of the testes size and the width of the MAG clear zone.

A number of possible reasons could be proposed as to why there was no observed phenotypic effect after treatment with RPL19 and UNN1 siRNA, even though silencing of the RPL9 and UNN1 (AFUN000965) transcripts occurred. One possible reason is that the silencing of the transcripts had not as yet been realized or translated on a phenotypic level. This could be a limitation of the study, and future studies should make continuous phenotypic observations for 24 h after the samples are treated with siRNA. However, since male maturation occurs within 36 h post ecdysis, the silencing was expected to have been realized or translated to a phenotypic level within 36 h of being treated with the siRNA. The observation that there was no phenotypic effect could also indicate that the UNN1 (AFUN000965) transcript does not play a role in the male sexual maturity and is involved with another developmental role. There is also a possibility that the RPL19 and UNN1 (AFUN000965) transcripts are not gender related and could be involved in both female and male mosquito development. The RPL19 transcript expression levels were used as an endogenous positive control and through real-time qPCR showed that the siRNA delivery and treatment was successful.

The qPCR results indicate that the delivery of the Mβ2M siRNA did not affect the expression of the RPL19 and UNN1 (AFUN000965) transcripts. A comparison was made to indicate any statistically significant difference in expression of the RPL19 and UNN1 (AFUN000965) transcripts between the untreated negative control and samples treated with exogenous siRNA (Mβ2M). The results in Section 5.4.1 then showed the difference in expression when the exogenous Mβ2M siRNA was conjugated to nanoparticles as compared to the untreated-NP. No statistically relevant differential expression was noted, indicating that the conjugation of

nanoparticles did not cause a change in expression in the RPL19 and UNN1 (AFUN000965) transcription levels.

The effect of the chitosan nanoparticles conjugated to the RPL19 and UNN1 siRNA (RPL19 siRNA-NP and UNN1 siRNA-NP) treatments of the *An. funestus* males clearly had an increased silencing effect of the RPL19 and the UNN1 (AFUN000965) transcripts, in the males. The RPL19 transcript had an increase in silencing by a log2 fold change of 0.593 when treated with RPL19 siRNA conjugated to chitosan nanoparticles compared to when only treated with RPL19 siRNA. Whereas the UNN1 (AFUN000965) transcript had an increase of silencing by 1.703 log2 fold-change when treated with UNN1 siRNA conjugated to chitosan nanoparticles when compared to samples only treated with UNN1 siRNA.

An increased silencing of the UNN1 siRNA-NP could not be stated with certainty, as the measuring of the increased silencing is measured by comparing the difference in the log2 fold change of the samples treated with siRNA not conjugated to nanoparticles to those that have been conjugated to nanoparticles. In this study, however, the samples treated with UNN1 siRNA did not show that silencing took place when compared to samples treated with the exogenous Mβ2M siRNA. A possible explanation could be that the male *An. funestus* mosquitoes did not feed on the siRNA, which may happen when working with live biological material. Another reason silencing was not observed in the sample treated with UNN1 siRNA is that the treatment was possibly contaminated by RNases either in the delivery phase or within the sample itself. The RNases could have destroyed the UNN1 siRNA before inducing a silencing effect on the UNN1 (AFUN000965) transcript. The possible destruction of the UNN1 siRNA by the RNASES could be a hypothesis to explain why silencing was not observed in the sample treated with UNN1 siRNA which is supported by the increased silencing of the UNN1 (AFUN000965) transcript when treated with the same siRNA but conjugated with chitosan nanoparticles.

Chitosan nanoparticles conjugated to the siRNA prevents the RNases from cleaving or destroying the siRNA (Zang *et al.*, 2010). The siRNA is then released over time by the chitosan nanoparticles whereby the siRNA can trigger the RNAi pathway and silence the targeted mRNA (Zang *et al.*, 2010). Chitosan nanoparticles could be offering protection to the siRNA, which is possibly the cause of the increased silencing of the siRNA when conjugated to chitosan nanoparticles. An increase in silencing of siRNA by conjugating chitosan nanoparticles to the

siRNA could be a useful tool in further RNAi knock-down studies in elucidating functional transcripts involved in important biological pathways, such as male sexual maturity.

Very few publications on the use of RNAi with the conjugation of chitosan nanoparticles on the anopheline species have been published. These were all performed on the *An. gambiae* species (Zhang *et al.*, 2010; Regna *et al.*, 2016; Taracena *et al.*, 2019). The conjugation of chitosan nanoparticles to the siRNA was shown to significantly increase the silencing effect of the siRNA in the *An. funestus* species. Allowing for the use of this powerful RNAi delivery tool in future studies of the functional genomics of the *An. funestus* mosquito.

6 Conclusions and future research

The severe disease burden of malaria, with its concomitant economic and social costs to Africa, remains unacceptable. Mechanisms such as insecticides to control the mosquito vectors need continuous renewal, particularly because of the ever-increasing insecticide resistance of malaria-carrying mosquitoes. Vector control methods traditionally target the female mosquitoes that are the main carriers of the disease. Very little is currently known about the biology of the *An. funestus* male mosquito. Specifically, the lack of knowledge on the molecular mechanisms involved in male sexual maturation is preventing research expanding into this field. In this study, both field work on swarming behaviour at a site in Zambia and functional genomics was used to initiate studies in understanding male sexual maturation in the main African malaria vector, *An. funestus*.

This study is based on two main hypotheses, namely, that immature *Anopheles funestus* males that have not undergone genitalia rotation will be unable to mate and will therefore not participate in swarming activities, and that a gene expression change occurs in the male *An. funestus* which leads to the rotation of the genitalia.

The studies on male swarms reported here, indicates that over 97% of the swarms composition contain sexually mature males and that sexually immature males do not participate in swarm formation (Zawada *et al*., 2018, and Chapter 2 of this thesis). The description in Chapter 2 is the first study that has molecularly identified *An. funestus* swarms to a species level. This study has also provided considerable insight into *An. funestus* swarm composition. This study has shown that the male *An. funestus* mosquito undergoes a physiological change (genitalia rotation) along with an associated behavioural change (swarming habits). The molecular mechanism that causes this physiological change is then further investigated to increase the understanding of male mosquitoes' mating biology.

The gene expression changes that occur in the male *An. funestus* which leads to the rotation of the genitalia was studied using functional genomic techniques.

RNA-sequencing of *An. funestus* Stage 3 males and comparison of the transcriptome to Stage 0 *An. funestus* males has provided hundreds of transcripts that could be involved in the male's sexual maturation. A large amount of data has been generated from the RNA-seq data and has provided a platform for future research. However, a more in-depth study of the RNA-seq data was out of the scope of this study. The down regulation of the CPRR1 (AFUN020892) transcript could indicate the possible softening of the endocuticle, in order to allow for the terminalia of the genitalia to fully rotate to sexual maturity. The biological function of CPRR1 in male *An. funestus* needs to be determined before any final conclusions can be drawn. The limited genomic information on *An. funestus* was challenging for this study and highlights the importance of improving whole genome sequencing information.

The qPCR studies performed in Chapter 4 confirmed the RNA-sequencing results from Chapter 3, by validating the six transcripts that were considered putatively involved with sexual maturation. Validation of these transcripts has shown that the RNA-seq data is robust and that the data can be used for future research. These selected validated transcripts could play a putative role in the sexual maturity of the male *An. funestus* mosquito due to their differential expression between Stage 0 and Stage 3. Further molecular techniques have to be undertaken to confirm this hypothesis.

The siRNA study did not indicate that the UNN1(AFUN000965) transcript, when silenced in the *An. funestus* males, had a role in genital rotation or affect the antennae fibrillae. Further studies can, however, be performed to confirm whether UNN1 is involved with other male sexual maturation events. These studies could include observing the knock-down effects over a longer period as well as determining a more overall expression analysis of the UNN1 (AFUN000965) transcript across all the stages of sexual development.

Conjugation of chitosan nanoparticles to siRNA as a knock-down technique has been evaluated and optimized for the *An. funestus* species. This technique of conjugating siRNA to chitosan nanoparticles increased the effect of siRNA knock-down, most likely due to the protection that the chitosan nanoparticles afford the siRNA from RNases (Zhang *et al.*, 2010). Future studies to elucidate putative gene functions in the *An. funestus* species could be based on the technique of

chitosan nanoparticle conjugation to siRNA. This is the first study, at the time of writing, that uses chitosan nanoparticles in conjunction with siRNA on the *An. funestus* species.

Males of the *An. funestus* species have been understudied for years because they are not directly involved in disease transmission. The importance of elucidating the functional transcripts of the male *An. funestus* is emphasised by this study. Dahalan *et al.* (2019) has shown recently that male anopheline mosquitoes can affect female susceptibility to *Plasmodium falciparum* by inducing changes in the female midgut which indicates that males do play a role in parasite transmission in previously unsuspected ways. Disturbance of the mating process will most likely directly impact on population reproductive success. Vector control strategies that target males may, therefore, have additional benefits in achieving the reduction of transmission. The lack of knowledge of the molecular mechanisms involved in male sexual maturation has until now prevented research expanding into this field. This study aimed to make a contribution in the area of male sexual maturation in mosquitoes, with the ultimate view of supporting new studies in transmission reduction.

Research performed and described in this thesis has significantly increased our understanding of the male *An. funestus* mosquito. More in-depth functional genomic studies of male anophelines were initiated. However, as indicated throughout this discussion, more research is needed for a better understanding of the male mosquito, and specifically its sexual maturation, as well as its mating behaviour in swarming, in order to contribute to novel control mechanisms of the vector. Such further studies include: more field studies on the swarming behaviour of the *An. funestus* group, and the investigation of mixed swarming and inter clade mating; further studies on specific transcripts that might be involved in genitalia rotation and male sexual maturation; bioinformatic analysis such as annotating previously unannotated transcripts, targeting the suite of transcripts for control of the sexual maturation and the mating process of anopheline males; optimization of techniques successful on *An. gambiae* for use on *An. funestus.*

Continually increasing our fundamental knowledge of the functional genomics of the male mosquito is one of the most important future tools in developing novel approaches and techniques to control the *An. funestus* mosquito as one of the main vectors of malaria in Africa.

7 References

1. Ageep, T.B., Damiens, D., Alsharif, B., Ahmed, A., Salih, E.H., Ahmed, F.T., Diabaté, A., Lees, R.S., Gilles, J.R. and El Sayed, B.B., 2014. Participation of irradiated *Anopheles arabiensis* males in swarms following field release in Sudan. *Malaria Journal*, *13*(1), pp.1-11.
2. Agnihotri, S.A., Mallikarjuna, N.N. and Aminabhavi, T.M., 2004. Recent advances on chitosan-based micro-and nanoparticles in drug delivery. *Journal of Controlled Release*, *100*(1), pp.5-28.
3. Agrawal, N., Dasaradhi, P.V.N., Mohmmed, A., Malhotra, P., Bhatnagar, R.K. and Mukherjee, S.K., 2003. RNA interference: biology, mechanism, and applications. *Microbiology and Molecular Biology Reviews*, *67*(4), pp.657-685.
4. Akey, D.H. and Jones, J.C., 1968. Coitus by Aedes aegypti males with unrotated and partially rotated terminalia. *Mosquito News*, *28*(4).
5. Andersen, S.O., Hojrup, P. and Roepstorff, P., 1995. Insect cuticular proteins. *Insect Biochemistry and Molecular Biology*, *25*(2), pp.153-176.
6. Anderson, P., Hallberg, E. and Subchev, M., 2000. Morphology of antennal sensilla auricillica and their detection of plant volatiles in the Herald moth, Scoliopteryx libatrix L. (Lepidoptera: Noctuidae). *Arthropod Structure and Development*, *29*(1), pp.33-41.
7. Andersen, C.L., Jensen, J.L. and Ørntoft, T.F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, *64*(15), pp.5245-5250.
8. Arts, G.J., Langemeijer, E., Tissingh, R., Ma, L., Pavliska, H., Dokic, K., Dooijes, R., Mešić, E., Clasen, R., Michiels, F. and van der Schueren, J., 2003. Adenoviral vectors expressing siRNAs for discovery and validation of gene function. *Genome Research*, *13*(10), pp.2325-2332.
9. Askling, H.H., Nilsson, J., Tegnell, A., Janzon, R. and Ekdahl, K., 2005. Malaria risk in travelers. *Emerging Infectious Diseases*, *11*(3), p.436.
10. Asmann, Y.W., Klee, E.W., Thompson, E.A., Perez, E.A., Middha, S., Oberg, A.L., Therneau, T.M., Smith, D.I., Poland, G.A., Wieben, E.D. and Kocher,

J.P.A., 2009. 3'tag digital gene expression profiling of human brain and universal reference RNA using Illumina Genome Analyzer. *BMC Genomics*, *10*(1), pp.1-11.

11. Ayala, D., Caro-Riaño, H., Dujardin, J.P., Rahola, N., Simard, F. and Fontenille, D., 2011. Chromosomal and environmental determinants of morphometric variation in natural populations of the malaria vector *Anopheles funestus* in Cameroon. *Infection, Genetics and Evolution*, *11*(5), pp.940-947.
12. Balabanidou, V., Grigoraki, L. and Vontas, J., 2018. Insect cuticle: a critical determinant of insecticide resistance. *Current Opinion in Insect Science*, *27*, pp.68-74.
13. Barrón, M.G., Paupy, C., Rahola, N., Akone-Ella, O., Ngangue, M.F., Wilson-Bahun, T.A., Pombi, M., Kengne, P., Costantini, C., Simard, F. and González, J., 2019. A new species in the major malaria vector complex sheds light on reticulated species evolution. *Scientific Reports*, *9*(1), pp.1-13.
14. Benedict, M.Q. and Robinson, A.S., 2003. The first releases of transgenic mosquitoes: an argument for the sterile insect technique. *Trends in Parasitology*, *19*(8), pp.349-355.
15. Benelli, G. and Beier, J.C., 2017. Current vector control challenges in the fight against malaria. *Acta Tropica*, *174*, pp.91-96.
16. Berry-Cabán, C.S., 2011. DDT and silent spring: fifty years after. *Journal of Military and Veterans Health*, *19*, pp.19-24.
17. Bio-Rad L. 2006. Real-Time PCR Applications Guide. Bio-Rad Laboratories.
18. Blandin, S.A., Wang-Sattler, R., Lamacchia, M., Gagneur, J., Lycett, G., Ning, Y., Levashina, E.A. and Steinmetz, L.M., 2009. Dissecting the genetic basis of resistance to malaria parasites in *Anopheles gambiae*. *Science*, *326*(5949), pp.147-150.
19. Boguski, M.S., Tolstoshev, C.M. and Bassett Jr, D.E., 1994. Gene discovery in dbEST. *Science*, *265*(5181), pp.1993-1994.
20. Bonizzoni, M., Afrane, Y., Dunn, W. A., Atieli, F. K., Zhou, G., Zhong, D., and Yan, G., 2012. Comparative transcriptome analyses of deltamethrin-resistant and-susceptible *Anopheles gambiae* mosquitoes from Kenya by RNA-Seq. *PloS One*, *7*(9), pp. 44607.
21. Bonizzoni, M., Ochomo, E., Dunn, W. A., Britton, M., Afrane, Y., Zhou, G., and Yan, G., 2015. RNA-seq analyses of changes in the *Anopheles gambiae*

transcriptome associated with resistance to pyrethroids in Kenya: identification of candidate-resistance genes and candidate-resistance SNPs. *Parasites and Vectors*, *8*(1), pp.1-13.

22. Bouyer, J. and Lefrançois, T., 2014. Boosting the sterile insect technique to control mosquitoes. *Trends in Parasitology*, *30*(6), pp.271-273.
23. Boyce, R., 1911. The history of yellow fever in West Africa. *British Medical Journal*, *1*(2615), pp.301.
24. Brogdon, W.G. and McAllister, J.C., 1998. Simplification of adult mosquito bioassays through use of time-mortality determinations in glass bottles. *Journal of the American Mosquito Control Association*, *14*(2), pp.159-164.
25. Brooke, B.D., Robertson, L., Kaiser, M.L., Raswiswi, E., Munhenga, G., Venter, N., Wood, O.R. and Koekemoer, L.L., 2015. Insecticide resistance in the malaria vector *Anopheles arabiensis* in Mamfene, KwaZulu-Natal. *South African Journal of Science*, *111*(11-12), pp.1-3.
26. Brown, A.W., 1986. Insecticide resistance in mosquitoes: a pragmatic review. *Journal of the American Mosquito Control Association*, *2*(2), pp.123-140.
27. Brown, A. E., Bugeon, L., Crisanti, A., and Catteruccia, F., 2003. Stable and heritable gene silencing in the malaria vector *Anopheles stephensi*. *Nucleic Acids Research*, *31*(15), pp.85-e85.
28. Brunhes, J., LeGoff, G. and Geoffroy, B., 1997, June. Afrotropical Anopheles. 1. Description of new species and modifications of taxonomical status (Diptera: Culicidae). In *Annales de la Societe Entomologique de France*, 33(2), pp. 173-183.
29. Bryant, W. B., Mills, M. K., Olson, B. J., and Michel, K., 2019. Small rna-seq analysis reveals mirna expression dynamics across tissues in the malaria vector, *Anopheles gambiae*. *G3: Genes, Genomes, Genetics*, *9*(5), pp. 1507-1517.
30. Bustin, S.A., 2010. Why the need for qPCR publication guidelines?—The case for MIQE. *Methods*, *50*(4), pp.217-226.
31. Bustin, S. and Huggett, J., 2017. qPCR primer design revisited. *Biomolecular Detection and Quantification*, *14*, pp.19-28.
32. Bustin, S. and Nolan, T., 2017. Talking the talk, but not walking the walk: RT-qPCR as a paradigm for the lack of reproducibility in molecular research. *European Journal of Clinical Investigation*, *47*(10), pp.756-774.

33. Bustin, S., 2017. The continuing problem of poor transparency of reporting and use of inappropriate methods for RT-qPCR. *Biomolecular Detection and Quantification*, *12*, pp.7-9.
34. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L. and Vandesompele, J., 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, 55(4), pp.611–622.
35. Callis, J., Carpenter, T., Sun, C.W. and Vierstra, R.D., 1995. Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in Arabidopsis thaliana ecotype Columbia. *Genetics*, *139*(2), pp.921-939.
36. Catteruccia, F. and Levashina, E.A., 2009. RNAi in the malaria vector, *Anopheles gambiae.* In *Therapeutic Applications of RNAi* (pp. 63-75). Humana Press.
37. Center for Disease Control. https://www.cdc.gov/. Last accessed 6 January 2021.
38. Chambers, G.M. and Klowden, M.J., 2001. Age of *Anopheles gambiae* Giles male mosquitoes at time of mating influences female oviposition. *Journal of Vector Ecology*, *26*(2), pp.196-201.
39. Charles J.P., 2010. The regulation of expression of insect cuticle protein genes. *Insect Biochemistry and Molecular Biology*, 40(3), pp.205-13.
40. Charlwood, J.D. and Jones, M.D.R., 1979. Mating behaviour in the mosquito, *Anopheles gambiae s. 1*. save: I. Close range and contact behaviour. *Physiological Entomology*, *4*(2), pp.111-120.
41. Charlwood, J.D., Thompson, R. and Madsen, H., 2003. Observations on the swarming and mating behaviour of *Anopheles funestus* from southern Mozambique. *Malaria Journal*, *2*(1), pp.1-10.
42. Chervoneva, I., Li, Y., Schulz, S., Croker, S., Wilson, C., Waldman, S.A. and Hyslop, T., 2010. Selection of optimal reference genes for normalization in quantitative RT-PCR. *BMC Bioinformatics*, *11*(1), pp.1-15.
43. Chevone, B.I. and Richards, A.G., 1976. Ultrastructure of the atypic muscles associated with terminalial inversion in male Aedes aegypti (L). *The Biological Bulletin*, *151*(2), pp.283-296.
44. Chima, R.I., Goodman, C.A. and Mills, A., 2003. The economic impact of malaria in Africa: a critical review of the evidence. *Health Policy*, *63*(1), pp.17-36.

