Anti-viral immunity in *Anopheles gambiae* mosquitoes

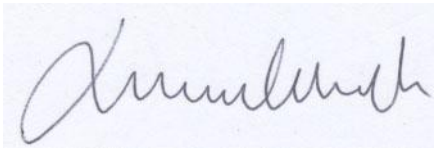
Joanna Waldock
Supervisor: Dr G Christophides

Division of Cell and Molecular Biology
Faculty of Natural Sciences
Imperial College London

A thesis submitted in accordance with the requirements of Imperial College London for the degree of Doctor of Philosophy, July 2010

Declaration of own work

I certify that this thesis and the research upon which it is based are a product of my own work. Any ideas, quotations, or publications from other authors are clearly stated in accordance with standard referencing practises and all scientific collaborations are acknowledged.



Joanna Waldock

29/06/2010

Abstract

Mosquito transmitted viruses (arboviruses) cause significant burden in much of the developing world. Little is known about mosquito responses to viral infection, and how these responses could be utilised to prevent spread of viral disease. *Anopheles gambiae*, the principal vector of human malaria, unusually transmits virtually no arboviruses, with one known exception - O'nyony-nyong Virus (ONNV). In this thesis the interactions between ONNV and the *A. gambiae* immune system were studied. Initially ONNV infection in *A. gambiae* mosquitoes and an *A. gambiae* derived cell line were characterised. The *in vivo* transcriptional responses of *A. gambiae* to viral infection were profiled using full genome microarrays, describing the global response to ONNV infection. This thesis demonstrates that the *A. gambiae* immune system does respond to viral infection, with genes covering all aspects of immunity being differentially regulated, from pathogen recognition to modulation of immune signalling, complement-mediated lysis/opsonisation and immune effector mechanisms.

Furthermore, this study identified four immune genes (a galectin, an MD2-like receptor and two lysozymes) regulated by ONNV infection that are capable of limiting virus during infection. These genes have novel roles in anti-viral immunity, and suggest previously uncharacterised mechanisms for targeting viral infection. Additionally, it is shown that *A. gambiae* uses a combination of core conserved anti-viral mechanisms, including RNAi, but does not utilise some signalling pathways reported to be anti-viral in other insects. This indicates that species specific mechanisms target viral infection. Finally this study demonstrates that foreign RNA acts as a Pathogen Associated Molecular Pattern (PAMP) in *A. gambiae* derived cells, and triggers transcriptional responses that dramatically reduce viral infection. In conclusion the data presented in this thesis demonstrate that *A. gambiae* responds to and is capable of limiting viral infection through conserved and novel immune mechanisms triggered by recognition of viral infection and foreign RNA.

Acknowledgements

I would first and foremost like to thank my supervisor, Dr George Christophides for all his guidance and help through the whole process of generating this thesis. I would like to thank Dr Brian Foy, for providing the viruses used during this thesis, and also to Rod Thompkins and Dr Irma Sanchez-Vargas for their training in the techniques used in this study. Many thanks go out to all the members of the Kafatos/Christophides laboratory who have helped me in innumerable ways throughout the years, in particular Dr Michael Povelones for answering *so* many questions and to Dr Ellen Bushell and Dr Rob Waterhouse who have kept me sane! Finally I would like to thank my husband, Nilly, for always having faith.

Collaborators and Contributions

Dr Brian Foy	Provided infectious clones of ONNV
Dr Amanda Jackson and Seth Redmond	Design of genome wide microarray platform
Dr Dina Vlachou and Dr Timm Schlegelmilsch	Assistance in analysis of microarray data
Dr Tibebu Habtewold and Kasia Sala	Mosquito rearing and animal work

Funding

This project was funded by a BBSRC Doctoral Training Grant to the Department of Life Sciences.