45. Choi, K.S., 2015. Diversity of clades I and II within *Anopheles funestus* sensu stricto in southern Africa. *Genes and Genomics*, *37*(2), pp.193-197.
46. Choi, K.S., Coetzee, M. and Koekemoer, L.L., 2013. Detection of clade types (clades I and II) within *Anopheles funestus sensu stricto* by the hydrolysis probe analysis (Taqman assay). *Parasites and Vectors*, *6*(1), pp.1-5.
47. Chuang, C.F. and Meyerowitz, E.M., 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, *97*(9), pp.4985-4990.
48. Clements, A.N., 1992. *The biology of mosquitoes: development, nutrition and reproduction* (Vol. 1, p. 509). London: Chapman & Hall.
49. Clements, A.N., 1999. *The biology of mosquitoes. Volume 2: sensory reception and behaviour*. CABI publishing.
50. Coats, J.R., 1990. Mechanisms of toxic action and structure-activity relationships for organochlorine and synthetic pyrethroid insecticides. *Environmental Health Perspectives*, *87*, pp.255-262.
51. Cobb, M., 2017. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biology*, *15*(9), pp. 2003243.
52. Coetzee, M., 2020. Key to the females of Afrotropical *Anopheles* mosquitoes (Diptera: Culicidae). *Malaria Journal*, *19*(1), pp.1-20.
53. Coetzee, M. and Fontenille, D., 2004. Advances in the study of *Anopheles funestus*, a major vector of malaria in Africa. *Insect Biochemistry and Molecular Biology*, *34*(7), pp.599-605.
54. Coetzee, M. and Koekemoer, L.L., 2013. Molecular systematics and insecticide resistance in the major African malaria vector *Anopheles funestus. Annual Review of Entomology*, *58*, pp.393-412.
55. Coetzee, M., Hunt, R.H., Wilkerson, R., Della Torre, A., Coulibaly, M.B. and Besansky, N.J., 2013a. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the Anopheles gambiae complex. *Zootaxa*, *3619*(3), pp.246-274.
56. Coetzee, M., Kruger, P., Hunt, R.H., Durrheim, D.N., Urbach, J. and Hansford, C.F., 2013b. Malaria in South Africa: 110 years of learning to control the disease. *South African Medical Journal*, *103*(10), pp.770-778.
57. Cohuet, A., Simard, F., Toto, J.C., Kengne, P., Coetzee, M. and Fontenille, D., 2003. Species identification within the *Anopheles funestus* group of malaria

vectors in Cameroon and evidence for a new species. *The American Journal of Tropical Medicine and Hygiene*, *69*(2), pp.200-205.

58. Collins, F.H. and Paskewitz, S.M., 1995. Malaria: current and future prospects for control. *Annual Review of Entomology*, *40*(1), pp.195-219.

59. Coluzzi, M., Petrarca, V. and di Deco, M.A., 1985. Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. *Italian Journal of Zoology*, *52*(1-2), pp.45-63.

60. Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X. and Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. *Genome Biology*, *17*(1), pp.1-19.

61. Da Silva, J.J., Mendes, J. and Lomônaco, C., 2009. Effects of sublethal concentrations of diflubenzuron and methoprene on *Aedes aegypti* (Diptera: Culicidae) fitness. *International Journal of Tropical Insect Science*, *29*(1), pp.17-23.

62. Dabiré, K.R., Sawadogo, P.S., Hien, D.F., Bimbilé-Somda, N.S., Soma, D.D., Millogo, A., Baldet, T., Gouagna, L.C., Simard, F., Lefèvre, T. and Diabaté, A., 2014. Occurrence of natural *Anopheles arabiensis* swarms in an urban area of Bobo-Dioulasso city, Burkina Faso, West Africa. *Acta Tropica*, *132*, pp.S35-S41.

63. Dahalan, F.A., Churcher, T.S., Windbichler, N. and Lawniczak, M.K., 2019. The male mosquito contribution towards malaria transmission: Mating influences the *Anopheles* female midgut transcriptome and increases female susceptibility to human malaria parasites. *PLoS Pathogens*, *15*(11), pp.e1008063.

64. Dahan, Y.L. and Koekemoer, L.L., 2014. Analysis of the genitalia rotation in the male *Anopheles funestus* (Diptera: Culicidae). *Acta Tropica*, *132*, pp. S20-S25.

65. Dahan-Moss, Y., Hendershot, A., Dhoogra, M., Julius, H., Zawada, J., Kaiser, M., Lobo, N.F., Brooke, B.D. and Koekemoer, L.L., 2020. Member species of the *Anopheles gambiae* complex can be misidentified as *Anopheles leesoni*. *Malaria Journal*, *19*(1), pp.1-9.

66. Dahan-Moss, Y.L. and Koekemoer, L.L., 2016. Analysis of esterase enzyme activity in adults of the major malaria vector *Anopheles funestus*. *Parasites and Vectors*, *9*(1), pp.1-8.

67. Dana, H., Chalbatani, G.M., Mahmoodzadeh, H., Karimloo, R., Rezaiean, O., Moradzadeh, A., Mehmandoost, N., Moazzen, F., Mazraeh, A., Marmari, V. and Ebrahimi, M., 2017. Molecular mechanisms and biological functions of siRNA. *International Journal of Biomedical Science*, *13*(2), pp.48.
68. Davis, E.E., 1976. A receptor sensitive to oviposition site attractants on the antennae of the mosquito, Aedes aegypti. *Journal of Insect Physiology*, *22*(10), pp.1371-1376.
69. De Marco, L., Sassera, D., Epis, S., Mastrantonio, V., Ferrari, M., Ricci, I., and Urbanelli, S., 2017. The choreography of the chemical defensome response to insecticide stress: insights into the *Anopheles stephensi* transcriptome using RNA-Seq. *Scientific Reports*, *7*(1), pp.1-12.
70. DeMeillon, B., 1986. The control of malaria with special reference to the contributions made by the staff of the South African Institute for Medical Research. *South African Medical Journal*, *70*(4), pp.67-69.
71. Della Torre, A., Costantini, C., Besansky, N.J. and Caccone, A., 2002. Speciation within *Anopheles gambiae*--the glass is half full. *Science*, *298*(5591), pp.115.
72. Department of Health, unpublished data, KZN
73. Derveaux, S., Vandesompele, J. and Hellemans, J., 2010. How to do successful gene expression analysis using real-time PCR. *Methods*, *50*(4), pp.227-230.
74. Diabaté, A., Brengues, C., Baldet, T., Dabire, K.R., Hougard, J.M., Akogbeto, M., Kengne, P., Simard, F., Guillet, P., Hemingway, J. and Chandre, F., 2004. The spread of the Leu-Phe kdr mutation through *Anopheles gambiae* complex in Burkina Faso: genetic introgression and de novo phenomena. *Tropical Medicine and International Health*, *9*(12), pp.1267-1273.
75. Diabaté, A., Dabire, R.K., Kengne, P., Brengues, C., Baldet, T., Ouari, A., Simard, F. and Lehmann, T., 2006. Mixed swarms of the molecular M and S forms of *Anopheles gambiae* (Diptera: Culicidae) in sympatric area from Burkina Faso. *Journal of Medical Entomology*, *43*(3), pp.480-483.
76. Diabaté, A., Dao, A., Diallo, M., Huestis, D.L. and Lehmann, T., 2011. Spatial distribution and male mating success of *Anopheles gambiae* swarms. *BMC Evolutionary Biology*, *11*(1), pp.1-11.
77. Diabaté, A., Dao, A., Yaro, A.S., Adamou, A., Gonzalez, R., Manoukis, N.C., Traoré, S.F., Gwadz, R.W. and Lehmann, T., 2009. Spatial swarm segregation and

reproductive isolation between the molecular forms of *Anopheles gambiae*. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1676), pp.4215-4222.

78. Dingle, H., 1972. Migration strategies of insects. *Science*, *175*(4028), pp.1327-1335.

79. Dorsett, Y. and Tuschl, T., 2004. siRNAs: applications in functional genomics and potential as therapeutics. *Nature Reviews Drug Discovery*, *3*(4), pp.318-329.

80. Durnez, L. and Coosemans, M., 2013. *Residual transmission of malaria: an old issue for new approaches*. InTech, pp. 671-704.

81. Dyck, V.A., Hendrichs, J. and Robinson, A.S., 2005. *Sterile insect technique: principles and practice in area-wide integrated pest management*. Taylor & Francis, pp. 1216.

82. Favia, G., Della Torre, A., Bagayoko, M., Lanfrancotti, A., Sagnon, N.F., Touré, Y.T. and Coluzzi, M., 1997. Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. *Insect Molecular Biology*, *6*(4), pp.377-383.

83. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), pp.806-811.

84. Foster, W.A. and Lea, A.O., 1975. Renewable fecundity of male Aedes aegypti following replenishment of seminal vesicles and accessory glands. *Journal of Insect Physiology*, *21*(5), pp.1085-1090.

85. Fukuto, T.R., 1990. Mechanism of action of organophosphorus and carbamate insecticides. *Environmental Health Perspectives*, *87*, pp.245-254.

86. Gantz, V.M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V.M., Bier, E. and James, A.A., 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proceedings of the National Academy of Sciences*, *112*(49), pp.e6736-e6743.

87. Garnham, P.C.C., 1966. Malaria parasites and other Haemosporidia. *Malaria Parasites and other Haemosporidia.* Blackwell Scientific Publications Ltd., Oxford, pp. xviii + 1114.

88. Garros, C., Koekemoer, L.L., Kamau, L., Awolola, T.S., Van Bortel, W., Coetzee, M., Coosemans, M. and Manguin, S., 2004. Restriction fragment length polymorphism method for the identification of major African and Asian malaria vectors within the *Anopheles funestus* and *An. minimus* groups. *The American Journal of Tropical Medicine and Hygiene*, *70*(3), pp.260-265.

89. Gatton, M.L., Martin, L.B. and Cheng, Q., 2004. Evolution of resistance to sulfadoxine-pyrimethamine in Plasmodium falciparum. *Antimicrobial Agents and Chemotherapy*, *48*(6), pp.2116-2123.

90. Gentile, G., Slotman, M., Ketmaier, V., Powell, J.R. and Caccone, A., 2001. Attempts to molecularly distinguish cryptic taxa in *Anopheles gambiae* ss. *Insect Molecular Biology*, *10*(1), pp.25-32.

91. Gerhard, D.S., Wagner, L., Feingold, E.A., Shenmen, C.M., Grouse, L.H., Schuler, G., Klein, S.L., Old, S., Rasooly, R., Good, P. and Guyer, M., 2004. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Research*, *14*(10 B), pp.2121-2127.

92. Giles, G.M.J., 1902. *A Handbook of the Gnats Or Mosquitoes: Giving the Anatomy and Life History of the Culicidæ Together with Descriptions of All Species Noticed Up to the Present Date*. John Bale, Sons & Company.

93. Gillies, M.T. and Coetzee, M., 1987. A supplement to the Anophelinae of Africa South of the Sahara. *Publication of the South African Institute for Medical Research*, *55*, pp.1-143.

94. Gillies, M.T. and DeMeillon, B., 1968. The Anophelinae of Africa south of the Sahara (Ethiopian zoogeographical region). *The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region)*. pp.343.

95. Govere, J., Durrheim, D., la Grange, K., Mabuza, A. and Booman, M., 2000. Community knowledge and perceptions about malaria and practices influencing malaria control in Mpumalanga Province, South Africa. *South African Medical Journal*, *90*(6), pp.611-618.

96. Griffith, M., Griffith, O.L., Mwenifumbo, J., Goya, R., Morrissy, A.S., Morin, R.D., Corbett, R., Tang, M.J., Hou, Y.C., Pugh, T.J. and Robertson, G., 2010. Alternative expression analysis by RNA sequencing. *Nature Methods*, *7*(10), pp.843.

97. Hackett, B.J., Gimnig, J., Guelbeogo, W., Costantini, C., Koekemoer, L.L., Coetzee, M., Collins, F.H. and Besansky, N.J., 2000. Ribosomal DNA internal transcribed spacer (ITS2) sequences differentiate *Anopheles funestus* and *An. rivulorum*, and uncover a cryptic taxon. *Insect Molecular Biology*, *9*(4), pp.369-374.

98. Hammond, A., Galizi, R., Kyrou, K., Simoni, A., Siniscalchi, C., Katsanos, D., Gribble, M., Baker, D., Marois, E., Russell, S. and Burt, A., 2016. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nature Biotechnology*, *34*(1), pp.78-83.

99. Hancock, P.A., Wiebe, A., Gleave, K.A., Bhatt, S., Cameron, E., Trett, A., Weetman, D., Smith, D.L., Hemingway, J., Coleman, M. and Gething, P.W., 2018. Associated patterns of insecticide resistance in field populations of malaria vectors across Africa. *Proceedings of the National Academy of Sciences*, *115*(23), pp.5938-5943.

100. Harbach, R.E., 2004. hypothesis of phylogenetic relationships. *Bulletin of Entomological Research*, *94*, pp.537-553.

101. Harbers, M. and Carninci, P., 2005. Tag-based approaches for transcriptome research and genome annotation. *Nature Methods*, *2*(7), pp.495-502.

102. Hargreaves, K., Koekemoer, L.L., Brooke, B.D., Hunt, R.H., Mthembu, J. and Coetzee, M., 2000. *Anopheles funestus* resistant to pyrethroid insecticides in South Africa. *Medical and Veterinary Entomology*, *14*(2), pp.181-189.

103. Harper, J.O., 1944. Notes on the Swarming of Males of A. funestus (Giles), in East Africa. *East African Medical Journal*, *21*(5).

104. Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.M., 1996. Real time quantitative PCR. *Genome Research*, *6*(10), pp.986-994.

105. Hellemans, J. and Vandesompele, J., 2014. Selection of reliable reference genes for RT-qPCR analysis. Humana Press, New York, NY. In *Quantitative Real-Time PCR*, pp. 19-26.

106. Hemingway, J., Ranson, H., Magill, A., Kolaczinski, J., Fornadel, C., Gimnig, J., Coetzee, M., Simard, F., Roch, D.K., Hinzoumbe, C.K. and Pickett, J., 2016. Averting a malaria disaster: will insecticide resistance derail malaria control?. *The Lancet*, *387*(10029), pp.1785-1788.

107. Hemingway, J., Vontas, J., Poupardin, R., Raman, J., Lines, J., Schwabe, C., Matias, A. and Kleinschmidt, I., 2013. Country-level operational implementation of the Global Plan for Insecticide Resistance Management. *Proceedings of the National Academy of Sciences*, *110*(23), pp.9397-9402.
108. Hendrichs, J., Vreysen, M.J.B., Enkerlin, W.R., Cayol, J.P., Dyck, V.A. and Robinson, A.S., 2005. Strategic options in using sterile insects for area-wide integrated pest management. *Sterile Insect*, pp.841.
109. Hershko, A., 1991. The ubiquitin pathway for protein degradation. *Trends in Biochemical Sciences*, *16*, pp.265-268.
110. Higuchi, R., Dollinger, G., Walsh, P.S. and Griffith, R., 1992. Simultaneous amplification and detection of specific DNA sequences. *Biotechnology*, *10*(4), pp.413-417.
111. Hoa, N.T., Keene, K.M., Olson, K.E. and Zheng, L., 2003. Characterization of RNA interference in an *Anopheles gambiae* cell line. *Insect Biochemistry and Molecular Biology*, *33*(9), pp.949-957.
112. Mehrtash, A. B., and Hochstrasser, M., 2019. Ubiquitin-dependent protein degradation at the endoplasmic reticulum and nuclear envelope. In *Seminars in cell & developmental biology* (Vol. 93, pp. 111-124). Academic Press.
113. Hougard, J.M., Duchon, S., Zaim, M. and Guillet, P., 2002. Bifenthrin: a useful pyrethroid insecticide for treatment of mosquito nets. *Journal of Medical Entomology*, *39*(3), pp.526-533.
114. Howell, P.I. and Knols, B.G., 2009. Male mating biology. *Malaria Journal*, *8*(2), pp.1-10.
115. Huho, B.J., Ng'habi, K.R., Killeen, G.F., Nkwengulila, G., Knols, B.G. and Ferguson, H.M., 2006. A reliable morphological method to assess the age of male *Anopheles gambiae*. *Malaria Journal*, *5*(1), pp.1-11.
116. Hunt, R.H., Brooke, B.D., Pillay, C., Koekemoer, L.L. and Coetzee, M., 2005. Laboratory selection for and characteristics of pyrethroid resistance in the malaria vector *Anopheles funestus*. *Medical and Veterinary Entomology*, *19*(3), pp.271-275.
117. Ibrahim, K.A., Eleyan, M., Abd El-Rahman, H.A., Khwanes, S.A. and Mohamed, R.A., 2020. Quercetin attenuates the oxidative injury–mediated upregulation of apoptotic gene expression and catecholaminergic neurotransmitters of the fetal

rats' brain following prenatal exposure to fenitrothion insecticide. *Neurotoxicity Research*, pp.1-12.

118. Irish, S.R., Kyalo, D., Snow, R.W. and Coetzee, M., 2020. Updated list of *Anopheles* species (Diptera: Culicidae) by country in the Afrotropical Region and associated islands. *Zootaxa*, *4747*(3).

119. Jensen, S., Gassama, M.P. and Heidmann, T., 1999. Cosuppression of I transposon activity in Drosophila by I-containing sense and antisense transgenes. *Genetics*, *153*(4), pp.1767-1774.

120. Jing, T.Z., Wang, Z.Y., Liu, K.Y. and Qi, F.H., 2006. A modified one-step procedure for rapid RNA isolation from insect. *Journal of Forestry Research*, *17*(2), pp.129-131.

121. Jones, L., Hamilton, A.J., Voinnet, O., Thomas, C.L., Maule, A.J. and Baulcombe, D.C., 1999. RNA–DNA interactions and DNA methylation in post-transcriptional gene silencing. *The Plant Cell*, *11*(12), pp.2291-2301.

122. Kaindoa, E.W., Ngowo, H.S., Limwagu, A.J., Tchouakui, M., Hape, E., Abbasi, S., Kihonda, J., Mmbando, A.S., Njalambaha, R.M., Mkandawile, G. and Bwanary, H., 2019. Swarms of the malaria vector *Anopheles funestus* in Tanzania. *Malaria Journal*, *18*(1), pp.1-10.

123. Kajla, M., Kakani, P., Choudhury, T.P., Kumar, V., Gupta, K., Dhawan, R., Gupta, L. and Kumar, S., 2017. *Anopheles stephensi* heme peroxidase HPX15 suppresses midgut immunity to support Plasmodium development. *Frontiers in Immunology*, *8*, pp.249.

124. Kelly-Hope, L., Ranson, H. and Hemingway, J., 2008. Lessons from the past: managing insecticide resistance in malaria control and eradication programmes. *The Lancet Infectious Diseases*, *8*(6), pp.387-389.

125. Kennerdell, J.R. and Carthew, R.W., 2000. Heritable gene silencing in Drosophila using double-stranded RNA. *Nature Biotechnology*, *18*(8), pp.896-898.

126. Killeen, G.F., 2014. Characterizing, controlling and eliminating residual malaria transmission. *Malaria Journal*, *13*(1), pp.1-22.

127. Klowden, M.J. and Beach, R., 1980. Large doses of ecdysterone may inhibit mosquito behavior nonspecifically. *Science, 208 (4447)*, pp.2062-1063.

128. Knols, B.G., Farenhorst, M., Andriessen, R., Snetselaar, J., Suer, R.A., Osinga, A.J., Knols, J.M., Deschietere, J., Ng'habi, K.R., Lyimo, I.N. and Kessy, S.T.,

2016. Eave tubes for malaria control in Africa: an introduction. *Malaria Journal*, *15*(1), pp.1-7.

129. Knox, T.B., Juma, E.O., Ochomo, E.O., Jamet, H.P., Ndungo, L., Chege, P., Bayoh, N.M., N'Guessan, R., Christian, R.N., Hunt, R.H. and Coetzee, M., 2014. An online tool for mapping insecticide resistance in major *Anopheles* vectors of human malaria parasites and review of resistance status for the Afrotropical region. *Parasites and Vectors*, *7*(1), pp.1-14.

130. Koekemoer, L.L., Coetzee, M. and Hunt, R.H., 1998. Hpall endonuclease distinguishes between two species in the *Anopheles funestus* group. *Insect Molecular Biology*, *7*(3), pp.273-277.

131. Koekemoer, L.L., Kamau, L., Garros, C., Manguin, S., Hunt, R.H. and Coetzee, M., 2006. Impact of the Rift Valley on restriction fragment length polymorphism typing of the major African malaria vector *Anopheles funestus* (Diptera: Culicidae). *Journal of Medical Entomology*, *43*(6), pp.1178-1184.

132. Koekemoer, L.L., Kamau, L., Hunt, R.H. and Coetzee, M., 2002. A cocktail polymerase chain reaction assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *The American Journal of Tropical Medicine and Hygiene*, *66*(6), pp.804-811.

133. Kyalo, D., Amratia, P., Mundia, C.W., Mbogo, C.M., Coetzee, M. and Snow, R.W., 2017. A geo-coded inventory of anophelines in the Afrotropical Region south of the Sahara: 1898-2016. *Wellcome Open Research*, *2*.

134. Kyrou, K., Hammond, A.M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A.K., Nolan, T. and Crisanti, A., 2018. A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nature Biotechnology*, *36*(11), pp.1062-1066.

135. Lees, R.S., Knols, B., Bellini, R., Benedict, M.Q., Bheecarry, A., Bossin, H.C., Chadee, D.D., Charlwood, J., Dabire, R.K., Djogbenou, L. and Egyir-Yawson, A., 2014. Improving our knowledge of male mosquito biology in relation to genetic control programmes. *Acta Tropica*, *132*, pp.s2-s11.