Table of contents

List of figures

List of tables

List of Appendices

Appendix 1

Appendix 2

List of abbreviations

A
Aa's – amino acids
AGO2 – Argonaute 2
AIDL – Arthopod Infectious Diseases Laboratory
AMP – Anti-microbial peptide
ATS – Alphavirus Transducing system

B
BHK – Baby hamster kidney cells
BSA – Bovine Serum Albumin

C
cDNA – complementary DNA
CEC - cecropin
CHIKV – Chickunguniya Virus
CLIP – Clip domain serine protease
cmRNA – copy messenger RNA
CRD – carbohydrate recognition domain
CTL – C-type lectin

D
DAPI - 4',6-Diamidino-2-Phenylindole
DCV – Drosophila C Virus
DEF - defensin
DENV – Dengue Virus
DMEM – Dulbecco's minimum essential media
DNA – deoxyribonucleic acid
dNTP – deoxyribonucleotide triphosphate
DOME – Domeless
Dpi – Days post infection
DPP - decapentaplegic
dsRNA – double stranded RNA

E
EDTA – Ethylenadiaminetetraacetic acid
EEEV – Eastern Equine Encephalitis Virus

F
FCS – Fetal Calf Serum
FHV – Flock House Virus
FREP – fibrinogen-like protein

G
GFP – Green fluorescent protein

GNBP – Gram negative binding protein
GO – gene ontology
GOI – Gene of interest

H
HIV – Human Immunodeficiency Virus
HOP – Hopskotch
Hpi – Hours post infection

I
IAP1 – inhibitor of apoptosis 1
IFA – immunofluorescence assay
IMD – Immune Deficiency
IM – intramuscular
IP - intraperitoneal
ITAM – immuno tyrosine activating motif

J
JAK/STAT – Janus Kinase/Signal Transducer and Activator of Transcription
JEV – Japanese Encephalitis Virus

K
KD – knock down

L
LB – lysogeny broth
LRIM – Leucine rich immune gene
LRR – leuice rich repeat

M
MEM – minimum essential media
mi-RNA – microRNA
miRNP – microRNA ribonucleoprotein complex
ML – MD2-like receptor
MMP – matrix metalloproteinase
MOI – multiplicity of infection
MTT – Thiozolyl Blue Tetrazolium Bromide

N
NoV – Nodamura Virus
nsP – non structural protein

O
ONNV – O'nyong-nyong Virus
ORF – Open reading frame

P

PBS – phosphate buffered saline
PCI – phenol/chloroform/iso-amyl alcohol
PCR – polymerase chain reaction
PFA - paraformaldehyde
PFU – Plaque forming units
PGRP – Peptidoglycan recognition protein
PIWI – P element induced wimpy testis
PkR – Protein kinase R
Ppan – Peter Pan
PRR – Pathogen recognition receptor
PRV – Pseudo Rabies Virus

Q

Qrt-PCR – Quantitative real time PCR

R

rasiRNA – repeat associated short interfering RNA
RBCs – red blood cells
RDRP – RNA dependent RNA polymerase
RISC – RNA induced silencing complex
RLH – Rig-1 like helicase
RLuc – Renilla/Lucinferase
RNA – ribonucleic acid
RNAse - ribonuclease
RNAi – RNA interference
Rpm – revolutions per minute
RRV – Ross River Virus
RT – room temperature

S

+ss – positive single stranded
SFV – Semliki Forest Virus
SINV – Sindbis Virus
siRNA – short interfering RNA
SRPN – Serpin
Sua - Suakoko

T

TEP – Thioester containing protein
TLR – Toll-like receptor
Tm – primer melting temperature
TSN – Tudor-SN

U

UPD – Unpaired

V

VERO – African Green monkey kidney cells
VIG – Vasa Intronic Gene
viRNA – virus specific short interfering RNA
Vir-1 – virally regulated 1

W

WEEV – Western Equine Encephalitis Virus
WNT – wnt signalling pathway
WNV – West Nile Virus