136. Letinić, B. D., Dahan-Moss, Y., & Koekemoer, L. L., 2020. Characterising the effect of Akirin knockdown on Anopheles arabiensis (Diptera: Culicidae) reproduction and survival, using RNA-mediated interference. *PloS One*, *15*(2), pp. 0228576.

137. Lewis, J.A., Wu, C.H., Levine, J.H. and Berg, H., 1980. Levamisole-resistant mutants of the nematode Caenorhabditis elegans appear to lack pharmacological acetylcholine receptors. *Neuroscience*, *5*(6), pp.967-989.
138. Li, H., Li, W.X. and Ding, S.W., 2002. Induction and suppression of RNA silencing by an animal virus. *Science*, *296*(5571), pp.1319-1321.
139. Li, J., Cai, C., Li, J., Li, J., Li, J., Sun, T., Wang, L., Wu, H. and Yu, G., 2018. Chitosan-based nanomaterials for drug delivery. *Molecules*, *23*(10), pp.2661.
140. Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods*, *25*(4), pp.402-408.
141. Maharaj, R., Mthembu, D.J. and Sharp, B.L., 2005. Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal. *South African Medical Journal*, *95*(11), pp.871-874.
142. Mahmood, F. and Reisen, W.K., 1982. *Anopheles stephensi (*Diptera: Culicidae): changes in male mating competence and reproductive system morphology associated with aging and mating. *Journal of Medical Entomology*, *19*(5), pp.573-588.
143. Manoukis, N.C., Diabate, A., Abdoulaye, A., Diallo, M., Dao, A., Yaro, A.S., Ribeiro, J.M. and Lehmann, T., 2009. Structure and dynamics of male swarms of *Anopheles gambiae*. *Journal of Medical Entomology*, *46*(2), pp.227-235.
144. Marchand, R.P., 1983. Field Observations On Swarming and Mating in *Anopheles gambiae* Mosquitoes in Tanzania1. *Netherlands Journal of Zoology*, *34*(3), pp.367-387.
145. McGettigan, P.A., 2013. Transcriptomics in the RNA-seq era. *Current Opinion in Chemical Biology*, *17*(1), pp.4-11.
146. Mezquita, J., Pau, M. and Mezquita, C., 1997. Characterization and expression of two chicken cDNAs encoding ubiquitin fused to ribosomal proteins of 52 and 80 amino acids. *Gene*, *195*(2), pp.313-319.
147. Michel, A.P., Ingrasci, M.J., Schemerhorn, B.J., Kern, M., Le Goff, G., Coetzee, M., Elissa, N., Fontenille, D., Vulule, J., Lehmann, T. and Sagnon, N.F., 2005. Rangewide population genetic structure of the African malaria vector *Anopheles funestus*. *Molecular Ecology*, *14*(14), pp.4235-4248.

148. Mitraka, E., Stathopoulos, S., Siden-Kiamos, I., Christophides, G.K. and Louis, C., 2013. Asaia accelerates larval development of *Anopheles gambiae*. *Pathogens and Global Health*, *107*(6), pp.305-311.
149. Mladenović, M., Arsić, B.B., Stanković, N., Mihović, N., Ragno, R., Regan, A., Milićević, J.S., Trtić-Petrović, T.M. and Mićić, R., 2018. The targeted pesticides as acetylcholinesterase inhibitors: Comprehensive cross-organism molecular modelling studies performed to anticipate the pharmacology of harmfulness to humans in vitro. *Molecules*, *23*(9), pp.2192.
150. Mongkol, W., Nguitragool, W., Sattabongkot, J., and Kubera, A., 2018. Blood-induced differential gene expression in *Anopheles dirus* evaluated using RNA sequencing. *Medical and Veterinary Entomology*, *32*(4), pp.399-406.
151. Moussian, B., 2013. The arthropod cuticle. In *Arthropod Biology and Evolution* (pp. 171-196). Springer, Berlin, Heidelberg.
152. Mozūraitis, R., Hajkazemian, M., Zawada, J.W., Szymczak, J., Pålsson, K., Sekar, V., Biryukova, I., Friedländer, M.R., Koekemoer, L.L., Baird, J.K. and Borg-Karlson, A.K., 2020. Male swarming aggregation pheromones increase female attraction and mating success among multiple African malaria vector mosquito species. *Nature Ecology and Evolution*, *4*(10), pp.1395-1401.
153. Nagpal, K., Singh, S.K. and Mishra, D.N., 2010. Chitosan nanoparticles: a promising system in novel drug delivery. *Chemical and Pharmaceutical Bulletin*, *58*(11), pp.1423-1430.
154. Nardini, L., Christian, R.N., Coetzer, N., Ranson, H., Coetzee, M. and Koekemoer, L.L., 2012. Detoxification enzymes associated with insecticide resistance in laboratory strains of *Anopheles arabiensis* of different geographic origin. *Parasites and Vectors*, *5*(1), pp.1-12.
155. Neafsey, D.E., Waterhouse, R.M., Abai, M.R., Aganezov, S.S., Alekseyev, M.A., Allen, J.E., Amon, J., Arcà, B., Arensburger, P., Artemov, G., Assour, L.A., 2015. Highly evolvable malaria vectors: the genomes of 16 *Anopheles* mosquitoes. *Science*, *347*, pp. 6217.
156. Negri, A., Ferrari, M., Nodari, R., Coppa, E., Mastrantonio, V., Zanzani, S., and Epis, S., 2019. Gene silencing through RNAi and antisense Vivo-Morpholino increases the efficacy of pyrethroids on larvae of *Anopheles stephensi*. *Malaria Journal*, *18*(1), pp.1-12.

157. Neville, A.C., 2012. Biology of the arthropod cuticle (Vol. 4). *Springer Science and Business Media*. New York.
158. Ng'habi, K.R., Huho, B.J., Nkwengulila, G., Killeen, G.F., Knols, B.G. and Ferguson, H.M., 2008. Sexual selection in mosquito swarms: may the best man lose?. *Animal Behaviour*, *76*(1), pp.105-112.
159. Nielsen, H.T., 1964. Swarming and some other habits of Mansonia perturbans and Psorophora ferox (Diptera: Culicidae). *Behaviour*, *24*(1-2), pp.67-88.
160. Nijhout, H.F., 1977. Control of antennal hair erection in male mosquitoes. *The Biological Bulletin*, *153*(3), pp.591-603.
161. Nolan, T., Papathanos, P., Windbichler, N., Magnusson, K., Benton, J., Catteruccia, F. and Crisanti, A., 2011. Developing transgenic *Anopheles* mosquitoes for the sterile insect technique. *Genetica*, *139*(1), pp.33-39.
162. Noriega, F.G. and Wells, M.A., 1993. A comparison of three methods for isolating RNA from mosquitoes. *Insect Molecular Biology*, *2*(1), pp.21-24.
163. O'Donnell, P.P. and Klowden, M.J., 1997. Methoprene affects the rotation of the male terminalia of Aedes aegypti mosquitoes. *Journal of the American Mosquito Control Association-Mosquito News*, *13*(1), pp.1-4.
164. Ogunah, J.A., Lalah, J.O. and Schramm, K.W., 2020. Malaria vector control strategies. What is appropriate towards sustainable global eradication?. *Sustainable Chemistry and Pharmacy*, *18*, pp.100339.
165. Oliva, C.F., Benedict, M.Q., Lempérière, G. and Gilles, J., 2011. Laboratory selection for an accelerated mosquito sexual development rate. *Malaria Journal*, *10*(1), pp.1-8.
166. Oliva, C.F., Jacquet, M., Gilles, J., Lempérière, G., Maquart, P.O., Quilici, S., Schooneman, F., Vreysen, M.J. and Boyer, S., 2012. The sterile insect technique for controlling populations of Aedes albopictus (Diptera: Culicidae) on Reunion Island: mating vigour of sterilized males. *PloS One*, *7*(11), pp.e49414.
167. Oliva, C.F., Damiens, D., Vreysen, M.J., Lemperière, G. and Gilles, J., 2013. Reproductive strategies of Aedes albopictus (Diptera: Culicidae) and implications for the sterile insect technique. *PLoS One*, *8*(11), pp.e78884.
168. Pennetier, C., Warren, B., Dabiré, K.R., Russell, I.J. and Gibson, G., 2010. "Singing on the wing" as a mechanism for species recognition in the malarial mosquito *Anopheles gambiae. Current Biology*, *20*(2), pp.131-136.

169. Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, *29*(9), pp.e45-e45.
170. Pondeville, E., David, J. P., Guittard, E., Maria, A., Jacques, J. C., Ranson, H., and Dauphin-Villemant, C., 2013. Microarray and RNAi analysis of P450s in *Anopheles gambiae* male and female steroidogenic tissues: CYP307A1 is required for ecdysteroid synthesis. *PloS One, 8*(12), pp.79861.
171. Puglise, J. M., Estep, A. S., and Becnel, J. J., 2016. Expression profiles and RNAi silencing of Inhibitor of Apoptosis transcripts in *Aedes, Anopheles,* and *Culex* mosquitoes (Diptera: Culicidae). *Journal of Medical Entomology*, *53*(2), pp.304-314.
172. Provost, M.W., Lum, P.T. and Branch, N., 1961. Rotation of male terminalia in Aedes taeniorhynchus (Diptera: Culicidae) as affected by temperature. *Annals of the Entomological Society of America*, *54*(6), pp.896-900.
173. Ranson, H., 2017. Current and future prospects for preventing malaria transmission via the use of insecticides. *Cold Spring Harbor Perspectives in Medicine*, *7*(11), pp.a026823.
174. Ranson, H., N'guessan, R., Lines, J., Moiroux, N., Nkuni, Z. and Corbel, V., 2011. Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control?. *Trends in Parasitology*, *27*(2), pp.91-98.
175. Rasheed, H., Ye, C., Meng, Y., Ran, Y., Li, J., & Su, X., 2019. Comparative transcriptomic analysis and endocuticular protein gene expression of alate adults, workers and soldiers of the termite Reticulitermes aculabialis. *BMC Genomics*, *20*(1), pp.1-12.
176. Rebers, J.E. and Willis, J.H., 2001. A conserved domain in arthropod cuticular proteins binds chitin. *Insect Biochemistry and Molecular Biology*, *31*(11), pp.1083-1093.
177. Rechsteiner, M., 1991. Natural substrates of the ubiquitin proteolytic pathway. *Cell*, *66*(4), pp.615-618.
178. Rees, D.M. and Onishi, K., 1951. Morphology of the terminalia and internal reproductive organs, and copulation in the mosquito, Culiseta inornata (Williston)(Diptera, Culicidae). *Proceedings of the Entomological Society of Washington*, *53*, pp.233-246.

179. Regna, K., Harrison, R.M., Heyse, S.A., Chiles, T.C., Michel, K. and Muskavitch, M.A., 2016. RNAi trigger delivery into *Anopheles gambiae* pupae. *Journal of Visualized Experiments* (109), pp.53738.
180. Richards, A.G., 1958. The cuticle of arthropods. *Ergebnisse der Biologie*, pp.1-26.
181. Ross, G.A., 1936. Insecticide as a Major Measure in Control of Malaria, being an Account of the Methods and Organisation put in Force in Natal and Zululand during the Past Six Years. *Insecticide as a Major Measure in Control of Malaria, being an Account of the Methods and Organisation put in Force in Natal and Zululand during the Past Six Years.* pp.114-133.
182. Roth, L.M., 1948. A study of mosquito behavior. An experimental laboratory study of the sexual behavior of Aedes aegypti (Linnaeus). *The American Midland Naturalist*, *40*(2), pp.265-352.
183. Russell, P.F. and Rao, T.R., 1942. On the ecology of larvae of *Anopheles culicifacies* Giles, in borrow-pits. *Bulletin of Entomological Research*, *32*(4), pp.341-361.
184. Sachs, J. and Malaney, P., 2002. The economic and social burden of malaria. *Nature*, *415*(6872), pp.680-685.
185. Sawadogo, P.S., Namountougou, M., Toé, K.H., Rouamba, J., Maïga, H., Ouédraogo, K.R., Baldet, T., Gouagna, L.C., Kengne, P., Simard, F. and Costantini, C., 2014. Swarming behaviour in natural populations of *Anopheles gambiae* and *An. coluzzii*: review of 4 years survey in rural areas of sympatry, Burkina Faso (West Africa). *Acta Tropica*, *132*, pp.S42-S52.
186. Sawadogo, S.P., Diabaté, A., Toé, H.K., Sanon, A., Lefevre, T., Baldet, T., Gilles, J., Simard, F., Gibson, G., Sinkins, S. and Dabiré, R.K., 2013. Effects of age and size on *Anopheles gambiae ss* male mosquito mating success. *Journal of Medical Entomology*, *50*(2), pp.285-293.
187. Sawadogo, S.P., Niang, A., Bilgo, E., Millogo, A., Maïga, H., Dabire, R.K., Tripet, F. and Diabaté, A., 2017. Targeting male mosquito swarms to control malaria vector density. *PloS One*, *12*(3), pp.e0173273.
188. Schefe, J.H., Lehmann, K.E., Buschmann, I.R., Unger, T. and Funke-Kaiser, H., 2006. Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's C T difference" formula. *Journal of Molecular Medicine*, *84*(11), pp.901-910.

189. Scott, J.A., Brogdon, W.G. and Collins, F.H., 1993. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. *The American Journal of Tropical Medicine and Hygiene*, *49*(4), pp.520-529.
190. Shakeel, M., Rodriguez, A., Tahir, U.B. and Jin, F., 2018. Gene expression studies of reference genes for quantitative real-time PCR: an overview in insects. *Biotechnology Letters*, *40*(2), pp.227-236.
191. Sharp, B.L. and Le Sueur, D., 1996. Malaria in South Africa the past, the present and selected implications for the future. *South African Medical Journal*, *86*(1).
192. Shi, Y. and He, M., 2014. Differential gene expression identified by RNA-Seq and qPCR in two sizes of pearl oyster (Pinctada fucata). *Gene*, *538*(2), pp.313-322.
193. Sinka, M.E., Rubio-Palis, Y., Manguin, S., Patil, A.P., Temperley, W.H., Gething, P.W., Van Boeckel, T., Kabaria, C.W., Harbach, R.E. and Hay, S.I., 2011. Correction: The dominant *Anopheles* vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasites and Vectors*, *4*, pp.210.
194. Sinkins, S.P. and Gould, F., 2006. Gene drive systems for insect disease vectors. *Nature Reviews Genetics*, *7*(6), pp.427-435.
195. Soderlund, D.M., Clark, J.M., Sheets, L.P., Mullin, L.S., Piccirillo, V.J., Sargent, D., Stevens, J.T. and Weiner, M.L., 2002. Mechanisms of pyrethroid neurotoxicity: implications for cumulative risk assessment. *Toxicology*, *171*(1), pp.3-59.
196. Sokhna, C., Ndiath, M.O. and Rogier, C., 2013. The changes in mosquito vector behaviour and the emerging resistance to insecticides will challenge the decline of malaria. *Clinical Microbiology and Infection*, *19*(10), pp.902-907.
197. Sougoufara, S., Doucouré, S., Sembéne, P.M.B., Harry, M. and Sokhna, C., 2017. Challenges for malaria vector control in sub-Saharan Africa: resistance and behavioral adaptations in *Anopheles* populations. *Journal of Vector Borne Diseases*, *54*(1), pp.4.
198. Spillings, B.L., Brooke, B.D., Koekemoer, L.L., Chiphwanya, J., Coetzee, M. and Hunt, R.H., 2009. A new species concealed by *Anopheles funestus* Giles, a major malaria vector in Africa. *The American Journal of Tropical Medicine and Hygiene*, *81*(3), pp.510-515.

199. Sternberg, S.H., LaFrance, B., Kaplan, M. and Doudna, J.A., 2015. Conformational control of DNA target cleavage by CRISPR–Cas9. *Nature*, *527*(7576), pp.110-113.
200. Taracena, M.L., Hunt, C.M., Benedict, M.Q., Pennington, P.M. and Dotson, E.M., 2019. Downregulation of female doublesex expression by oral-mediated RNA interference reduces number and fitness of *Anopheles gambiae* adult females. *Parasites and Vectors*, *12*(1), pp.1-11.
201. Theobald, F.V., 1903. A Monograph of the Culicidae or Mosquitoes. III. *London: Brit*. pp;. 359.
202. Tiyaboonchai, W., 2013. Chitosan nanoparticles: a promising system for drug delivery. *Naresuan University Journal: Science and Technology*, *11*(3), pp.51-66.
203. Togawa, T., Dunn, W.A., Emmons, A.C., Nagao, J. and Willis, J.H., 2008. Developmental expression patterns of cuticular protein genes with the R&R Consensus from *Anopheles gambiae. Insect Biochemistry and Molecular Biology*, *38*(5), pp.508-519.
204. Van Bakel, H. and Hughes, T.R., 2009. Establishing legitimacy and function in the new transcriptome. *Briefings in Functional Genomics and Proteomics*, *8*(6), pp.424-436.
205. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*(7), pp.1-12.
206. Vannini, L. and Willis, J.H., 2017. Localization of RR-1 and RR-2 cuticular proteins within the cuticle of *Anopheles gambiae*. *Arthropod Structure and Development*, *46*(1), pp.13-29.
207. Varshavsky, A., 2011. The N-end rule pathway and regulation by proteolysis. *Protein Science*, *20*(8), pp.1298-1345.
208. Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W., 1995. Serial analysis of gene expression. *Science*, *270*(5235), pp.484-487.
209. Vezenegho, S.B., Chiphwanya, J., Hunt, R.H., Coetzee, M., Bass, C. and Koekemoer, L.L., 2013. Characterization of the *Anopheles funestus* group, including *Anopheles funestus*-like, from Northern Malawi. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *107*(12), pp.753-762.

210. Vierstra, R.D., 1996. Proteolysis in plants: mechanisms and functions. *Post-Transcriptional Control of Gene Expression in Plants*, pp.275-302.
211. Vontas, J. and Mavridis, K., 2019. Vector population monitoring tools for insecticide resistance management: Myth or fact?. *Pesticide Biochemistry and Physiology*, *161*, pp.54-60.
212. Waite, J.L., Swain, S., Lynch, P.A., Sharma, S.K., Haque, M.A., Montgomery, J. and Thomas, M.B., 2017. Increasing the potential for malaria elimination by targeting zoophilic vectors. *Scientific Reports*, *7*(1), pp.1-10.
213. Wang, Z., Gerstein, M. and Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), pp.57-63.
214. Waterhouse, P.M., Graham, M.W. and Wang, M.B., 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceedings of the National Academy of Sciences*, *95*(23), pp.13959-13964.
215. Weedall, G.D., Mugenzi, L.M., Menze, B.D., Tchouakui, M., Ibrahim, S.S., Amvongo-Adjia, N., Irving, H., Wondji, M.J., Tchoupo, M., Djouaka, R. and Riveron, J.M., 2019. A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy. *Science Translational Medicine*, *11*(484).
216. White, G.B., 1985. *Anopheles bwambae* sp. n., a malaria vector in the Semliki Valley, Uganda, and its relationships with other sibling species of the An. gambiae complex (Diptera: Culicidae). *Systematic Entomology*, *10*(4), pp.501-522.
217. Wianny, F. and Zernicka-Goetz, M., 2000. Specific interference with gene function by double-stranded RNA in early mouse development. *Nature Cell Biology*, *2*(2), pp.70-75.
218. Wiebe, A., Longbottom, J., Gleave, K., Shearer, F.M., Sinka, M.E., Massey, N.C., Cameron, E., Bhatt, S., Gething, P.W., Hemingway, J. and Smith, D.L., 2017. Geographical distributions of African malaria vector sibling species and evidence for insecticide resistance. *Malaria Journal*, *16*(1), pp.1-10.
219. Willis, J.H., 2010. Structural cuticular proteins from arthropods: annotation, nomenclature, and sequence characteristics in the genomics era. *Insect Biochemistry and Molecular Biology*, *40*(3), pp.189-204.