Y

YFV – Yellow Fever Virus

1 Introduction

1.1 The importance of mosquitoes to human health

Mosquitoes belong to the order of the two winged flies called the Diptera, Suborder Nematocera, Family Culicidae, where some 3000 or so species have been painstakingly grouped and ordered. Mosquitoes have been living alongside, and biting, man for around one million years [1]. Though studied by many people for centuries, it is only in the past 100 years that research into mosquitoes has been intensified and become of great importance to human health. This explosion of work was brought about by "the discovery that these flying visits could be the prelude to sickness and death; the realisation that this lowly creature, this insignificant gnat could be responsible for whole epidemics, epidemics that in the past had helped shape the course of human history" [1]. The idea that mosquitoes were responsible for disease transmission has arisen through a number of early observations; the first reference is from a Brahmin priest called Susruta, in 500 BC India, who declared that malaria was spread by mosquitoes[1]. More recently, in 1572 Henry Hawkes wrote a report from Mexico in which he said that bites from mosquitoes inflicted newcomers to the area with sickness [1]. In 1848 Josiah Nott, an American doctor, suggested that both Yellow Fever and Malaria were spread by mosquitoes. In a world where the concept of disease was based on miasmas or bad air, and where germ theory had yet to be developed, these ideas were largely forgotten. In 1878 Sir Patrick Manson showed that the filarial worm *Wuchereria bancrofti* (the causative agent of elephantiasis) lived part of its life cycle in the mosquito, however he attributed the transmission of the disease to contaminated water from mosquito egg laying rather than direct transmission by the mosquito itself [1]. It was not until the discovery, and proof, by Ronald Ross in 1897 that mosquitoes were responsible for the transmission of malaria, a disease that has infected and killed people numbering hundreds of millions, that the study of mosquitoes changed forever. In parallel, in 1881 a series of papers by Carlos Finlay on the spread of yellow fever by *Aedes aegypti*, and in 1900 a series of experiments by Sir Walter Reed proved that the transmission of yellow fever was carried out by mosquitoes alone[1]. Since the turn of the last century, mosquitoes have been

shown to transmit more than 100 viruses, and dozens of parasitic infections, making them among the most important vectors of disease known to man (see Table 1).

Table 1. Examples of the human disease burden of mosquito species across the world

Disease	Infectious agent	Mosquito vectors
Parasites		
Malaria	Protozoa;	*Anopheles spp*
	Plasmodium falciparum	
	Plasmodium vivax	*Anopheles spp*
	Plasmodium ovale	*Anopheles spp*
	Plasmodium malariae	*Anopheles spp*
Elephantiasis	Helminth;	*Anopheles, Culex* and *Aedes spp.*
	Wuchereria bancrofti	
Viruses		
Eastern Equine Encephalitis (EEEV)	Virus	*Culiseta melanura Culex Melanoconion*
Venuzuelan Equine Encephalitis (VEEV)	Virus	*Culex Melanoconion spp*
Western Equine Encephalitis (WEEV)	Virus	*Culex tarsalis, Culex quinquefasciatus*
Chickungunya (CHICKV)	Virus	*Aedes spp*
O'nyong-nyong (ONNV)	Virus	*Anopheles gambiae/funestus*
Ross River (RRV)	Virus	*Culex annulirostris Oculerotatis vigilax,*
Dengue Fever 1-4 (DENV1-4)	Virus	*Aedes aegypti, Aedes albopictus*
Japanese Encephalitis (JEV)	Virus	*Culex tritaeniorhynchus, Culex spp*
St Louis Encephalitis (StLEV)	Virus	*Culex quinquefasciatus*
West Nile (WNV)	Virus	*Culex spp*
Yellow Fever (YFV)	Virus	*Aedes, Sabethes and Haemagogus spp.*

1.2 Arboviral disease

Viruses comprise the majority of infectious agents transmitted by mosquitoes. Arthropod-borne viruses (arboviruses) are defined as those multiplying within arthropod tissues after ingestion of infected vertebrate blood and transmitted by subsequent bites to other susceptible vertebrates [2]. Arboviruses mainly impact upon the developing world causing disease in humans, animals and livestock. They represent an emerging and resurgent group of pathogens, increasing in both prevalence and geographical distribution [3]. Of those, many are mosquito-borne, including Chikungunya virus (CHIKV), Dengue virus (DENV), West Nile virus (WNV) and Yellow fever virus (YFV). There are two main mosquito transmitted groups of arboviruses that are important for human health; the Flaviviruses and the Alphaviruses.

1.2.1 Flaviviruses

The flavivirus genus contains over 70 viruses, including DENV, Japanese encephalitis (JEV), WNV and YFV [4]. The name flavivirus is derived from the word yellow, referring to the jaundice caused by YFV [5]. The flaviviruses have a positive sense single stranded (+ss) ribonucleic acid (RNA) genome, encoding 10 genes; 3 structural proteins (capsid, envelope and membrane) and 7 non-structural proteins essential for genome replication [5].

YFV is widespread across Africa, and since the introduction of the disease believed to have happen during the slave trade, in South America. Sylvatic (jungle) cycles of transmission from primates are the sources of current epidemics, with multiple mosquito vectors being involved. Urban cycles of YFV are transmitted by *Ae. aegypti* from person to person. *Ae. aegypti* control programmes, as proposed by Finlay and Reed were successful in the eradication of urban cycles of YFV in Cuba, in Panama allowing the completion of the panama canal, in Rio de Janerio and subsequently in much of South America [1, 4].

Although an effective vaccine is available, developed in the 1930's by Max Theiler[1], only 32 of 44 countries where YFV is endemic use the vaccine, and in many of these countries vaccine coverage is less than 50% [4]. Currently no effective anti-viral treatment exists, and although the majority of cases are mild, fatality rates from severe YFV are around 50% [4]. The WHO data indicates that 15-20% of reported cases of YFV are fatal. The presence of capable mosquito vectors in much of Asia raises concerns that epidemics may arise from imported cases of the disease.

Dengue Fever, caused by a group of 4 DENV flaviviruses, is among the most important emerging infectious diseases: estimates range from 50-100 million [6] to 70-500 million [7] cases each year with more than 20,000 deaths, mainly in children [6-7]. Currently, DENV can be found in almost all tropical and sub-tropical regions, with regular epidemics in South-East Asia, the Caribbean and South America; no cases have been reported in temperate European countries in recent years [8]. DENV is predominantly transmitted by *Ae. aegypti*, however, *Ae. albopictus* is believed to be responsible for outbreaks in Japan, Indonesia, the Seychelles, Thailand, Malaysia, Indian Ocean Islands, Hawaii and China [9]. It is predicted that indigenous DENV transmission in Europe should be expected soon due to the presence of competent vectors and the frequent import of the disease [8]. To date no vaccine exists for DENV. One of the complicating factors is the presence of 4 distinct serotypes within the DENV group. Severity of disease often increases upon secondary infection with another DENV serotype, believed to be caused by antibody-dependent enhancement; antibodies bind to, but are unable to neutralise the virus, forming non-neutralizing antibody-virus complexes that bind to the Fc receptors of circulating mononuclear cells. This facilitates virus uptake and results in higher viral loads than in primary infection [4]. Thus a vaccine must provide long term immunity to all 4 DENV serotypes to be effective. Currently 5 DENV vaccines are in clinical or pre-clinical trials [7], and there are hopes that one of these will be widely availably by the mid-2010's [4].

A number of other flaviviruses impact upon human health, mainly from the Japanese encephalitis serogroup. Viruses from this serogroup share clinical and ecological features such as maintenance in enzootic cycles, usually between birds and *Culex* mosquito species [4]. Humans are 'dead end' hosts of these viruses i.e. human-mosquito-human transmission is not observed, and infection often results in encephalitic disease [4, 10]. JEV is endemic to parts of Asia, where most of the 50,000 annually reported cases orginate [10]. WNV is endemic in much of Africa, Europe, the Middle East, Central Asia, and has expanded across North America within the last decade[4]. Other members of the group can be found throughout Australia (Kunjin Virus, Murray Valley encephalitis, Alfuy viru), Africa (Usutu virus, Koutango virus, Yaounde virus) and South America (Cacipacore virus, St Louis encephalitis)[10]. The Flaviviruses represent an important group of diverse human pathogens, with a health burden across much of the world.

1.2.2 Alphaviruses

The alphaviruses are a genus of 30 enveloped RNA viruses that cause disease in human and domestic animals. They are transmitted by mosquitoes or other haematophagous insects (Table 1)[11]. Similarly to the Flaviviruses, the Alphavirus genome consists of +ssRNA, encoding 4 non-structural proteins essential for genome replication, capsid and 2 envelope glycoproteins [11]. Viruses from this group with particular interest in terms of human health are CHIKV, Ross River virus (RRV), Sindbis virus (SINV), Eastern equine encephalitis (EEE), Venezuelan encephalitis (VEE) and O'nyong-nyong (ONNV).