220. Willis, J.H., Iconomidou, V.A., Smith, R.F. and Hamodrakas, S.J., 2005. Comprehensive Insect Science. *Cuticular Proteins*, *4*, pp.79-109.
221. Winnebeck, E.C., Millar, C.D. and Warman, G.R., 2010. Why does insect RNA look degraded?. *Journal of Insect Science*, *10*(1), pp.159.
222. World Health Organization. 2018. World Malaria Report 2018. https://www.who.int/. Last accessed 7 January 2021.
223. World Health Organization. 2020. World Malaria Report 2020: 20 years of global progress and challenges. https://www.who.int/. Last accessed 7 January 2021.
224. Wu, A.R., Neff, N.F., Kalisky, T., Dalerba, P., Treutlein, B., Rothenberg, M.E., Mburu, F.M., Mantalas, G.L., Sim, S., Clarke, M.F. and Quake, S.R., 2014. Quantitative assessment of single-cell RNA-sequencing methods. *Nature Methods*, *11*(1), pp.41.
225. Xie, F.Y., Woodle, M.C. and Lu, P.Y., 2006. Harnessing in vivo siRNA delivery for drug discovery and therapeutic development. *Drug Discovery Today*, *11*(1-2), pp.67-73.
226. Yahouédo, G.A., Chandre, F., Rossignol, M., Ginibre, C., Balabanidou, V., Mendez, N.G.A., Pigeon, O., Vontas, J. and Cornelie, S., 2017. Contributions of cuticle permeability and enzyme detoxification to pyrethroid resistance in the major malaria vector *Anopheles gambiae*. *Scientific Reports*, *7*(1), pp.1-10.
227. Yang, C.H., Yang, P.C., Zhang, S.F., Shi, Z.Y., Kang, L. and Zhang, A.B., 2017. Identification, expression pattern, and feature analysis of cuticular protein genes in the pine moth Dendrolimus punctatus (Lepidoptera: Lasiocampidae). *Insect Biochemistry and Molecular Biology*, *83*, pp.94-106.
228. Yoder, J.A., Walsh, C.P. and Bestor, T.H., 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, *13*(8), pp.335-340.
229. Yuval, B. and Bouskila, A., 1993. Temporal dynamics of mating and predation in mosquito swarms. *Oecologia*, *95*(1), pp.65-69.
230. Yuval, B., Holliday-Hanson, M.L., Washino, R.K., 1994. Energy budget of swarming male mosquitoes. *Ecological Entomology*, *9(1),* pp.74-78.
231. Zawada, J.W., Dahan-Moss, Y.L., Muleba, M., Dabire, R.K., Maïga, H., Venter, N., Davies, C., Hunt, R.H., Coetzee, M. and Koekemoer, L.L., 2018. Molecular and physiological analysis of *Anopheles funestus* swarms in Nchelenge, Zambia. *Malaria Journal*, *17*(1), pp.1-7.

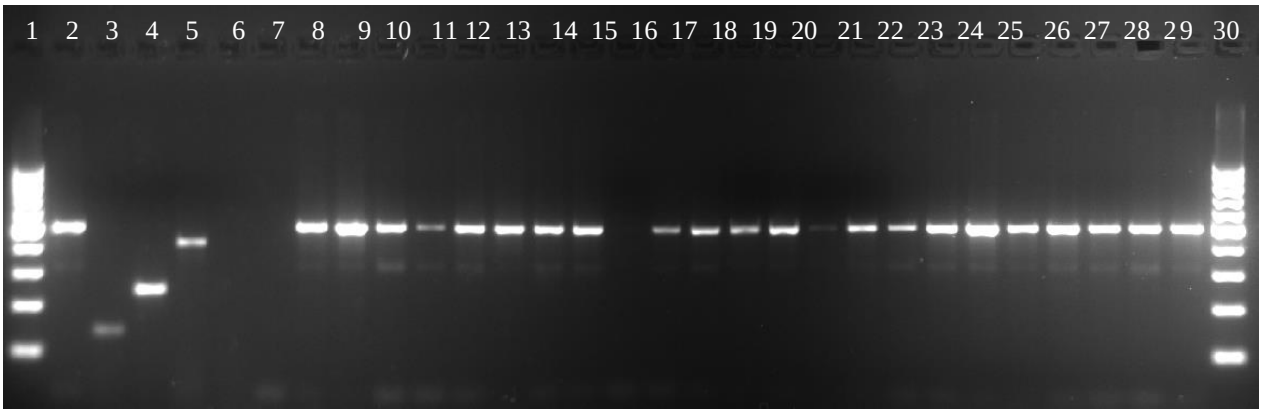
232. Zhang, X., Zhang, J. and Zhu, K.Y., 2010. Chitosan/double-stranded RNA nanoparticle-mediated RNA interference to silence chitin synthase genes through larval feeding in the African malaria mosquito (*Anopheles gambiae*). *Insect Molecular Biology*, *19*(5), pp.683-693.

Websites:

1. IRMapper
http://www.irmapper.com. Last accessed 6 January 2021.
2. OligoAnalyzer Tool
https://eu.idtdna.com/calc/analyzer. Last accessed 17 January 2020.
3. VectorBase
https://vectorbase.org/vectorbase/app. Last accessed 7 January 2021.
4. National Center for Biotechnology and Information
https://www.ncbi.nlm.nih.gov/sra/SRX7649914. Last accessed 7 January 2021.
5. beiRESOURCES.
https://www.beiresources.org/portals/2/MR4/MR4_Publications/Methods%20in%20Anopheles%20Research%202014/2014MethodsinAnophelesResearchManualFullVersionv2tso.pdf. Last accessed 7 January 2021.
6. https://earth.google.com/web/search/S09%C2%B019%E2%80%9928.6%E2%80%9D+E028%C2%B047%E2%80%9906.9%E2%80%9D. Last accessed 27 January 2021.
7. Uniprot (https://www.uniprot.org/uniprot/Q9FWW6). Last accessed 22 February 2021.

8 Appendix

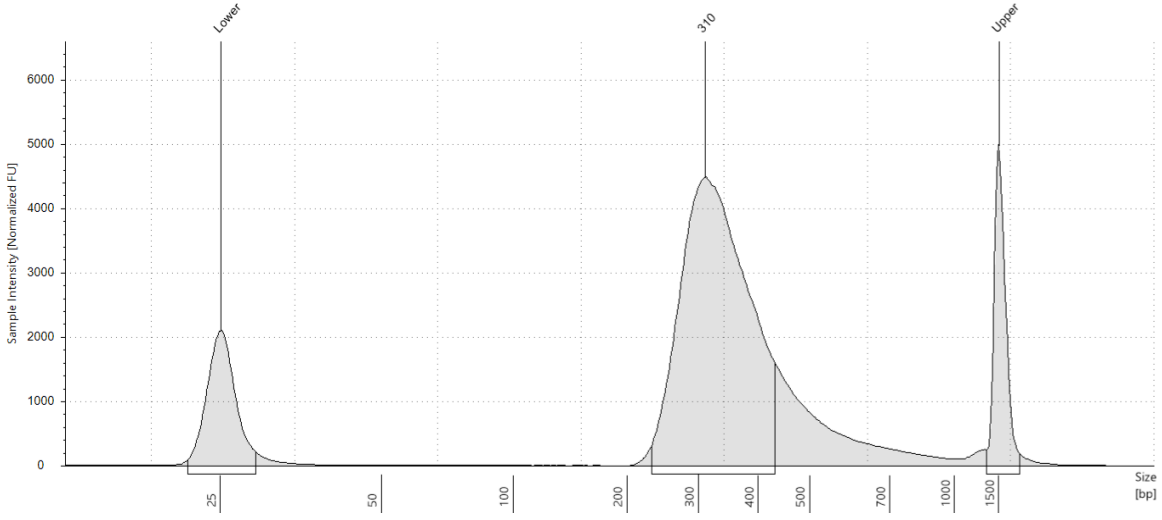
8.1 Example *An. funestus* PCR results of swarm 4



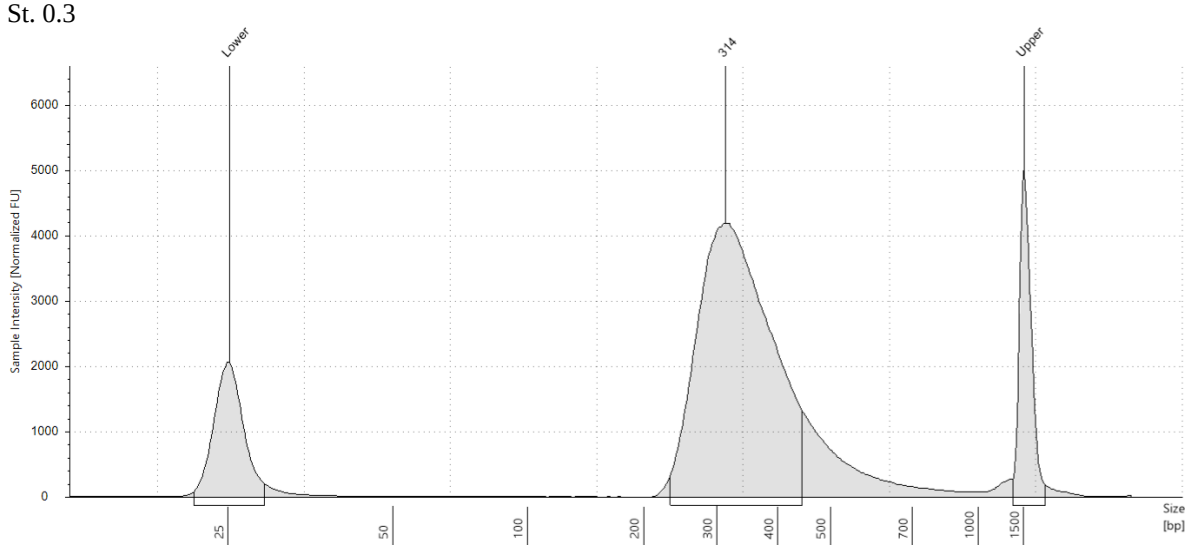
Lanes 1 and 30 represent molecular weight markers, lane 2-6 represent controls (*An. funestus, An. leesoni, An. parensis, An. rivulorum,* no template control, PCR negative control, respectively) , Lanes 8 to 30 represent samples from swarm 4 collected from Nchelenge, Zambia.

8.2 Bioanalyzer results: Bioanalyzer graphs of samples

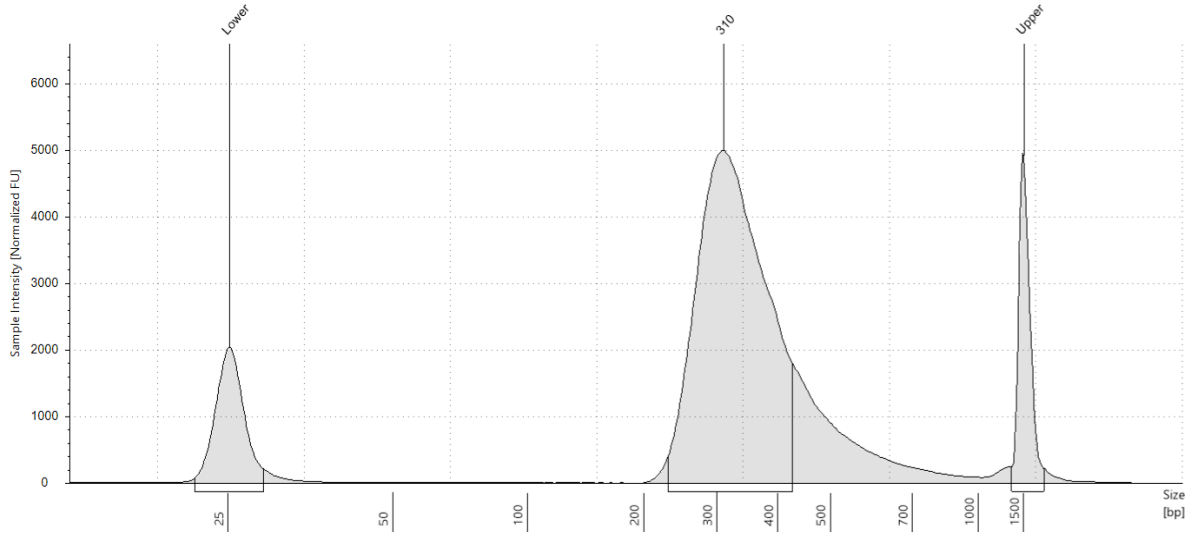
St 0.2



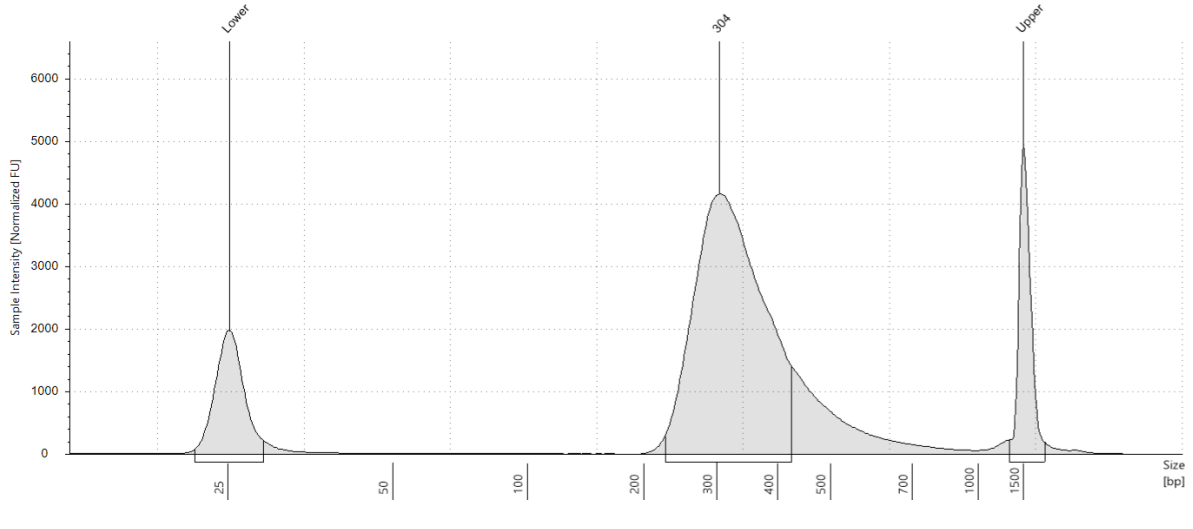
St. 0.3



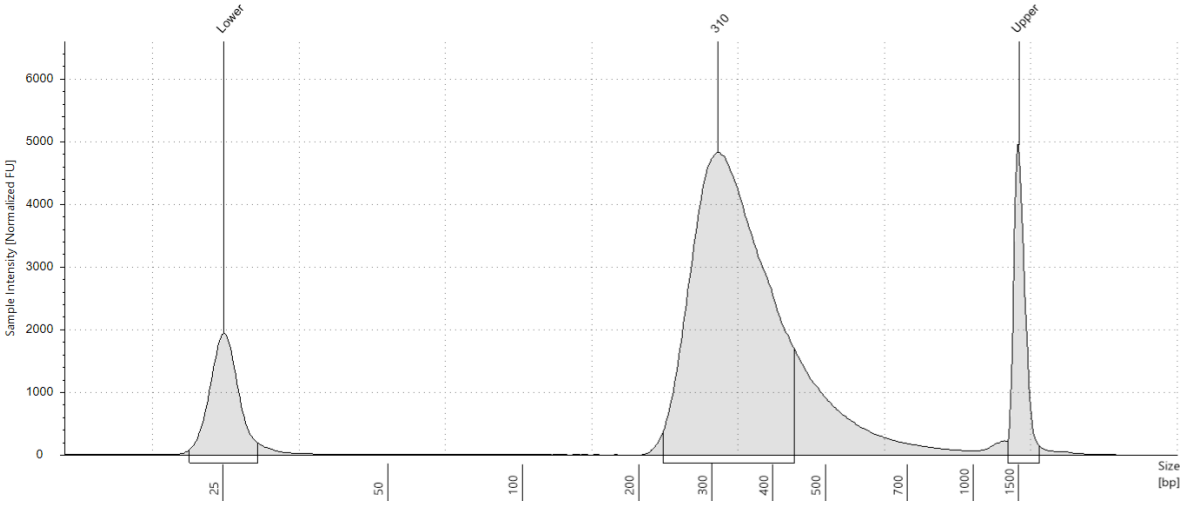
St. 3.1



St. 3.2



St. 3.3



8.3 cDNA concentrations: Working out of cDNA concentrations for qPCR

Sample	Vol. of cDNA (µL)	Vol. of Nuclease-free water (µL)	Final Concentration of 10 µL (ng/µL)
0.1	3.792	6.208	100
0.2	1.788	8.212	100
0.3	1.358	8.642	100
3.1	1.930	8.07	100
3.2	2.242	7.758	100
3.3	2.237	7.763	100

8.4 **RNA-seq data**

Gene_ID	Name	Description	Avg. FC 3/0	Gene_ID	Name	Description	Avg. FC 3/0
AFUN009542	.	.	-114,72	AFUN018927	.	serine protease inhibitor (serpin) [Source:Projected from Anopheles gambiae (AGAP004198) VB Community Annotation]	-2,59
AFUN019559	.	Polyubiquitin-B [Source:Projected from Anopheles gambiae (AGAP001970) VB Community Annotation]	-39,52	AFUN021977	.	.	-2,59
AFUN004371	.	.	-34,64	AFUN019535	.	.	-2,58
AFUN002627	.	Flavin-containing monooxygenase FMO GS-OX-like [Source:Projected from Anopheles gambiae (AGAP010399) VB Community Annotation]	-33,61	AFUN016476	.	.	-2,58
AFUN018979	.	.	-30,70	AFUN019349	.	.	-2,57
AFUN018851	.	troponin C [Source:Projected from Anopheles gambiae (AGAP006178) VB Community Annotation]	-26,50	AFUN020982	.	.	-2,57
AFUN007942	.	.	-26,23	AFUN005308	.	.	-2,57
AFUN019730	.	.	-26,04	AFUN008316	.	.	-2,57
AFUN021084	.	.	-23,20	AFUN002908	.	.	-2,56
AFUN007286	.	.	-22,38	AFUN009723	.	Elongation of very long chain fatty acids protein [Source:UniProtKB/TrEMBL%3BAcc:A0A182RTX0]	-2,56
AFUN003703	.	.	-22,20	AFUN021644	.	.	-2,56
AFUN006524	.	cuticular protein 3 from fifty-one aa family [Source:Projected from Anopheles gambiae (AGAP004690) VB Community Annotation]	-21,72	AFUN021135	.	.	-2,56
AFUN021553	.	.	-20,99	AFUN008813	.	.	-2,55
AFUN007919	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP002994) VB Community Annotation]	-20,78	AFUN006627	.	alpha-aminoadipic semialdehyde synthase [Source:Projected from Anopheles gambiae (AGAP008632) VB Community Annotation]	-2,53
AFUN005191	.	Dimeric dihydrodiol dehydrogenase [Source:Projected from Anopheles gambiae (AGAP009225) VB Community Annotation]	-20,48	AFUN022365	.	.	-2,53
AFUN021422	.	.	-19,58	AFUN002414	Obp6	odorant-binding protein 6 [Source:VB Community Annotation]	-2,53
AFUN001082	.	.	-19,03	AFUN005746	.	.	-2,53
AFUN019978	.	.	-18,94	AFUN005155	.	Aminomethyltransferase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RG19]	-2,53
AFUN009991	.	.	-18,56	AFUN003300	.	.	-2,53
AFUN004294	.	.	-18,00	AFUN002269	.	.	-2,53

AFUN008772	.	tRNA pseudouridine synthase [Source:Projected from Anopheles gambiae (AGAP009903) VB Community Annotation]	**-17,88**	AFUN000298	.	ecdysoneless homolog [Source:Projected from Anopheles gambiae (AGAP005814) VB Community Annotation]	**-2,53**
AFUN008663	.	cuticular protein 3 in CPLCA family [Source:Projected from Anopheles gambiae (AGAP006148) VB Community Annotation]	**-16,80**	AFUN014684	.	.	**-2,52**
AFUN019566	.	.	**-15,68**	AFUN001242	.	Alpha-actinin%2C sarcomeric [Source:Projected from Anopheles gambiae (AGAP001497) UniProtKB/Swiss-Prot%3BAcc:Q7PKQ5]	**-2,51**
AFUN016208	.	Sulfotransferase (Sult) [Source:VB Community Annotation]	**-15,28**	AFUN003824	.	.	**-2,51**
AFUN004703	.	.	**-14,91**	AFUN004740	.	.	**-2,50**
AFUN008616	.	cuticular protein [Source:Projected from Anopheles gambiae (AGAP003308) VB Community Annotation]	**-14,82**	AFUN000144	.	.	**-2,50**
AFUN016626	.	.	**-13,05**	AFUN020277	.	Ras-related protein Rab-23 [Source:Projected from Anopheles gambiae (AGAP011565) VB Community Annotation]	**-2,50**
AFUN021590	.	.	**-12,89**	AFUN007436	.	peptidoglycan recognition protein (long) [Source:Projected from Anopheles gambiae (AGAP005552) VB Community Annotation]	**-2,50**
AFUN008335	.	.	**-12,31**	AFUN011893	.	.	**-2,49**
AFUN021226	.	chitinase [Source:Projected from Anopheles gambiae (AGAP002457) VB Community Annotation]	**-11,97**	AFUN001482	.	.	**-2,49**
AFUN002477	.	Synaptic vesicle protein [Source:Projected from Anopheles gambiae (AGAP004350) VB Community Annotation]	**-11,82**	AFUN021978	.	.	**-2,48**
AFUN001291	.	D-amino-acid oxidase [Source:Projected from Anopheles gambiae (AGAP001116) VB Community Annotation]	**-11,64**	AFUN001672	.	26S proteasome 26S non-ATPase subunit N7 [Source:Projected from Anopheles gambiae (AGAP002061) VB Community Annotation]	**-2,47**
AFUN021425	.	.	**-11,55**	AFUN020478	.	all-trans-retinol dehydrogenase (NAD+) [Source:Projected from Anopheles gambiae (AGAP008125) VB Community Annotation]	**-2,47**
AFUN020892	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP005999) VB Community Annotation]	**-11,39**	AFUN008677	.	.	**-2,47**
AFUN021754	.	elongation of very long chain fatty acids protein [Source:Projected from Anopheles gambiae (AGAP010695) VB Community Annotation]	**-11,05**	AFUN005649	.	.	**-2,47**
AFUN008446	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae	**-10,99**	AFUN018674	.	.	**-2,46**

		(AGAP005451) VB Community Annotation]					
AFUN021080	.	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012322) VB Community Annotation]	**-10,72**	AFUN020454	.	.	**-2,46**
AFUN014716	.	.	**-10,57**	AFUN007058	.	26S proteasome 26S non-ATPase subunit N12 [Source:Projected from Anopheles gambiae (AGAP002606) VB Community Annotation]	**-2,46**
AFUN003006	.	.	**-10,19**	AFUN002871	.	.	**-2,46**
AFUN019696	.	.	**-10,06**	AFUN008437	.	.	**-2,46**
AFUN012022	.	cuticular protein CPLCG family [Source:Projected from Anopheles gambiae (AGAP008459) VB Community Annotation]	**-9,98**	AFUN010681	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006009) VB Community Annotation]	**-2,46**
AFUN001952	.	uncoupling protein 2%2C mitochondrial [Source:Projected from Anopheles gambiae (AGAP011676) VB Community Annotation]	**-9,75**	AFUN011295	SRPN9	serine protease inhibitor (serpin) 9 [Source:VB Community Annotation]	**-2,46**
AFUN021976	.	.	**-9,38**	AFUN019073	.	.	**-2,45**
AFUN018761	.	.	**-9,32**	AFUN015947	.	sugar transporter ERD6-like [Source:Projected from Anopheles gambiae (AGAP003494) VB Community Annotation]	**-2,45**
AFUN004506	.	Synaptic vesicle protein [Source:Projected from Anopheles gambiae (AGAP013063) VB Community Annotation]	**-9,22**	AFUN022205	.	takeout [Source:Projected from Anopheles gambiae (AGAP004263) VB Community Annotation]	**-2,45**
AFUN022364	.	.	**-9,14**	AFUN005685	.	.	**-2,45**
AFUN006563	.	.	**-8,79**	AFUN016420	.	ubiquitin-protein ligase E3 A [Source:VB Community Annotation]	**-2,44**
AFUN010477	.	cuticular protein 1 from CPFL family [Source:Projected from Anopheles gambiae (AGAP010902) VB Community Annotation]	**-8,65**	AFUN001576	.	sinuous [Source:Projected from Anopheles gambiae (AGAP003809) VB Community Annotation]	**-2,44**
AFUN018736	.	.	**-8,65**	AFUN008340	.	Bifunctional purine biosynthesis protein [Source:Projected from Anopheles gambiae (AGAP001423) VB Community Annotation]	**-2,44**
AFUN007615	.	ecdysis-triggering hormone [Source:Projected from Anopheles gambiae (AGAP007062) VB Community Annotation]	**-8,46**	AFUN009997	.	cuticular protein CPLCP8 [Source:Projected from Anopheles gambiae (AGAP008893) VB Community Annotation]	**-2,44**
AFUN020938	.	.	**-8,32**	AFUN016506	.	sodium/potassium/calcium exchanger [Source:VB Community Annotation]	**-2,44**
AFUN021964	.	cuticular protein CPLCG family [Source:Projected from Anopheles gambiae (AGAP008445) VB Community Annotation]	**-8,27**	AFUN007479	.	.	**-2,43**
AFUN011306	.	.	**-8,13**	AFUN018987	.	.	**-2,42**
AFUN002973	.	cysteine sulfinic acid decarboxylase [Source:Projected from	**-8,09**	AFUN008339	.	heat shock protein 90kDa beta [Source:Projected from Anopheles gambiae	**-2,42**

		Anopheles gambiae (AGAP002425) VB Community Annotation]				(AGAP001424) VB Community Annotation]	
AFUN000713	.	protein flightin [Source:Projected from Anopheles gambiae (AGAP007249) VB Community Annotation]	**-7,76**	AFUN018725	.	.	**-2,42**
AFUN021062	.	kelch-like protein 1/4/5 [Source:Projected from Anopheles gambiae (AGAP001513) VB Community Annotation]	**-7,75**	AFUN007150	.	.	**-2,42**
AFUN008661	.	.	**-7,70**	AFUN019526	.	.	**-2,41**
AFUN002384	.	.	**-7,58**	AFUN019029	.	.	**-2,41**
AFUN016422	.	.	**-7,57**	AFUN016397	.	Beta-galactosidase [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HED6]	**-2,41**
AFUN011588	.	.	**-7,56**	AFUN006682	.	C-type lectin (CTL) - selectin like [Source:Projected from Anopheles gambiae (AGAP000929) VB Community Annotation]	**-2,40**
AFUN002198	.	myosin heavy chain [Source:Projected from Anopheles gambiae (AGAP010147) VB Community Annotation]	**-7,53**	AFUN001984	.	.	**-2,40**
AFUN004704	.	.	**-7,51**	AFUN018479	.	.	**-2,40**
AFUN003096	.	Elongation of very long chain fatty acids protein [Source:UniProtKB/TrEMBL%3BAcc:A0A182RA82]	**-7,48**	AFUN008583	.	.	**-2,40**
AFUN004125	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP006283) VB Community Annotation]	**-7,42**	AFUN004012	.	Retinaldehyde binding protein [Source:Projected from Anopheles gambiae (AGAP008133) UniProtKB/TrEMBL%3BAcc:A0A1S4GY47]	**-2,39**
AFUN018885	.	.	**-7,29**	AFUN003170	.	.	**-2,39**
AFUN004214	.	Haloacid dehalogenase-like hydrolase domain-containing protein 1A [Source:Projected from Anopheles gambiae (AGAP003372) VB Community Annotation]	**-7,24**	AFUN000053	.	cardioacceleratory peptide [Source:Projected from Anopheles gambiae (AGAP009729) VB Community Annotation]	**-2,39**
AFUN019066	.	heme peroxidase [Source:Projected from Anopheles gambiae (AGAP013327) VB Community Annotation]	**-7,23**	AFUN006809	.	.	**-2,39**
AFUN015909	.	cytochrome P450 [Source:Projected from Anopheles gambiae (AGAP005660) VB Community Annotation]	**-7,20**	AFUN021421	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP003381) VB Community Annotation]	**-2,39**
AFUN016216	.	cuticular protein [Source:VB Community Annotation]	**-7,19**	AFUN006084	.	20S proteasome subunit beta [Source:Projected from Anopheles gambiae (AGAP001697) VB Community Annotation]	**-2,39**
AFUN019843	.	.	**-7,11**	AFUN011064	SRPN19	Serine protease inhibitor (serpin) 19 [Source:UniProtKB/TrEMBL%3BAcc:A0A182RXP6]	**-2,39**
AFUN010480	.	cuticular protein 2 from fifty-one aa family [Source:Projected from Anopheles gambiae (AGAP010901) VB Community Annotation]	**-7,08**	AFUN021975	.	.	**-2,39**

AFUN002039	.	phosphatidylinositol transfer protein SEC14 [Source:Projected from Anopheles gambiae (AGAP007353) VB Community Annotation]	**-7,06**	AFUN016503	.	.	**-2,38**
AFUN004722	.	.	**-7,02**	AFUN006098	.	.	**-2,38**
AFUN007540	.	.	**-7,01**	AFUN020768	.	.	**-2,38**
AFUN005747	.	cuticular protein 4 from fifty-one aa family [Source:Projected from Anopheles gambiae (AGAP000382) VB Community Annotation]	**-6,82**	AFUN010361	.	Serine hydroxymethyltransferase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RVQ1]	**-2,38**
AFUN005432	.	.	**-6,79**	AFUN019195	.	.	**-2,37**
AFUN021869	.	C-4 methylsterol oxidase [Source:Projected from Anopheles gambiae (AGAP002767) VB Community Annotation]	**-6,59**	AFUN003123	.	DnaJ homolog subfamily B member [Source:Projected from Anopheles gambiae (AGAP010239) VB Community Annotation]	**-2,37**
AFUN014262	.	.	**-6,52**	AFUN007752	.	leishmanolysin-like peptidase [Source:Projected from Anopheles gambiae (AGAP002955) VB Community Annotation]	**-2,37**
AFUN002821	.	.	**-6,41**	AFUN019905	.	.	**-2,37**
AFUN002506	.	Cytoplasmic actin [Source:Projected from Anopheles gambiae (AGAP011515) UniProtKB/TrEMBL%3BAcc:A0A1S4H6W1]	**-6,40**	AFUN003738	Prosalpha4	26S proteasome alpha 4 subunit [Source:VB Community Annotation]	**-2,37**
AFUN018850	.	.	**-6,35**	AFUN016057	.	cellular retinaldehyde binding protein [Source:VB Community Annotation]	**-2,35**
AFUN022263	.	matrix metalloprotease [Source:Projected from Anopheles gambiae (AGAP003929) VB Community Annotation]	**-6,33**	AFUN003991	.	.	**-2,35**
AFUN000279	.	.	**-6,31**	AFUN009920	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006830) VB Community Annotation]	**-2,34**
AFUN004178	.	phosphoenolpyruvate carboxykinase (GTP) [Source:Projected from Anopheles gambiae (AGAP003350) VB Community Annotation]	**-6,30**	AFUN019791	.	Canopy homolog [Source:Projected from Anopheles gambiae (AGAP008615) VB Community Annotation]	**-2,34**
AFUN015730	.	.	**-6,29**	AFUN007335	.	.	**-2,34**
AFUN021916	.	.	**-6,26**	AFUN007183	.	venom allergen [Source:Projected from Anopheles gambiae (AGAP010550) VB Community Annotation]	**-2,34**
AFUN019457	.	.	**-6,22**	AFUN009688	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP003246) VB Community Annotation]	**-2,34**
AFUN008195	.	.	**-6,22**	AFUN020166	.	Chitinase [Source:Projected from Anopheles gambiae (AGAP000198) UniProtKB/TrEMBL%3BAcc:A0A1S4G846]	**-2,33**
AFUN010999	.	.	**-6,16**	AFUN011662	.	glycerol-3-phosphate O-acyltransferase 1/2 [Source:Projected from	**-2,32**

						Anopheles gambiae (AGAP003687) VB Community Annotation]	
AFUN015955	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006001) VB Community Annotation]	**-6,12**	AFUN009596	.	Differentially expressed in FDCP 6-like protein [Source:Projected from Anopheles gambiae (AGAP012332) VB Community Annotation]	**-2,32**
AFUN015954	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006000) VB Community Annotation]	**-6,06**	AFUN018623	.	.	**-2,32**
AFUN002142	.	Amidophosphoribosyltransferase [Source:Projected from Anopheles gambiae (AGAP000179) UniProtKB/TrEMBL%3BAcc:Q7QEW3]	**-6,01**	AFUN016523	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP002784) VB Community Annotation]	**-2,32**
AFUN019930	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP000345) VB Community Annotation]	**-6,01**	AFUN002796	.	.	**-2,32**
AFUN002141	.	phosphoribosylaminoimidazole carboxylase [Source:Projected from Anopheles gambiae (AGAP000180) VB Community Annotation]	**-6,00**	AFUN002314	.	stearoyl-CoA desaturase (delta-9 desaturase) [Source:Projected from Anopheles gambiae (AGAP003418) VB Community Annotation]	**-2,31**
AFUN003859	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP000344) VB Community Annotation]	**-5,97**	AFUN008299	Prosalpha7	26S proteasome alpha 7 subunit [Source:VB Community Annotation]	**-2,31**
AFUN016444	HPX1	Heme peroxidase 1 [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HDW6]	**-5,93**	AFUN005813	.	Branched-chain-amino-acid aminotransferase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RHV1]	**-2,31**
AFUN021836	.	Stearoyl-CoA desaturase [Source:Projected from Anopheles gambiae (AGAP013071) VB Community Annotation]	**-5,92**	AFUN007563	.	.	**-2,31**
AFUN018571	.	.	**-5,92**	AFUN021510	.	zinc carboxypeptidase A [Source:Projected from Anopheles gambiae (AGAP011434) VB Community Annotation]	**-2,31**
AFUN021078	.	.	**-5,88**	AFUN020376	.	.	**-2,31**
AFUN006309	.	cuticular protein CPCFC family [Source:Projected from Anopheles gambiae (AGAP007980) VB Community Annotation]	**-5,84**	AFUN005184	.	.	**-2,30**
AFUN019108	.	.	**-5,84**	AFUN008327	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP001433) VB Community Annotation]	**-2,30**
AFUN004344	.	.	**-5,82**	AFUN020009	.	.	**-2,30**
AFUN001201	.	tripartite motif-containing protein [Source:Projected from Anopheles gambiae (AGAP001544) VB Community Annotation]	**-5,81**	AFUN000369	.	cecropin anti-microbial peptide [Source:Projected from Anopheles gambiae (AGAP006722) VB Community Annotation]	**-2,30**
AFUN007551	.	.	**-5,77**	AFUN021994	.	.	**-2,30**
AFUN004638	.	Metalloendopeptidase [Source:UniProtKB/TrEMBL%3BAcc:A0A182REL1]	**-5,75**	AFUN018915	.	.	**-2,30**

AFUN010012	.	.	-5,72	AFUN020098	.	dystroglycan [Source:Projected from Anopheles gambiae (AGAP005162) VB Community Annotation]	-2,29
AFUN020299	.	.	-5,71	AFUN007187	.	.	-2,29
AFUN000839	.	.	-5,67	AFUN018958	.	mitochondrial 2-oxoglutarate/malate transporter (solute carrier family 25) [Source:Projected from Anopheles gambiae (AGAP012097) VB Community Annotation]	-2,29
AFUN006910	.	.	-5,61	AFUN022247	.	Sphingomyelin phosphodiesterase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RXV4]	-2,29
AFUN018637	.	.	-5,56	AFUN005697	.	.	-2,29
AFUN019728	.	.	-5,56	AFUN002199	.	cytochrome b5-related [Source:Projected from Anopheles gambiae (AGAP010149) VB Community Annotation]	-2,28
AFUN018576	.	.	-5,56	AFUN003252	.	Long-chain acyl-CoA synthetase [Source:Projected from Anopheles gambiae (AGAP011603) UniProtKB/TrEMBL%3BAcc:A0A1S4H733]	-2,28
AFUN018641	.	.	-5,47	AFUN010802	.	.	-2,27
AFUN009995	.	.	-5,35	AFUN006981	.	.	-2,27
AFUN005393	.	.	-5,35	AFUN022206	.	takeout [Source:Projected from Anopheles gambiae (AGAP004262) VB Community Annotation]	-2,27
AFUN004286	.	suppressor of G2 allele of SKP1 [Source:Projected from Anopheles gambiae (AGAP006161) VB Community Annotation]	-5,33	AFUN005208	.	.	-2,27
AFUN004527	.	Envelysin [Source:Projected from Anopheles gambiae (AGAP004620) VB Community Annotation]	-5,29	AFUN007641	.	26S proteasome 26S non-ATPase subunit N5 [Source:Projected from Anopheles gambiae (AGAP005535) VB Community Annotation]	-2,27
AFUN019697	.	.	-5,24	AFUN004753	.	short/branched chain acyl-CoA dehydrogenase [Source:Projected from Anopheles gambiae (AGAP009783) VB Community Annotation]	-2,26
AFUN010464	.	.	-5,16	AFUN000792	.	.	-2,26
AFUN021105	.	fructose-1%2C6-biphosphatase [Source:Projected from Anopheles gambiae (AGAP009173) VB Community Annotation]	-5,15	AFUN000631	.	.	-2,26
AFUN000107	.	.	-5,15	AFUN000484	.	.	-2,26
AFUN009101	.	serine-type enodpeptidase [Source:Projected from Anopheles gambiae (AGAP002543) VB Community Annotation]	-5,12	AFUN009446	.	solute carrier family 30 (zinc transporter)%2C member 5/7 [Source:Projected from Anopheles gambiae (AGAP004461) VB Community Annotation]	-2,25
AFUN019219	.	.	-5,09	AFUN019574	.	.	-2,25
AFUN010811	.	.	-5,07	AFUN000062	.	.	-2,25

AFUN020168	.	.	-5,06	AFUN008114	.	Ubiquitin domain-containing protein [Source:Projected from Anopheles gambiae (AGAP001747) VB Community Annotation]	-2,25
AFUN018632	.	.	-4,99	AFUN021894	.	sodium-coupled monocarboxylate transporter [Source:Projected from Anopheles gambiae (AGAP013429) VB Community Annotation]	-2,25
AFUN022012	.	.	-4,99	AFUN021002	.	nucleoside diphosphate-linked moiety X motif [Source:Projected from Anopheles gambiae (AGAP002403) VB Community Annotation]	-2,25
AFUN019590	.	.	-4,95	AFUN000168	.	26S proteasome 26S non-ATPase subunit N8 [Source:Projected from Anopheles gambiae (AGAP012135) VB Community Annotation]	-2,24
AFUN007092	.	glucosamine--fructose-6-phosphate aminotransferase (isomerizing) [Source:Projected from Anopheles gambiae (AGAP011352) VB Community Annotation]	-4,95	AFUN009226	.	.	-2,24
AFUN000190	.	.	-4,94	AFUN005259	.	ATP-binding cassette [Source:UniProtKB/TrEMBL%3BAcc:A0A182RGB5]	-2,24
AFUN020593	.	.	-4,94	AFUN007897	.	.	-2,24
AFUN021424	.	.	-4,90	AFUN021552	.	Cellular retinaldehyde-binding protein [Source:Projected from Anopheles gambiae (AGAP009381) UniProtKB/TrEMBL%3BAcc:A0A1S4H0Z1]	-2,24
AFUN002468	.	cellular retinaldehyde binding protein [Source:Projected from Anopheles gambiae (AGAP004339) VB Community Annotation]	-4,90	AFUN001096	.	.	-2,23
AFUN002911	.	.	-4,89	AFUN022316	.	26S proteasome 26S non-ATPase subunit N9 [Source:Projected from Anopheles gambiae (AGAP001651) VB Community Annotation]	-2,23
AFUN008586	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP001648) VB Community Annotation]	-4,89	AFUN006945	.	.	-2,23
AFUN011265	.	.	-4,87	AFUN009404	.	.	-2,23
AFUN018690	.	.	-4,86	AFUN003200	.	.	-2,23
AFUN005101	CYP325F2	cytochrome P450 [Source:VB Community Annotation]	-4,85	AFUN021361	.	fatty acyl-CoA elongase [Source:Projected from Anopheles gambiae (AGAP003195) VB Community Annotation]	2,00
AFUN021423	.	.	-4,85	AFUN002723	.	Eukaryotic translation initiation factor 3 subunit J [Source:UniProtKB/TrEMBL%3BAcc:A0A182R963]	2,00
AFUN008291	.	.	-4,82	AFUN005415	.	Tetraspanin [Source:UniProtKB/TrEMBL%3BAcc:A0A182RGQ6]	2,00