CHIKV was first isolated in the 1950s in East Africa [12]. Sporadic outbreaks were reported during the 1960's-1980's throughout much of Asia and in countries of Southern and Central Africa [13]. The primary vector was thought to be *Ae. aegypti*, while *Ae. albopictus* (the Asian tiger mosquito) acted as a secondary vector [14]. The Asian tiger mosquito has been spreading in geographical range rapidly; first identified in America in 1983, it invaded 36 States and several South American countries within 25 years, and currently is spreading through parts of Africa and Europe [9]. A huge epidemic of CHIKV occurred in the Indian Ocean in 2005-2006,

with over 3 million cases being reported in India and surrounding islands [15]. In 2006, around 35% of the population in the French island La Reunion was infected [13]. Unusually during this epidemic, very few asymptomatic infections, and a relatively high mortality rate for CHIKV (1:1000) were observed, likely due to the fact that infected population had little or no immunity [4]. Fears emerged for outbreaks of CHIKV and other arboviral diseases in North America and Europe, as the Asian tiger mosquito (found in many parts of Europe) was identified as the major vector [13]. With increased tourism and the existence of local mosquito populations capable of transmitting arboviruses, we may expect epidemics with high morbidity in non-immune populations. One year after the epidemic of CHIKV in Asia, around 200 cases were reported in 2 small Italian towns on the Adriatic coast, demonstrating the first example of locally transmitted CHIKV by the Asian Tiger mosquito in Europe [16-17]. This mosquito was introduced in Italy in 1990 through the importation of used tyres and has subsequently become endemic and widespread [18]. The most important feature of the 2005-2006 epidemic was a single amino acid substitution in the envelope protein of CHICKV that allows efficient transmission by both *Ae. aegypti* and *Ae. Albopictus* [4]. It was this variant that was locally transmitted in Italy in 2007, and it is believed this strain could cause further outbreaks in Europe and the Americas where *Ae. albopictus* is becoming widespread [4].

1.2.2.1 ONNV epidemics – 1959 to the present day

ONNV (family *Togaviridae*, genus *Alphavirus*, species *O'nyong-nyong*) is an unusual and rather elusive arthropod-borne virus [19]. Literally translated as 'weakening of the joints' or 'joint breaker', the disease causes fever, acute joint pain, lymphadenitus and rash [20]. In 1959 ONNV caused a huge epidemic beginning in north-west Uganda, and spreading to Kenya, throughout Tanzania, Malawi and Mosambique [21] infecting over 2 million people between 1959 and 1962 in East Africa alone [22-23]. Interestingly the disease affected all age groups with the same ferocity, indicating that no natural immunity was present in the population from previous infections [1]. Initially the disease was believed to be CHIKV [22] (a closely related alphavirus) that was responsible for an epidemic in Tanzania 6 years

previously [1], however serological tests confirmed the outbreak to be a different virus, named ONNV (Gulu strain)[24]. The principal vectors of the epidemic were determined as either *Anopheles gambiae* or *Anopheles funestus* mosquitoes [25] making ONNV the first and only discovered virus that is naturally transmitted by *Anopheline* mosquitoes. Curiously the disease all but vanished for 37 years, re-emerging in Uganda in 1996 [26-28]. During this time ONNV was sporadically identified in a small number of people [26, 28-29]. No animal reservoir of the disease could be found, and isolation of ONNV has only ever been possible from humans or mosquitoes [30]. An epidemiology study of ONNV carried out in Kenya by Marshall *et al 1982* [31] declared that ONNV 'transmission had virtually or actually ceased'.

In 1996 ONNV suddenly re-appeared in Uganda. The SG650 strain of ONNV was isolated during this epidemic. Mosquito collections during the outbreak indicate *A. funestus* as the principal vector of the virus however only a small number of mosquitoes were found to be ONNV positive, and only 2 *A. gambiae* mosquitoes were caught during the study [26]. All modern research on ONNV derives from this 1996 outbreak strain. The virus was sequenced [27], and was found to be the same as Igbo ora virus, identified in a small number of people since the 1959 outbreak [28]. The virus is closely related to CHICKV, and is placed in the same seriological group as Semliki Forest virus (SFV) and RRV [27, 32].