AFUN008665	.	cuticular protein 1 in CPLCA family [Source:Projected from Anopheles gambiae (AGAP006145) VB Community Annotation]	-4,82	AFUN008197	.	Phosphoglycerate mutase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RPL4]	2,01
AFUN001498	.	.	-4,80	AFUN021213	.	Transglutaminase [Source:Projected from Anopheles gambiae (AGAP009099) UniProtKB/TrEMBL%3BAcc:D2K8K8]	2,01
AFUN001232	.	.	-4,79	AFUN016480	.	.	2,01
AFUN007013	LRIM17	leucine rich immune protein (Coil-less) [Source:VB Community Annotation]	-4,78	AFUN021794	.	.	2,01
AFUN001233	.	.	-4,76	AFUN018794	.	.	2,01
AFUN002628	.	Flavin-containing monooxygenase [Source:UniProtKB/TrEMBL%3BAcc:A0A182R8W8]	-4,72	AFUN019771	.	.	2,01
AFUN018569	.	.	-4,68	AFUN016321	.	histone deacetylase [Source:VB Community Annotation]	2,01
AFUN009919	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006829) VB Community Annotation]	-4,66	AFUN020974	.	Succinate-semialdehyde dehydrogenase I [Source:Projected from Anopheles gambiae (AGAP003165) VB Community Annotation]	2,01
AFUN016174	Obp8	Odorant-binding protein 8 [Source:UniProtKB/TrEMBL%3BAcc:A0A1I8JUP1]	-4,64	AFUN000189	.	.	2,02
AFUN009112	.	glucose dehydrogenase (acceptor) [Source:Projected from Anopheles gambiae (AGAP002552) VB Community Annotation]	-4,63	AFUN016615	.	.	2,02
AFUN021588	.	.	-4,63	AFUN010296	.	.	2,02
AFUN011055	LRIM15	leucine rich immune protein (TM) [Source:VB Community Annotation]	-4,62	AFUN004800	.	JNK1/MAPK8-associated membrane protein [Source:Projected from Anopheles gambiae (AGAP001461) VB Community Annotation]	2,02
AFUN011127	.	.	-4,60	AFUN004813	.	.	2,03
AFUN019624	.	.	-4,57	AFUN019226	.	.	2,03
AFUN009544	.	Transferrin [Source:Projected from Anopheles gambiae (AGAP011169) UniProtKB/TrEMBL%3BAcc:A0A1S4H5X3]	-4,56	AFUN005051	.	Mitochondrial pyruvate carrier [Source:UniProtKB/TrEMBL%3BAcc:A0A182RFR9]	2,04
AFUN021350	.	.	-4,50	AFUN018795	.	.	2,04
AFUN005924	.	.	-4,48	AFUN021284	.	F-box and leucine-rich repeat protein [Source:Projected from Anopheles gambiae (AGAP009335) VB Community Annotation]	2,04
AFUN008829	GSTS1	Glutathione S-transferase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RRD4]	-4,47	AFUN004070	.	.	2,05
AFUN008662	.	cuticular protein unclassified [Source:Projected from Anopheles gambiae (AGAP006149) VB Community Annotation]	-4,47	AFUN020178	.	.	2,05
AFUN004032	.	Lipase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RCW3]	-4,47	AFUN000554	.	.	2,05

AFUN004581	.	4-nitrophenylphosphatase [Source:Projected from Anopheles gambiae (AGAP005972) VB Community Annotation]	-4,41	AFUN021720	.	.	2,05
AFUN021296	.	.	-4,38	AFUN007478	.	chitinase [Source:Projected from Anopheles gambiae (AGAP006414) VB Community Annotation]	2,05
AFUN005982	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP006261) VB Community Annotation]	-4,37	AFUN010256	.	Selenophosphate synthetase [Source:Projected from Anopheles gambiae (AGAP011375) VB Community Annotation]	2,06
AFUN016426	.	chitinase [Source:Projected from Anopheles gambiae (AGAP005634) VB Community Annotation]	-4,37	AFUN020895	.	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:B5AII3]	2,06
AFUN018663	.	.	-4,36	AFUN004950	.	.	2,06
AFUN016185	.	ATP-binding cassette transporter (ABC transporter) family C member [Source:VB Community Annotation]	-4,34	AFUN002075	.	Coenzyme Q10 homolog B (S. cerevisiae) [Source:Projected from Anopheles gambiae (AGAP007391) VB Community Annotation]	2,06
AFUN021443	.	thioester-containing protein [Source:Projected from Anopheles gambiae (AGAP008368) VB Community Annotation]	-4,33	AFUN003238	.	.	2,06
AFUN016578	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP009214) VB Community Annotation]	-4,33	AFUN009953	.	phosphotriesterase-related protein [Source:Projected from Anopheles gambiae (AGAP011061) VB Community Annotation]	2,07
AFUN003027	.	.	-4,32	AFUN021399	.	carbonic anhydrase I [Source:Projected from Anopheles gambiae (AGAP013402) VB Community Annotation]	2,07
AFUN007552	.	.	-4,30	AFUN007146	.	NADH-cytochrome b5 reductase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RLM8]	2,07
AFUN020128	.	.	-4,26	AFUN000268	.	.	2,08
AFUN006308	.	Short chain dehydrogenase/reductase family 16c member [Source:Projected from Anopheles gambiae (AGAP007981) UniProtKB/TrEMBL%3BAcc:A0A1S4GWV7]	-4,25	AFUN002139	.	.	2,08
AFUN009167	.	Sugar transporter ERD6-like [Source:Projected from Anopheles gambiae (AGAP001236) VB Community Annotation]	-4,24	AFUN005934	.	.	2,08
AFUN018872	.	.	-4,23	AFUN003744	.	aminoacyl tRNA synthase complex-interacting multifunctional protein [Source:Projected from Anopheles gambiae (AGAP005418) VB Community Annotation]	2,08
AFUN016471	.	.	-4,18	AFUN014447	.	.	2,08
AFUN009201	.	organic cation transporter [Source:Projected from Anopheles gambiae (AGAP001202) VB Community Annotation]	-4,14	AFUN009228	.	.	2,09

AFUN015859	.	.	-4,14	AFUN021811	.	dehydrogenase/reductase SDR family member 11 precursor [Source:Projected from Anopheles gambiae (AGAP005499) VB Community Annotation]	2,09
AFUN019447	.	26S proteasome 26S non-ATPase subunit N6 [Source:Projected from Anopheles gambiae (AGAP001440) VB Community Annotation]	-4,13	AFUN021925	.	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase [Source:Projected from Anopheles gambiae (AGAP002662) UniProtKB/TrEMBL%3BAcc:Q7QCJ7]	2,09
AFUN020443	.	.	-4,11	AFUN001481	.	.	2,09
AFUN007346	.	.	-4,10	AFUN018812	.	Mitochondrial pyruvate carrier [Source:UniProtKB/TrEMBL%3BAcc:A0A182RFR9]	2,10
AFUN018501	.	.	-4,06	AFUN001336	.	.	2,10
AFUN009783	.	.	-4,05	AFUN001326	.	.	2,10
AFUN018531	.	.	-4,04	AFUN000349	.	.	2,10
AFUN003467	.	.	-4,01	AFUN009831	.	DNA-directed RNA polymerase%2C mitochondrial [Source:Projected from Anopheles gambiae (AGAP007778) VB Community Annotation]	2,11
AFUN006060	.	Class B Scavenger Receptor (CD36 domain). [Source:Projected from Anopheles gambiae (AGAP005725) VB Community Annotation]	-3,98	AFUN020247	.	.	2,11
AFUN003893	.	.	-3,97	AFUN004150	.	aldose reductase [Source:Projected from Anopheles gambiae (AGAP011068) VB Community Annotation]	2,11
AFUN020241	.	glucose dehydrogenase (acceptor) [Source:Projected from Anopheles gambiae (AGAP002551) VB Community Annotation]	-3,97	AFUN021433	.	.	2,11
AFUN004759	.	cuticular protein [Source:Projected from Anopheles gambiae (AGAP009790) VB Community Annotation]	-3,94	AFUN008767	.	.	2,11
AFUN019710	.	.	-3,92	AFUN021839	.	dumpy-like protein [Source:Projected from Anopheles gambiae (AGAP003027) VB Community Annotation]	2,12
AFUN010704	.	PI-PLC X domain-containing protein [Source:Projected from Anopheles gambiae (AGAP010008) VB Community Annotation]	-3,88	AFUN020457	.	dual specificity phosphatase [Source:Projected from Anopheles gambiae (AGAP006593) VB Community Annotation]	2,12
AFUN022366	.	a disintegrin and metalloproteinase with thrombospondin motifs [Source:Projected from Anopheles gambiae (AGAP004201) VB Community Annotation]	-3,87	AFUN015971	.	.	2,12
AFUN011927	.	heat shock 70kDa protein [Source:Projected from Anopheles gambiae (AGAP004192) VB Community Annotation]	-3,84	AFUN015168	.	.	2,13

AFUN018949	.	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP012325) VB Community Annotation]	-3,83	AFUN003355	.	.	2,13
AFUN010798	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP012487) VB Community Annotation]	-3,78	AFUN019561	.	.	2,13
AFUN019437	.	.	-3,75	AFUN008304	.	.	2,13
AFUN019837	.	.	-3,73	AFUN008423	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP013184) VB Community Annotation]	2,14
AFUN009421	.	.	-3,72	AFUN016383	.	Transporter [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HF51]	2,14
AFUN016059	.	.	-3,72	AFUN005423	.	.	2,14
AFUN019620	.	.	-3,71	AFUN020900	.	.	2,14
AFUN019180	.	.	-3,70	AFUN021922	.	ubiquinone biosynthesis protein COQ9%2C mitochondrial [Source:Projected from Anopheles gambiae (AGAP004842) VB Community Annotation]	2,15
AFUN008935	.	.	-3,70	AFUN021219	.	.	2,16
AFUN018918	.	.	-3,69	AFUN011793	.	.	2,16
AFUN010737	.	Cullin-associated NEDD8-dissociated protein [Source:Projected from Anopheles gambiae (AGAP009970) UniProtKB/TrEMBL%3BAcc:A0A1S4H2L0]	-3,69	AFUN020953	.	excitatory amino acid transporter 2 isoform [Source:Projected from Anopheles gambiae (AGAP006718) VB Community Annotation]	2,16
AFUN008113	.	.	-3,67	AFUN007418	.	trypsinogen [Source:Projected from Anopheles gambiae (AGAP008277) VB Community Annotation]	2,16
AFUN007717	CYP325H1	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:A0A182RN86]	-3,67	AFUN009967	.	.	2,17
AFUN005673	.	.	-3,62	AFUN016430	LRIM27	Leucine rich immune protein (Coil-less) [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HDV4]	2,17
AFUN002681	.	solute carrier family 26%2C other [Source:Projected from Anopheles gambiae (AGAP010344) VB Community Annotation]	-3,59	AFUN019087	.	.	2,18
AFUN006783	.	.	-3,58	AFUN015802	CYP6P1	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:B5AII1]	2,18
AFUN021948	.	.	-3,56	AFUN021883	.	carboxypeptidase B [Source:Projected from Anopheles gambiae (AGAP006208) VB Community Annotation]	2,18
AFUN002143	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP000177) VB Community Annotation]	-3,55	AFUN019690	.	.	2,19
AFUN016351	.	.	-3,55	AFUN021608	.	.	2,19
AFUN001612	.	myosin light chain [Source:Projected from	-3,54	AFUN008739	.	.	2,19

		Anopheles gambiae (AGAP001622) VB Community Annotation]					
AFUN016062	.	Angiotensin-converting enzyme [Source:UniProtKB/TrEMBL%3BAcc:A0A1I8JUB2]	-3,54	AFUN011345	.	.	2,20
AFUN009830	.	.	-3,53	AFUN005409	.	2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase [Source:Projected from Anopheles gambiae (AGAP009648) VB Community Annotation]	2,20
AFUN016607	.	Signal transducer and activator of transcription [Source:UniProtKB/TrEMBL%3BAcc:A0A3F2YXQ7]	-3,52	AFUN014088	.	.	2,20
AFUN008572	.	.	-3,51	AFUN019375	.	.	2,20
AFUN021642	.	.	-3,51	AFUN018679	.	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP007286) VB Community Annotation]	2,21
AFUN021291	.	adenylate kinase [Source:Projected from Anopheles gambiae (AGAP009317) VB Community Annotation]	-3,45	AFUN004330	mRpL28	39S ribosomal protein L28%2C mitochondrial [Source:UniProtKB/TrEMBL%3BAcc:A0A182RDQ8]	2,21
AFUN005342	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP007042) VB Community Annotation]	-3,44	AFUN005986	.	.	2,22
AFUN008664	.	.	-3,43	AFUN010723	.	.	2,22
AFUN016214	.	aldehyde dehydrogenase family 7 member A1 [Source:VB Community Annotation]	-3,43	AFUN001121	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP002811) UniProtKB/TrEMBL%3BAcc:A0A1S4GG73]	2,22
AFUN000530	.	.	-3,43	AFUN010391	.	.	2,22
AFUN016188	.	.	-3,40	AFUN010292	.	.	2,22
AFUN008212	.	Elongation of very long chain fatty acids protein [Source:UniProtKB/TrEMBL%3BAcc:A0A182RPM9]	-3,38	AFUN016156	.	.	2,23
AFUN007437	.	cellular retinaldehyde binding protein [Source:Projected from Anopheles gambiae (AGAP005553) VB Community Annotation]	-3,37	AFUN002277	.	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP006080) VB Community Annotation]	2,24
AFUN011567	.	.	-3,37	AFUN004098	.	.	2,24
AFUN020510	.	tyrosine 3-monooxygenase [Source:Projected from Anopheles gambiae (AGAP006023) VB Community Annotation]	-3,37	AFUN008349	.	Niemann-Pick C2 protein [Source:Projected from Anopheles gambiae (AGAP002851) VB Community Annotation]	2,24
AFUN010481	.	.	-3,37	AFUN021145	.	.	2,24
AFUN015905	CYP4J10	cytochrome P450 [Source:VB Community Annotation]	-3,36	AFUN016600	.	.	2,25
AFUN021134	.	cellular retinaldehyde binding protein [Source:Projected from Anopheles gambiae (AGAP007356) VB Community Annotation]	-3,36	AFUN009413	.	iron-sulfur cluster assembly protein [Source:Projected from Anopheles gambiae (AGAP002730) VB Community Annotation]	2,25
AFUN015979	.	ATP-binding cassette transporter (ABC transporter) family G member	-3,32	AFUN002193	.	.	2,25

		[Source:Projected from Anopheles gambiae (AGAP009467) UniProtKB/TrEMBL%3BAcc: A0A1S4H190]					
AFUN000101	.	.	-3,32	AFUN022121	.	Aldo-keto reductase family 1%2C member C3 [Source:Projected from Anopheles gambiae (AGAP012851) VB Community Annotation]	2,25
AFUN020720	.	chymotrypsin like protease [Source:Projected from Anopheles gambiae (AGAP009000) VB Community Annotation]	-3,31	AFUN000965	.	.	2,26
AFUN016606	.	.	-3,30	AFUN006976	.	.	2,26
AFUN008504	.	scavenger receptor class B member [Source:Projected from Anopheles gambiae (AGAP004845) VB Community Annotation]	-3,30	AFUN002546	.	.	2,27
AFUN021398	.	carbonic anhydrase III [Source:Projected from Anopheles gambiae (AGAP002360) VB Community Annotation]	-3,30	AFUN000501	.	.	2,27
AFUN020282	.	.	-3,28	AFUN019067	.	HIG1 hypoxia inducible domain family%2C member 1B [Source:Projected from Anopheles gambiae (AGAP010716) VB Community Annotation]	2,27
AFUN021362	.	sensory appendage protein [Source:Projected from Anopheles gambiae (AGAP008052) VB Community Annotation]	-3,28	AFUN002418	.	DnaJ homolog subfamily C member [Source:Projected from Anopheles gambiae (AGAP003533) VB Community Annotation]	2,28
AFUN019227	.	.	-3,27	AFUN021905	.	Niemann-Pick Type C-2 [Source:Projected from Anopheles gambiae (AGAP002848) VB Community Annotation]	2,28
AFUN018581	.	.	-3,26	AFUN005104	.	glycine N-methyltransferase [Source:Projected from Anopheles gambiae (AGAP002198) VB Community Annotation]	2,28
AFUN007455	.	.	-3,26	AFUN001345	.	.	2,29
AFUN022209	.	secretory phospholipase A2 [Source:Projected from Anopheles gambiae (AGAP003846) VB Community Annotation]	-3,24	AFUN021120	.	.	2,29
AFUN004960	.	Ubiquitin-related modifier 1 homolog [Source:UniProtKB/TrEMBL%3BAcc:A0A182RFH8]	-3,24	AFUN014166	.	.	2,29
AFUN018860	.	aspartyl-tRNA synthetase [Source:Projected from Anopheles gambiae (AGAP005576) VB Community Annotation]	-3,24	AFUN010220	.	.	2,29
AFUN016125	.	.	-3,23	AFUN019062	.	.	2,29
AFUN002067	.	Fem-1 homolog b [Source:Projected from Anopheles gambiae (AGAP007382) VB Community Annotation]	-3,21	AFUN014168	.	.	2,30
AFUN008429	.	.	-3,21	AFUN008298	.	organic solute transporter subunit alpha [Source:Projected	2,30

						from Anopheles gambiae (AGAP001974) VB Community Annotation]	
AFUN022369	.	.	-3,21	AFUN022298	.	.	2,30
AFUN001865	.	.	-3,20	AFUN005981	.	.	2,30
AFUN002334	.	.	-3,19	AFUN015822	.	cecropin anti-microbial peptide [Source:Projected from Anopheles gambiae (AGAP000693) VB Community Annotation]	2,31
AFUN009965	.	dsRNase [Source:Projected from Anopheles gambiae (AGAP011047) VB Community Annotation]	-3,19	AFUN011354	.	Class A Scavenger Receptor (SRCR domain) with Lysyl Oxidase domain. [Source:Projected from Anopheles gambiae (AGAP004118) VB Community Annotation]	2,32
AFUN009238	.	facilitated trehalose transporter Tret1-2 homolog [Source:Projected from Anopheles gambiae (AGAP008900) VB Community Annotation]	-3,19	AFUN020019	.	3-glucan binding protein [Source:Projected from Anopheles gambiae (AGAP004455) VB Community Annotation]	2,32
AFUN000645	.	.	-3,18	AFUN004808	.	Geranylgeranyl transferase type-2 subunit beta [Source:UniProtKB/TrEMBL%3BAcc:A0A182RF30]	2,32
AFUN010886	.	acyl-protein thioesterase 1%2C2 [Source:Projected from Anopheles gambiae (AGAP012883) VB Community Annotation]	-3,18	AFUN004043	.	uridine kinase [Source:Projected from Anopheles gambiae (AGAP003874) VB Community Annotation]	2,33
AFUN015896	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP003252) VB Community Annotation]	-3,18	AFUN021048	.	juvenile hormone-inducible protein [Source:Projected from Anopheles gambiae (AGAP001553) VB Community Annotation]	2,33
AFUN011053	.	.	-3,18	AFUN008709	.	.	2,33
AFUN003294	.	.	-3,17	AFUN014158	.	.	2,35
AFUN004601	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP005995) VB Community Annotation]	-3,16	AFUN006849	.	.	2,35
AFUN002696	.	.	-3,16	AFUN015790	CYP6AG1	cytochrome P450 [Source:VB Community Annotation]	2,35
AFUN007826	.	.	-3,16	AFUN021449	.	troponin C [Source:Projected from Anopheles gambiae (AGAP006181) VB Community Annotation]	2,35
AFUN000486	.	.	-3,16	AFUN015919	CYP6Z1	cytochrome P450 [Source:VB Community Annotation]	2,36
AFUN007657	.	Elongation of very long chain fatty acids protein [Source:UniProtKB/TrEMBL%3BAcc:A0A182RN28]	-3,14	AFUN008297	.	.	2,36
AFUN007986	.	cytosolic carboxypeptidase [Source:Projected from Anopheles gambiae (AGAP001814) VB Community Annotation]	-3,14	AFUN007893	.	Argininosuccinate synthase [Source:Projected from Anopheles gambiae (AGAP003015) UniProtKB/Swiss-Prot%3BAcc:Q7PR38]	2,36
AFUN009922	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006833) VB Community Annotation]	-3,13	AFUN016590	.	.	2,36
AFUN005105	.	.	-3,13	AFUN016381	.	nutrient amino acid transporter [Source:VB Community Annotation]	2,36

AFUN005790	.	kinesin family member [Source:Projected from Anopheles gambiae (AGAP007592) VB Community Annotation]	-3,10	AFUN019032	.	.	2,37
AFUN005257	.	Mitogen-activated protein kinase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RGB3]	-3,10	AFUN008478	.	.	2,37
AFUN022358	.	yellow protein [Source:Projected from Anopheles gambiae (AGAP003096) VB Community Annotation]	-3,08	AFUN010951	.	StAR-related lipid transfer protein [Source:Projected from Anopheles gambiae (AGAP004792) VB Community Annotation]	2,37
AFUN003358	.	.	-3,08	AFUN020552	.	sensory appendage protein [Source:Projected from Anopheles gambiae (AGAP008054) VB Community Annotation]	2,39
AFUN001655	.	myosin alkali light chain [Source:Projected from Anopheles gambiae (AGAP001569) VB Community Annotation]	-3,08	AFUN001599	.	.	2,39
AFUN020600	.	.	-3,07	AFUN016591	.	trypsin [Source:Projected from Anopheles gambiae (AGAP008294) VB Community Annotation]	2,39
AFUN004727	.	.	-3,06	AFUN020106	.	.	2,39
AFUN020921	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP010122) VB Community Annotation]	-3,06	AFUN020875	.	HtrA serine peptidase [Source:Projected from Anopheles gambiae (AGAP000240) VB Community Annotation]	2,40
AFUN008292	.	.	-3,05	AFUN009205	.	glycogen debranching enzyme [Source:Projected from Anopheles gambiae (AGAP001200) VB Community Annotation]	2,40
AFUN020897	.	.	-3,03	AFUN022044	.	.	2,41
AFUN008447	.	.	-3,03	AFUN008039	.	Nucleoside diphosphate kinase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RP58]	2,41
AFUN020037	.	hyaluronoglucosaminidase [Source:Projected from Anopheles gambiae (AGAP011479) VB Community Annotation]	-3,02	AFUN021853	.	juvenile hormone acid methyltransferase [Source:Projected from Anopheles gambiae (AGAP005256) VB Community Annotation]	2,42
AFUN020947	.	chymotrypsin [Source:Projected from Anopheles gambiae (AGAP006710) VB Community Annotation]	-3,02	AFUN016094	.	.	2,42
AFUN010592	.	.	-3,02	AFUN008710	.	.	2,43
AFUN020371	.	.	-3,02	AFUN000502	.	.	2,43
AFUN003856	.	.	-3,01	AFUN014167	.	.	2,44
AFUN018700	.	.	-3,01	AFUN007480	.	serine protease (non-digestive) [Source:Projected from Anopheles gambiae (AGAP006416) VB Community Annotation]	2,44
AFUN006657	.	.	-3,01	AFUN001882	Gr33	gustatory receptor [Source:VB Community Annotation]	2,47
AFUN006639	.	26S proteasome 26S non-ATPase subunit N11 [Source:Projected from Anopheles gambiae (AGAP008617) VB Community Annotation]	-3,00	AFUN004246	.	Acetyl-coenzyme A synthetase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RDH4]	2,47