1.2.2.2 Virion structure, replication and transmission of ONNV

The alphaviruses have been well characterised, primarily based on SINV, RRV and SFV, with high conservation across the Alphavirus family. Thus the virion structure, replication and transmission of ONNV are based on that of the Alphavirus family, although some unknown differences may occur. The virion consists of a single copy of the genome complexed with multiple copies of capsid protein, forming the nucleocapsid core [11]. The nucleocapsid core is surrounded by a host derived lipid bilayer embedded with multiple dimers of two glycoproteins; E1 and E2. E1 and E2

dimers are arranged in trimers, each forming a spike on the surface of the virion [11]. The replication cycle of the alphaviruses in vertebrate cells has been extensively studied and is reviewed by Strauss and Strauss [11]. The general replication cycle begins with entry to a cell, either via receptor mediated endocytosis, or direct fusion of the virus with the plasma membrane [33]. Virus particles in endocytic vesicles are believed to fuse with the endosomal membrane after exposure to the acidic pH of the vesicle. Once free in the cell cytoplasm, viral replication begins [11]. The +ss viral RNA encodes 4 non-structural proteins (called nsP1-4) and 5 structural proteins (capsid, E1 and E2, E3 and 6K). The genome is organised as a single open reading frame (ORF) that generates one of two possible polypeptides (see figure 1) in most alphaviruses. This ORF comprises the 5' two thirds of the genome. P123 is believed to function as a proteinase that acts *in trans* (acting on other proteins) to process polypeptides involved in the replication process. P1234 is formed by read-through of an opal stop codon used to produce P123, and forms the active replicase. P1234 is cleaved to form nsP1-4, as well as a series of intermediates all with distinct important functions in replication [11]. ONNV does not have an opal stop codon between nsP3 and nsP4, and produces only a single P1234 polypeptide [27]. A minus strand copy of the genome is produced, and is transcribed into the 26S subgenomic RNA, in which the structural ORF resides. P130 polypeptide is translated and is cleaved *in cis* (cleaved by itself) to produce capsid protein, a precursor of E2 later cleaved into E2 and E3, the small 6K protein and E1 (see figure 1). In the early stages of infection, the non-structural genes are translated, and genome replication begins. This process occurs in the cytoplasm of the infected cell, although replication complexes are associated with internal membrane structures within the cell [11].

Each non-structural protein has a distinct function, along with their polypeptide intermediates[34]. nsP1 is required for the initiation or continuation of minus strand RNA synthesis. It is also believed to be responsible for capping genomic and subgenomic RNAs during transcription, having been shown to have both methyltransferase and guanyltransferase activity [11]. The nsP2 gene is believed to be an RNA helicase, belonging to the same superfamily as the DEAD box helicases, with a role in RNA unwinding during replication and transcription [11]. nsP2 is also

thought to be the proteinase responsible for cleaving the non-structural polypeptides. During SFV infection, around half the nsP2 produced is localised to the nucleus, where it's function is unknown [11]. The nsP3 gene has a poorly understood function in viral replication [35]. It has two