AFUN010548	.	.	-3,00	AFUN006763	.	muskelin [Source:Projected from Anopheles gambiae (AGAP004884) VB Community Annotation]	2,47
AFUN016579	.	.	-3,00	AFUN016135	.	.	2,47
AFUN006396	.	matrix metalloprotease [Source:Projected from Anopheles gambiae (AGAP006904) VB Community Annotation]	-3,00	AFUN002138	.	.	2,48
AFUN002913	.	.	-2,99	AFUN007484	.	.	2,48
AFUN016235	.	serine-threonine kinase receptor-associated protein [Source:VB Community Annotation]	-2,99	AFUN000091	.	Iron--sulfur cluster insertion protein erpA [Source:Projected from Anopheles gambiae (AGAP010072) VB Community Annotation]	2,49
AFUN018583	.	.	-2,98	AFUN015810	GSTE4	Glutathione S-transferase [Source:UniProtKB/TrEMBL%3BAcc:A0A182S4T9]	2,49
AFUN022202	.	.	-2,98	AFUN008843	.	zinc finger protein GLIS2 [Source:Projected from Anopheles gambiae (AGAP006736) VB Community Annotation]	2,49
AFUN014940	.	.	-2,98	AFUN016368	.	lysosomal Pro-X carboxypeptidase [Source:VB Community Annotation]	2,50
AFUN019423	.	.	-2,98	AFUN009100	LRIM20	leucine rich immune protein (Coil-less) [Source:VB Community Annotation]	2,50
AFUN022057	.	.	-2,97	AFUN022152	.	.	2,51
AFUN020715	.	.	-2,97	AFUN016248	.	.	2,52
AFUN008095	.	.	-2,96	AFUN015739	CYP307A1	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:A0A182S4L5]	2,52
AFUN019721	.	.	-2,96	AFUN017757	LSU_rRNA_eukarya	Eukaryotic large subunit ribosomal RNA	2,52
AFUN021649	.	.	-2,96	AFUN009998	.	.	2,53
AFUN010444	.	.	-2,95	AFUN002945	.	.	2,53
AFUN020490	.	spaetzle-like cytokine [Source:Projected from Anopheles gambiae (AGAP008360) VB Community Annotation]	-2,95	AFUN001636	.	.	2,53
AFUN002544	.	.	-2,95	AFUN002661	.	.	2,54
AFUN022342	.	purine-nucleoside phosphorylase [Source:Projected from Anopheles gambiae (AGAP005945) VB Community Annotation]	-2,94	AFUN015853	PPO4	Prophenoloxidase 4 [Source:UniProtKB/TrEMBL%3BAcc:A0A182S4Y3]	2,54
AFUN021428	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP003390) VB Community Annotation]	-2,93	AFUN011641	.	.	2,55
AFUN020209	.	.	-2,92	AFUN006850	.	Niemann-Pick C2 protein [Source:Projected from Anopheles gambiae (AGAP012352) VB Community Annotation]	2,56
AFUN019863	.	.	-2,92	AFUN018665	.	.	2,56
AFUN008426	GSTU2	glutathione S-transferase [Source:VB Community Annotation]	-2,91	AFUN007328	.	alanine-glyoxylate aminotransferase 2-like [Source:Projected from Anopheles gambiae	2,58

						(AGAP000313) VB Community Annotation]	
AFUN009132	.	.	-2,91	AFUN021469	.	acid phosphatase [Source:Projected from Anopheles gambiae (AGAP004591) VB Community Annotation]	2,59
AFUN018720	.	.	-2,90	AFUN005926	.	inorganic phosphate cotransporter [Source:Projected from Anopheles gambiae (AGAP008931) VB Community Annotation]	2,59
AFUN022282	.	cystinosin [Source:Projected from Anopheles gambiae (AGAP004115) VB Community Annotation]	-2,90	AFUN006135	CYP4C36	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:A0A182RIS0]	2,59
AFUN001587	.	.	-2,89	AFUN018698	.	ubiquitin-conjugating enzyme E2 H [Source:Projected from Anopheles gambiae (AGAP001068) VB Community Annotation]	2,59
AFUN004974	.	.	-2,89	AFUN006903	.	prostaglandin reductase [Source:Projected from Anopheles gambiae (AGAP000414) VB Community Annotation]	2,59
AFUN016504	.	.	-2,89	AFUN000613	.	5-aminolevulinate synthase [Source:UniProtKB/TrEMBL%3BAcc:A0A182R372]	2,60
AFUN015894	CYP4H26	cytochrome P450 [Source:VB Community Annotation]	-2,88	AFUN004873	.	.	2,60
AFUN006579	.	ubiquitin-conjugating enzyme E2 R [Source:Projected from Anopheles gambiae (AGAP008680) VB Community Annotation]	-2,87	AFUN020656	.	acetyl-CoA carboxylase / biotin carboxylase [Source:Projected from Anopheles gambiae (AGAP005175) VB Community Annotation]	2,61
AFUN007571	.	.	-2,87	AFUN016241	.	Carboxylic ester hydrolase [Source:VB Community Annotation]	2,61
AFUN008350	.	.	-2,87	AFUN021057	.	protein phosphatase 1%2C regulatory (inhibitor) subunit [Source:Projected from Anopheles gambiae (AGAP001520) VB Community Annotation]	2,61
AFUN021779	.	cyanogenic beta-glucosidase [Source:Projected from Anopheles gambiae (AGAP006426) VB Community Annotation]	-2,87	AFUN018778	.	.	2,62
AFUN002122	.	.	-2,87	AFUN000100	.	.	2,62
AFUN004701	.	.	-2,86	AFUN016614	.	.	2,62
AFUN003620	.	UDP-N-acetylglucosamine pyrophosphorylase [Source:Projected from Anopheles gambiae (AGAP007889) UniProtKB/TrEMBL%3BAcc:A0A1S4GWJ6]	-2,84	AFUN014058	.	.	2,62
AFUN016373	.	.	-2,84	AFUN021041	.	.	2,62
AFUN021089	.	.	-2,84	AFUN004258	.	.	2,63
AFUN005374	.	phosphoserine phosphatase [Source:Projected from Anopheles gambiae (AGAP012247) VB Community Annotation]	-2,83	AFUN006702	CYP6AF1	cytochrome P450 [Source:VB Community Annotation]	2,63
AFUN006859	CYP18A1	cytochrome P450 [Source:VB Community Annotation]	-2,82	AFUN009018	.	.	2,63
AFUN018952	.	.	-2,82	AFUN004322	.	cytochrome b5 outer mitochondrial membrane	2,64

						isoform [Source:Projected from Anopheles gambiae (AGAP008363) VB Community Annotation]	
AFUN008082	.	Dihydropyrimidine dehydrogenase [NADP(+)] [Source:UniProtKB/TrEMBL%3BAcc:A0A182RP99]	-2,82	AFUN020278	.	odorant-binding protein [Source:Projected from Anopheles gambiae (AGAP002905) VB Community Annotation]	2,64
AFUN003806	.	.	-2,81	AFUN001956	.	facilitated glucose transporter (solute carrier family 2) [Source:Projected from Anopheles gambiae (AGAP011680) VB Community Annotation]	2,64
AFUN004557	.	Phosphoserine aminotransferase [Source:Projected from Anopheles gambiae (AGAP004598) UniProtKB/TrEMBL%3BAcc:Q5TRW7]	-2,81	AFUN020407	.	Phosphate transporter [Source:UniProtKB/TrEMBL%3BAcc:A0A182RX55]	2,64
AFUN007971	.	tropomyosin [Source:Projected from Anopheles gambiae (AGAP001799) VB Community Annotation]	-2,80	AFUN002978	CYP314A1	cytochrome P450 [Source:VB Community Annotation]	2,64
AFUN019611	.	.	-2,80	AFUN019403	.	.	2,65
AFUN001697	.	.	-2,80	AFUN016429	LRIM5	leucine-rich immune protein (Short) [Source:VB Community Annotation]	2,66
AFUN009675	.	.	-2,79	AFUN001787	.	.	2,66
AFUN004369	.	.	-2,79	AFUN022022	.	acyl-CoA-binding protein [Source:Projected from Anopheles gambiae (AGAP007460) VB Community Annotation]	2,68
AFUN020143	.	.	-2,79	AFUN022272	.	.	2,69
AFUN019490	.	Serine protease inhibitor (serpin) 9 [Source:UniProtKB/TrEMBL%3BAcc:A0A182RYC2]	-2,78	AFUN020726	.	.	2,70
AFUN021310	.	.	-2,78	AFUN021661	.	.	2,70
AFUN006011	.	.	-2,78	AFUN022284	.	CLIP-domain serine protease [Source:Projected from Anopheles gambiae (AGAP004148) VB Community Annotation]	2,70
AFUN004002	.	argininosuccinate lyase [Source:Projected from Anopheles gambiae (AGAP008141) VB Community Annotation]	-2,78	AFUN001497	.	.	2,73
AFUN010015	.	.	-2,77	AFUN020939	.	glyoxalase domain-containing protein [Source:Projected from Anopheles gambiae (AGAP000367) VB Community Annotation]	2,73
AFUN006340	.	.	-2,77	AFUN002338	.	.	2,73
AFUN009937	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006864) VB Community Annotation]	-2,77	AFUN010490	.	RBR-type E3 ubiquitin transferase [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HDH3]	2,74
AFUN001868	.	.	-2,75	AFUN006606	.	.	2,74
AFUN004884	CYP305A1	cytochrome P450 [Source:VB Community Annotation]	-2,75	AFUN022305	.	.	2,74
AFUN001921	.	all-trans-retinol dehydrogenase (NAD+) [Source:Projected from Anopheles gambiae	-2,74	AFUN020004	.	.	2,75

		(AGAP010232) VB Community Annotation]					
AFUN018999	.	.	-2,74	AFUN010660	.	Mpv17-like protein [Source:Projected from Anopheles gambiae (AGAP004046) VB Community Annotation]	2,76
AFUN011604	.	.	-2,74	AFUN015801	CYP6P2	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:B5AII2]	2,77
AFUN004461	.	.	-2,73	AFUN017339	SSU_rRNA_eukarya	Eukaryotic small subunit ribosomal RNA	2,77
AFUN015858	.	.	-2,72	AFUN010441	.	.	2,78
AFUN010170	.	.	-2,71	AFUN002547	.	Eupolytin [Source:Projected from Anopheles gambiae (AGAP011477) UniProtKB/TrEMBL%3BAcc:A0A1S4H7M9]	2,79
AFUN022218	.	lysosomal acid lipase [Source:Projected from Anopheles gambiae (AGAP013368) VB Community Annotation]	-2,71	AFUN002514	.	carboxylesterase [Source:Projected from Anopheles gambiae (AGAP011509) VB Community Annotation]	2,79
AFUN002875	.	mitochondrial carnitine/acylcarnitine carrier protein [Source:Projected from Anopheles gambiae (AGAP009128) VB Community Annotation]	-2,70	AFUN001339	.	.	2,79
AFUN004189	.	asparagine synthase (glutamine-hydrolysing) [Source:Projected from Anopheles gambiae (AGAP003833) VB Community Annotation]	-2,70	AFUN022078	.	meiotic chromosome segregation protein [Source:Projected from Anopheles gambiae (AGAP003239) VB Community Annotation]	2,79
AFUN019283	.	.	-2,69	AFUN019026	.	fibrinogen and fibronectin [Source:Projected from Anopheles gambiae (AGAP012000) VB Community Annotation]	2,80
AFUN004029	.	.	-2,69	AFUN016226	.	salivary gland protein [Source:VB Community Annotation]	2,81
AFUN019503	.	.	-2,68	AFUN018819	.	.	2,81
AFUN020482	.	Tropomyosin [Source:Projected from Anopheles gambiae (AGAP001797) VB Community Annotation]	-2,68	AFUN020427	.	inorganic phosphate cotransporter [Source:Projected from Anopheles gambiae (AGAP012251) VB Community Annotation]	2,81
AFUN021493	.	phosphoribosylamine --glycine ligase / phosphoribosylglycinamide formyltransfer [Source:Projected from Anopheles gambiae (AGAP009786) VB Community Annotation]	-2,68	AFUN008888	.	female reproductive tract protease GLEANR_896 [Source:Projected from Anopheles gambiae (AGAP008861) VB Community Annotation]	2,81
AFUN015982	.	ATP-binding cassette transporter (ABC transporter) family G member [Source:Projected from Anopheles gambiae (AGAP009472) UniProtKB/TrEMBL%3BAcc:A0A1S4H231]	-2,68	AFUN021736	.	.	2,82
AFUN010508	.	synaptosomal-associated protein%2C 29kDa [Source:Projected from Anopheles gambiae (AGAP011135) VB Community Annotation]	-2,68	AFUN017736	LSU_rRNA_eukarya	Eukaryotic large subunit ribosomal RNA	2,82

AFUN008585	.	polyadenylate-binding protein-interacting protein [Source:Projected from Anopheles gambiae (AGAP011627) VB Community Annotation]	-2,68	AFUN019789	.	.	2,84
AFUN002929	.	26S proteasome 26S non-ATPase subunit N3 [Source:Projected from Anopheles gambiae (AGAP009082) VB Community Annotation]	-2,67	AFUN021595	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP009874) VB Community Annotation]	2,85
AFUN004424	.	.	-2,66	AFUN009036	.	.	2,86
AFUN006204	.	tubulin polyglutamylase TTLL4 [Source:Projected from Anopheles gambiae (AGAP009150) VB Community Annotation]	-2,66	AFUN018766	.	.	2,86
AFUN009676	.	Cysteine dioxygenase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RTS3]	-2,66	AFUN018633	.	.	2,87
AFUN010121	.	.	-2,66	AFUN020265	.	CUG-BP- and ETR3-like factor [Source:Projected from Anopheles gambiae (AGAP004950) VB Community Annotation]	2,88
AFUN010339	.	F-box and leucine-rich repeat protein 2/20 [Source:Projected from Anopheles gambiae (AGAP007807) UniProtKB/TrEMBL%3BAcc:A0A1S4GWA2]	-2,65	AFUN002137	.	Malic enzyme [Source:UniProtKB/TrEMBL%3BAcc:A0A182R7I7]	2,88
AFUN019025	.	.	-2,65	AFUN004325	.	.	2,89
AFUN006705	.	dsRNase [Source:Projected from Anopheles gambiae (AGAP011025) VB Community Annotation]	-2,65	AFUN007272	.	protein turtle [Source:Projected from Anopheles gambiae (AGAP008943) VB Community Annotation]	2,89
AFUN007090	.	.	-2,65	AFUN004972	.	.	2,90
AFUN004480	.	.	-2,65	AFUN022249	.	.	2,90
AFUN019836	.	.	-2,64	AFUN004637	.	.	2,91
AFUN002557	.	.	-2,64	AFUN009128	.	Nat13 protein [Source:Projected from Anopheles gambiae (AGAP002568) VB Community Annotation]	2,91
AFUN007492	.	.	-2,64	AFUN006610	GAM1	gambicin [Source:VB Community Annotation]	2,91
AFUN016217	.	Cuticular protein [Source:UniProtKB/TrEMBL%3BAcc:A0A1I8JUT6]	-2,63	AFUN006656	.	long-chain-fatty-acid--CoA ligase ACSBG [Source:Projected from Anopheles gambiae (AGAP008596) VB Community Annotation]	2,92
AFUN015846	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP005996) VB Community Annotation]	-2,63	AFUN007532	.	.	2,94
AFUN008088	.	.	-2,62	AFUN001383	CYP9J5	cytochrome P450 [Source:VB Community Annotation]	2,94
AFUN010958	.	fatty acyl-CoA reductase [Source:Projected from Anopheles gambiae (AGAP004784) VB Community Annotation]	-2,61	AFUN010593	.	.	2,95
AFUN021474	.	fatty acyl-CoA reductase [Source:Projected from Anopheles gambiae	-2,60	AFUN008757	.	salivary gland protein [Source:Projected from Anopheles gambiae	2,95

		(AGAP005984) VB Community Annotation]				(AGAP006507) VB Community Annotation]	
AFUN003878	.	E2 ubiquitin-conjugating enzyme E2D%2C effete [Source:Projected from Anopheles gambiae (AGAP000145) VB Community Annotation]	-2,60	AFUN020975	.	.	2,96
AFUN002720	.	glutaminyl-peptide cyclotransferase precursor [Source:Projected from Anopheles gambiae (AGAP006610) VB Community Annotation]	-2,60	AFUN011058	.	.	2,96
AFUN003140	Prosalpha6	26S proteasome alpha 6 subunit [Source:VB Community Annotation]	-2,60	AFUN017741	LSU_rRNA_eukarya	Eukaryotic large subunit ribosomal RNA	2,97
AFUN010004	.	.	-2,60	AFUN003850	GAPDH	glyceraldehyde 3-phosphate dehydrogenase [Source:VB Community Annotation]	2,97
AFUN020651	.	.	-2,23	AFUN010489	.	Synaptic vesicle glycoprotein 2B [Source:Projected from Anopheles gambiae (AGAP010892) VB Community Annotation]	2,98
AFUN010813	.	.	-2,22	AFUN009189	.	.	3,00
AFUN018974	.	.	-2,22	AFUN022155	.	.	3,00
AFUN003023	.	.	-2,22	AFUN001632	.	Ser/Thr protein phosphatase/nucleotidase [Source:Projected from Anopheles gambiae (AGAP001601) VB Community Annotation]	3,00
AFUN022223	.	alpha-tocopherol transfer protein-like protein [Source:Projected from Anopheles gambiae (AGAP004762) VB Community Annotation]	-2,22	AFUN019282	.	.	3,03
AFUN016538	.	.	-2,22	AFUN021431	.	.	3,06
AFUN018748	.	.	-2,22	AFUN015785	CYP6AA2	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:B5AIH8]	3,06
AFUN010746	.	.	-2,21	AFUN005177	.	ubiquitin thioesterase CYLD [Source:Projected from Anopheles gambiae (AGAP008412) VB Community Annotation]	3,07
AFUN019386	.	.	-2,21	AFUN008871	.	formyltetrahydrofolate dehydrogenase [Source:Projected from Anopheles gambiae (AGAP009591) VB Community Annotation]	3,07
AFUN001331	.	.	-2,21	AFUN005510	.	.	3,09
AFUN010828	.	RING-box protein [Source:Projected from Anopheles gambiae (AGAP005202) VB Community Annotation]	-2,21	AFUN016064	.	angiotensin-converting enzyme [Source:VB Community Annotation]	3,11
AFUN007773	.	.	-2,21	AFUN001368	.	ornithine decarboxylase [Source:Projected from Anopheles gambiae (AGAP012308) VB Community Annotation]	3,11
AFUN021732	.	outer membrane lipoprotein Blc [Source:Projected from Anopheles gambiae (AGAP002592) VB Community Annotation]	-2,21	AFUN016627	.	NADPH oxidase [Source:Projected from Anopheles gambiae (AGAP003772) VB Community Annotation]	3,13

AFUN004285	.	TatD DNase family protein [Source:Projected from Anopheles gambiae (AGAP006162) VB Community Annotation]	-2,21	AFUN020678	.	glucose dehydrogenase (acceptor) [Source:Projected from Anopheles gambiae (AGAP002557) VB Community Annotation]	3,15
AFUN007358	.	.	-2,20	AFUN007917	.	ATP-dependent Clp protease ATP-binding subunit clpX-like%2C mitochondrial [Source:Projected from Anopheles gambiae (AGAP002996) VB Community Annotation]	3,16
AFUN001254	.	.	-2,20	AFUN005237	.	Kynurenine formamidase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RG95]	3,18
AFUN019662	.	.	-2,20	AFUN002520	.	.	3,20
AFUN002774	.	BTB/POZ domain-containing protein [Source:Projected from Anopheles gambiae (AGAP006777) VB Community Annotation]	-2,20	AFUN001196	.	sodium-coupled monocarboxylate transporter [Source:Projected from Anopheles gambiae (AGAP001550) VB Community Annotation]	3,20
AFUN011451	.	WD repeat-containing protein [Source:Projected from Anopheles gambiae (AGAP000678) VB Community Annotation]	-2,20	AFUN018961	.	.	3,23
AFUN003663	.	E3 ubiquitin protein ligase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RBU8]	-2,20	AFUN021886	.	.	3,26
AFUN020743	.	.	-2,20	AFUN006491	.	.	3,26
AFUN004717	.	cullin [Source:Projected from Anopheles gambiae (AGAP003605) VB Community Annotation]	-2,20	AFUN010918	CYP6N1	cytochrome P450 [Source:VB Community Annotation]	3,26
AFUN020990	.	26S proteasome regulatory subunit T6 [Source:Projected from Anopheles gambiae (AGAP007243) VB Community Annotation]	-2,20	AFUN006921	.	diphthamide biosynthesis protein [Source:Projected from Anopheles gambiae (AGAP011288) VB Community Annotation]	3,27
AFUN000403	.	.	-2,19	AFUN001018	.	duplex-specific nuclease [Source:Projected from Anopheles gambiae (AGAP011696) VB Community Annotation]	3,28
AFUN021064	.	.	-2,19	AFUN002641	.	Serine--pyruvate aminotransferase [Source:UniProtKB/TrEMBL%3BAcc:A0A182R8Y1]	3,28
AFUN007741	Pelle	TOLL pathway signalling Ser/Thr kinase [Source:VB Community Annotation]	-2,19	AFUN020023	.	influenza virus NS1A-binding protein [Source:Projected from Anopheles gambiae (AGAP003823) VB Community Annotation]	3,31
AFUN006918	.	Protein flightless-1 [Source:Projected from Anopheles gambiae (AGAP011292) VB Community Annotation]	-2,19	AFUN011197	.	.	3,34
AFUN000738	.	DE-cadherin [Source:Projected from Anopheles gambiae (AGAP007203) VB Community Annotation]	-2,19	AFUN006671	PPO2	Prophenoloxidase 2 [Source:UniProtKB/TrEMBL%3BAcc:A0A182RKA3]	3,37
AFUN016197	.	.	-2,19	AFUN002058	.	glucuronosyltransferase [Source:Projected from Anopheles gambiae (AGAP007374) VB Community Annotation]	3,38
AFUN003557	.	ATP-binding cassette transporter (ABC transporter) family C member [Source:Projected from	-2,18	AFUN001382	CYP9J5	cytochrome P450 [Source:VB Community Annotation]	3,38

		Anopheles gambiae (AGAP009835) VB Community Annotation]					
AFUN021539	.	growth arrest and DNA-damage-inducible protein [Source:Projected from Anopheles gambiae (AGAP007651) VB Community Annotation]	-2,18	AFUN019376	.	.	3,39
AFUN002845	.	.	-2,18	AFUN022260	.	.	3,42
AFUN018584	.	.	-2,18	AFUN019146	.	.	3,43
AFUN002121	.	26S proteasome non-ATPase regulatory subunit [Source:Projected from Anopheles gambiae (AGAP000201) UniProtKB/TrEMBL%3BAcc:Q7QEX9]	-2,18	AFUN005543	.	.	3,43
AFUN010017	.	defense protein [Source:Projected from Anopheles gambiae (AGAP008878) VB Community Annotation]	-2,18	AFUN008711	.	alpha-N-acetylglucosaminidase [Source:Projected from Anopheles gambiae (AGAP011750) VB Community Annotation]	3,45
AFUN020333	.	.	-2,18	AFUN001129	.	Ecdysteroid-regulated protein [Source:Projected from Anopheles gambiae (AGAP002804) VB Community Annotation]	3,45
AFUN021339	.	.	-2,18	AFUN018524	.	.	3,48
AFUN009096	.	.	-2,18	AFUN015854	PPO5	prophenoloxidase 5 [Source:VB Community Annotation]	3,48
AFUN020801	.	Ras-related and estrogen-regulated growth inhibitor-like protein [Source:Projected from Anopheles gambiae (AGAP003577) VB Community Annotation]	-2,18	AFUN005947	.	.	3,49
AFUN008458	.	.	-2,17	AFUN001894	.	.	3,54
AFUN018750	.	.	-2,17	AFUN019141	.	.	3,54
AFUN009006	.	.	-2,17	AFUN022154	.	F-box and leucine-rich repeat protein [Source:Projected from Anopheles gambiae (AGAP011930) VB Community Annotation]	3,56
AFUN006962	.	neutral ceramidase [Source:Projected from Anopheles gambiae (AGAP000973) VB Community Annotation]	-2,17	AFUN006915	DEF1	defensin anti-microbial peptide [Source:VB Community Annotation]	3,56
AFUN000046	.	protein disulfide-isomerase A6 [Source:Projected from Anopheles gambiae (AGAP001919) VB Community Annotation]	-2,17	AFUN022227	.	Niemann-Pick Type C-2 [Source:Projected from Anopheles gambiae (AGAP013106) VB Community Annotation]	3,60
AFUN006385	.	chitinase [Source:Projected from Anopheles gambiae (AGAP006898) VB Community Annotation]	-2,17	AFUN015792	CYP6P9A	cytochrome P450 [Source:VB Community Annotation]	3,61
AFUN020357	.	decaprenyl-diphosphate synthase subunit [Source:Projected from Anopheles gambiae (AGAP000591) VB Community Annotation]	-2,16	AFUN003608	.	excitatory amino acid transporter 2 isoform [Source:Projected from Anopheles gambiae (AGAP007903) VB Community Annotation]	3,62
AFUN021023	.	.	-2,16	AFUN005224	.	lysophosphatidate acyltransferase [Source:Projected from Anopheles gambiae	3,63

						(AGAP009415) VB Community Annotation]	
AFUN002275	CYP301A1	cytochrome P450 [Source:VB Community Annotation]	-2,16	AFUN004162	.	.	3,63
AFUN019937	.	Similar to mandelate racemase/muconate lactonizing protein [Source:Projected from Anopheles gambiae (AGAP007421) VB Community Annotation]	-2,15	AFUN014744	.	.	3,63
AFUN022367	.	.	-2,15	AFUN001169	.	.	3,64
AFUN010295	.	.	-2,15	AFUN003583	.	.	3,66
AFUN015251	.	.	-2,15	AFUN005319	.	.	3,73
AFUN019313	.	.	-2,14	AFUN022061	.	.	3,75
AFUN002179	.	Fatty acyl-CoA reductase [Source:Projected from Anopheles gambiae (AGAP002279) UniProtKB/TrEMBL%3BAcc:Q7QKD9]	-2,14	AFUN011947	.	.	3,77
AFUN001420	.	BCL2/adenovirus E1B 19 kDa protein-interacting protein [Source:Projected from Anopheles gambiae (AGAP001411) VB Community Annotation]	-2,14	AFUN008006	.	Lipid storage droplets surface-binding protein [Source:Projected from Anopheles gambiae (AGAP002890) VB Community Annotation]	3,78
AFUN015123	.	.	-2,14	AFUN021924	.	.	3,80
AFUN008495	.	.	-2,13	AFUN005751	Tsf1	Transferrin [Source:UniProtKB/TrEMBL%3BAcc:A0A182RHN9]	3,91
AFUN000597	.	.	-2,13	AFUN004114	.	.	3,92
AFUN015893	.	Elongation of very long chain fatty acids protein [Source:UniProtKB/TrEMBL%3BAcc:A0A182S523]	-2,13	AFUN010408	.	MFS transporter%2C VNT family%2C synaptic vesicle glycoprotein [Source:Projected from Anopheles gambiae (AGAP000128) VB Community Annotation]	3,92
AFUN007900	.	26S proteasome regulatory subunit T3 [Source:Projected from Anopheles gambiae (AGAP003008) VB Community Annotation]	-2,13	AFUN000380	Pgk	Phosphoglycerate kinase [Source:UniProtKB/TrEMBL%3BAcc:A0A182R2I9]	3,93
AFUN010859	.	.	-2,13	AFUN005035	.	.	3,96
AFUN021960	.	.	-2,13	AFUN016172	.	Defective in cullin neddylation protein [Source:UniProtKB/TrEMBL%3BAcc:A0A1I8JUP2]	3,97
AFUN006868	.	homeobox protein abdominal-A homolog [Source:Projected from Anopheles gambiae (AGAP004662) VB Community Annotation]	-2,13	AFUN009834	.	.	3,98
AFUN018532	.	.	-2,13	AFUN021117	.	membrane dipeptidase [Source:Projected from Anopheles gambiae (AGAP011654) VB Community Annotation]	3,98
AFUN016209	.	sulfotransferase [Source:Projected from Anopheles gambiae (AGAP029785) VB Community Annotation]	-2,13	AFUN003779	.	.	3,98
AFUN019827	.	.	-2,12	AFUN019639	.	.	4,03
AFUN021229	.	.	-2,12	AFUN016382	.	Transporter [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HEC4]	4,03

AFUN000789	.	fimbrin [Source:Projected from Anopheles gambiae (AGAP004071) VB Community Annotation]	-2,12	AFUN020976	.	.	4,06
AFUN004914	.	Proteasome subunit beta [Source:UniProtKB/TrEMBL%3BAcc:A0A182RFD2]	-2,12	AFUN009418	.	Glycerol-3-phosphate dehydrogenase [Source:UniProtKB/TrEMBL%3BAcc:A0A182RT17]	4,12
AFUN021863	.	.	-2,12	AFUN011754	.	stearoyl-CoA desaturase (delta-9 desaturase) [Source:Projected from Anopheles gambiae (AGAP001713) VB Community Annotation]	4,12
AFUN020252	.	.	-2,11	AFUN009037	.	.	4,17
AFUN001205	.	.	-2,11	AFUN020614	.	trypsin type [Source:Projected from Anopheles gambiae (AGAP004900) VB Community Annotation]	4,20
AFUN011760	.	.	-2,11	AFUN004426	.	.	4,21
AFUN002111	.	elongation of very long chain fatty acids [Source:Projected from Anopheles gambiae (AGAP011812) VB Community Annotation]	-2,11	AFUN008655	.	alpha-crystallin chain B [Source:Projected from Anopheles gambiae (AGAP007160) VB Community Annotation]	4,24
AFUN020148	.	transcription factor CP2 and related proteins [Source:Projected from Anopheles gambiae (AGAP005564) VB Community Annotation]	-2,10	AFUN019627	.	.	4,27
AFUN021410	.	alpha-glucosidase [Source:Projected from Anopheles gambiae (AGAP012400) VB Community Annotation]	-2,10	AFUN007531	.	.	4,28
AFUN001204	.	.	-2,10	AFUN015786	CYP6AA1	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:B5AIH9]	4,32
AFUN015732	.	.	-2,10	AFUN016455	.	D7 short form salivary protein [Source:VB Community Annotation]	4,33
AFUN009548	.	26S proteasome regulatory subunit T1 [Source:Projected from Anopheles gambiae (AGAP011174) UniProtKB/TrEMBL%3BAcc:Q7QHG0]	-2,10	AFUN008882	.	mitochondrial uncoupling protein (solute carrier family 25) [Source:Projected from Anopheles gambiae (AGAP009603) VB Community Annotation]	4,38
AFUN004479	Prosalpha3	26S proteasome alpha 3 subunit [Source:VB Community Annotation]	-2,10	AFUN018553	.	.	4,40
AFUN003595	.	xanthine dehydrogenase/oxidase [Source:Projected from Anopheles gambiae (AGAP007918) VB Community Annotation]	-2,09	AFUN015823	.	cecropin anti-microbial peptide [Source:Projected from Anopheles gambiae (AGAP000692) VB Community Annotation]	4,42
AFUN019712	.	26S proteasome regulatory subunit T2 [Source:Projected from Anopheles gambiae (AGAP003215) VB Community Annotation]	-2,09	AFUN003124	.	late trypsin [Source:Projected from Anopheles gambiae (AGAP010240) VB Community Annotation]	4,56
AFUN007188	.	.	-2,09	AFUN019139	.	.	4,63
AFUN021427	.	cuticular protein RR-2 family [Source:Projected from Anopheles gambiae (AGAP003375) VB Community Annotation]	-2,08	AFUN000569	.	.	4,63
AFUN011715	Crc	Calreticulin [Source:UniProtKB/TrEMBL%3BAcc:A0A182RZI4]	-2,08	AFUN002611	.	Alkaline phosphatase [Source:Projected from Anopheles gambiae (AGAP001684)	4,66

						UniProtKB/TrEMBL%3BAcc:Q7PY02]	
AFUN003885	.	.	-2,08	AFUN016592	.	trypsin [Source:Projected from Anopheles gambiae (AGAP008296) VB Community Annotation]	4,66
AFUN010657	.	MAD%2C mothers against decapentaplegic interacting protein [Source:Projected from Anopheles gambiae (AGAP004049) VB Community Annotation]	-2,08	AFUN011483	.	.	4,73
AFUN022149	.	.	-2,08	AFUN021295	.	.	4,82
AFUN018567	.	.	-2,08	AFUN002545	.	Metalloendopeptidase [Source:UniProtKB/TrEMBL%3BAcc:A0A182R8N6]	4,83
AFUN003365	.	.	-2,08	AFUN021888	.	.	4,85
AFUN004059	.	.	-2,08	AFUN007883	.	.	4,98
AFUN001869	.	.	-2,08	AFUN004474	.	.	5,00
AFUN000508	.	.	-2,07	AFUN021889	.	.	5,05
AFUN020419	.	.	-2,07	AFUN004964	.	.	5,10
AFUN002221	.	Serine/threonine-protein phosphatase [Source:UniProtKB/TrEMBL%3BAcc:A0A182R7S1]	-2,07	AFUN004370	.	.	5,15
AFUN006414	.	.	-2,07	AFUN016374	NimB	nimrod B2 [Source:VB Community Annotation]	5,15
AFUN018617	.	.	-2,07	AFUN015889	CYP6P9b	cytochrome P450 [Source:VB Community Annotation]	5,21
AFUN005836	.	.	-2,06	AFUN021509	.	carboxypeptidase [Source:Projected from Anopheles gambiae (AGAP011443) VB Community Annotation]	5,21
AFUN010633	.	.	-2,06	AFUN020519	.	.	5,31
AFUN008754	.	.	-2,06	AFUN008000	.	lipophorin [Source:Projected from Anopheles gambiae (AGAP001826) VB Community Annotation]	5,59
AFUN009926	.	.	-2,05	AFUN018874	.	.	5,69
AFUN021211	.	.	-2,05	AFUN018787	.	.	5,77
AFUN018759	.	.	-2,05	AFUN003702	.	.	5,90
AFUN010604	HPX2	heme peroxidase 2 [Source:VB Community Annotation]	-2,05	AFUN006107	.	low molecular weight phosphotyrosine protein phosphatase [Source:Projected from Anopheles gambiae (AGAP009266) VB Community Annotation]	5,97
AFUN010396	.	Wnt inhibitory factor 1000117] [Source:Projected from Anopheles gambiae (AGAP000117) VB Community Annotation]	-2,05	AFUN021795	.	.	6,00
AFUN009147	.	dehydrogenase/reductase SDR family member 11 precursor [Source:Projected from Anopheles gambiae (AGAP010311) VB Community Annotation]	-2,05	AFUN016115	.	.	6,19
AFUN021107	.	.	-2,05	AFUN008613	.	.	6,20
AFUN008275	Prosalpha2	26S proteasome alpha 2 subunit [Source:VB Community Annotation]	-2,04	AFUN007240	.	.	6,22

AFUN016565	.	neuronal guanine nucleotide exchange factor [Source:Projected from Anopheles gambiae (AGAP007452) VB Community Annotation]	-2,04	AFUN003917	LRIM4	leucine-rich immune protein (Long) [Source:VB Community Annotation]	6,35
AFUN021454	.	.	-2,04	AFUN007842	.	.	6,61
AFUN020660	.	.	-2,04	AFUN011465	.	cecropin anti-microbial peptide [Source:Projected from Anopheles gambiae (AGAP000694) VB Community Annotation]	6,61
AFUN011341	.	.	-2,04	AFUN001630	.	mitochondrial 18 kDa protein [Source:Projected from Anopheles gambiae (AGAP001603) VB Community Annotation]	6,86
AFUN020871	.	.	-2,04	AFUN007395	.	acylphosphatase [Source:Projected from Anopheles gambiae (AGAP008311) VB Community Annotation]	6,87
AFUN010293	.	F-box protein [Source:Projected from Anopheles gambiae (AGAP011911) UniProtKB/TrEMBL%3BAcc:A0A1S4H7V0]	-2,03	AFUN019746	.	protein Mpv17 [Source:Projected from Anopheles gambiae (AGAP001403) VB Community Annotation]	6,94
AFUN006529	.	nuclear receptor subfamily 6 group A [Source:Projected from Anopheles gambiae (AGAP004693) VB Community Annotation]	-2,03	AFUN016001	.	.	7,08
AFUN006717	.	.	-2,03	AFUN007797	.	.	7,33
AFUN020693	.	UNC93-like protein MFSD11 [Source:Projected from Anopheles gambiae (AGAP000311) VB Community Annotation]	-2,03	AFUN007549	CYP9K1	Cytochrome P450 [Source:UniProtKB/TrEMBL%3BAcc:A0A182RMS7]	8,03
AFUN021063	.	.	-2,03	AFUN005303	.	.	8,09
AFUN022239	.	mating-induced stimulator of oogenesis MISO [Source:Projected from Anopheles gambiae (AGAP002620) VB Community Annotation]	-2,02	AFUN011059	.	.	8,20
AFUN001975	.	AarF domain containing kinase [Source:Projected from Anopheles gambiae (AGAP006050) VB Community Annotation]	-2,02	AFUN016449	.	Trypsin [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HDX0]	8,44
AFUN007103	.	.	-2,02	AFUN008514	.	.	8,46
AFUN019422	.	.	-2,02	AFUN008582	.	.	8,46
AFUN005624	.	.	-2,02	AFUN009918	.	cuticular protein RR-1 family [Source:Projected from Anopheles gambiae (AGAP006828) VB Community Annotation]	8,75
AFUN021467	.	sodium-independent sulfate anion transporter [Source:Projected from Anopheles gambiae (AGAP013218) VB Community Annotation]	-2,02	AFUN018668	.	.	9,50
AFUN009139	.	ubiquitin-conjugating enzyme E2 L3 [Source:Projected from Anopheles gambiae (AGAP001056) VB Community Annotation]	-2,02	AFUN018799	.	.	11,15

AFUN019003	.	.	-2,01	AFUN008930	.	cuticular protein CPLCG family [Source:Projected from Anopheles gambiae (AGAP008444) VB Community Annotation]	11,18
AFUN018485	.	.	-2,01	AFUN000200	.	calcium/calmodulin-dependent protein kinase (CaM kinase) II [Source:Projected from Anopheles gambiae (AGAP000028) VB Community Annotation]	11,30
AFUN019454	.	.	-2,01	AFUN019222	.	.	12,20
AFUN002920	.	.	-2,01	AFUN016384	.	Transporter [Source:UniProtKB/TrEMBL%3BAcc:A0A1Y9HDQ9]	14,43
AFUN001682	.	ATP-binding cassette transporter (ABC transporter) family G member [Source:Projected from Anopheles gambiae (AGAP002050) VB Community Annotation]	-2,01	AFUN011407	.	myofilin variant C [Source:Projected from Anopheles gambiae (AGAP004161) VB Community Annotation]	17,78
AFUN002944	.	.	-2,01	AFUN001929	.	.	19,32
AFUN004836	.	farnesyl diphosphate synthase [Source:Projected from Anopheles gambiae (AGAP007104) VB Community Annotation]	-2,01	AFUN010906	.	dual specificity phosphatase [Source:Projected from Anopheles gambiae (AGAP008228) VB Community Annotation]	19,85
AFUN010356	.	Protein quiver [Source:UniProtKB/TrEMBL%3BAcc:A0A182RVP6]	-2,01	AFUN020303	.	heme peroxidase [Source:Projected from Anopheles gambiae (AGAP003714) VB Community Annotation]	23,00
AFUN019355	.	.	-2,00	AFUN018474	.	.	35,38
AFUN020769	.	.	-2,00	AFUN020285	.	.	36,90
AFUN009289	.	.	-2,00	AFUN003379	.	.	57,01
AFUN007855	.	stearoyl-CoA desaturase (delta-9 desaturase) [Source:Projected from Anopheles gambiae (AGAP003050) VB Community Annotation]	-2,00	AFUN001928	.	.	62,94

8.5 **Plagiarism declaration**

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Jacek Zawada (Student number: 1122748) am a student registered for the degree of Doctor of Philosophy in the academic year 2021.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: _____ Date: 28th April 2021