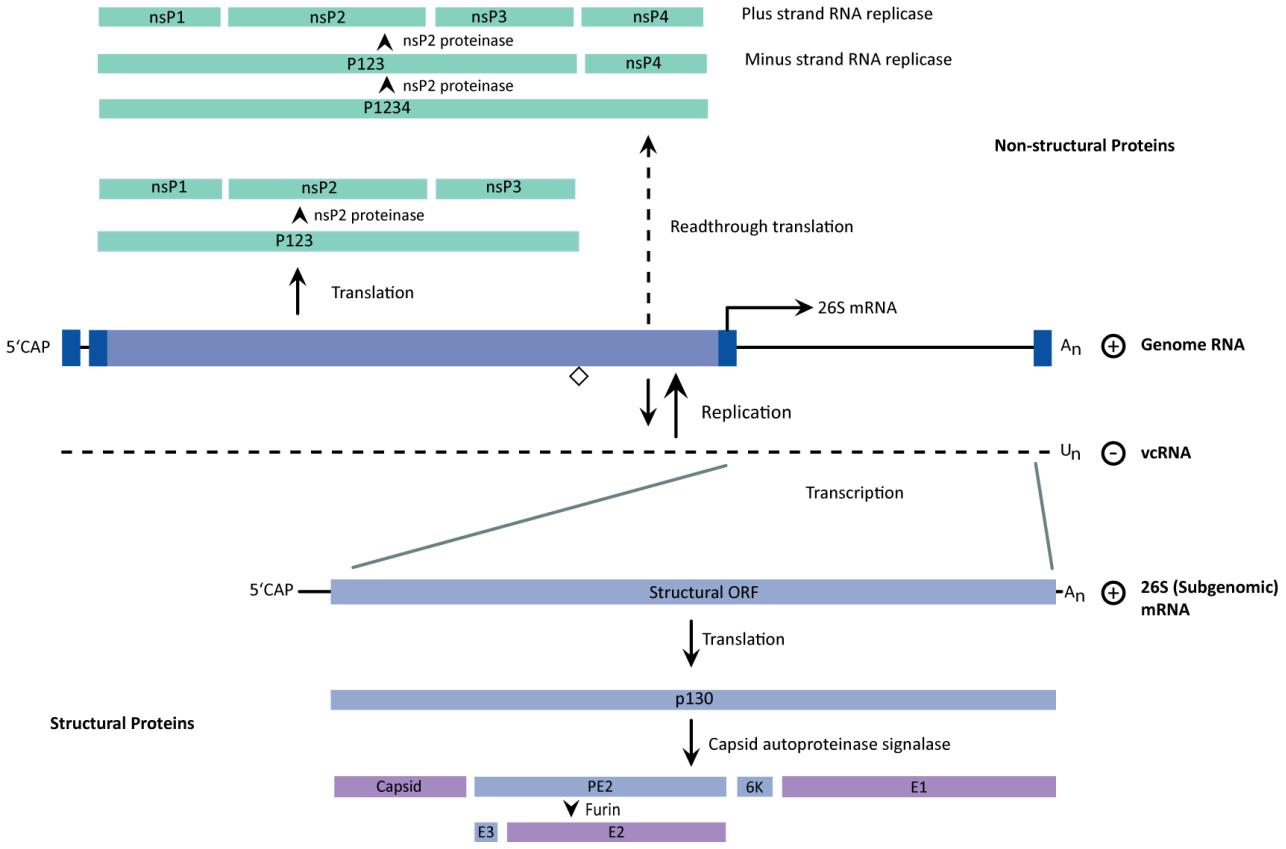


Figure 1. Genome organisation of the Alphaviruses – adapted from Strauss and Strauss [11]. The 49S genomic RNA is illustrated schematically in the centre, with its translated ORF shown as a medium blue box. Small dark blue boxes are conserved sequence elements; the open diamond denotes the leaky opal termination codon. The non-structural polyproteins and their processed products are shown above. Termination at the opal codon produces P123, whose major function in replication is believed to be as a proteinase that acts *in trans* to process the polyproteins in active RNA replicases; this proteinase domain is found in the nsP2 region. Readthrough of the opal stop codon produces P1234, which can form the active replicase. The 26S subgenomic mRNA is expanded below to show the structural ORF (light blue box) and its translation products. Structural polypeptides present in the virion are shaded purple. Viral complementary (vc)RNA is the minus strand compliment of genomic RNA [11].

domains; a highly conserved N-terminal macro domain with unknown function, and a highly variable and heavily phosphorylated C-terminal domain also with an unknown function [35]. nsP4 is the RNA-dependent RNA polymerase (RDRP) of the virus and is required for replication [11]. In vertebrate cells, early in infection, positive strand RNA (destined to be new genome copies) and negative strand RNA (the template for generation of the 26S subgenomic RNA and thus the structural genes) are generated, however minus strand synthesis ceases after a short time, and only positive strand RNA is generated throughout the rest of the infection. It has been suggested that the switch from minus to positive strand synthesis may be regulated by the cleavage of polyproteins, with P123 + nsP4 generating negative strand RNA, and nsP1, nsP2, nsP3 and nsP4 generating positive strand RNA with both replication complexes requiring host factors [11].

The structural proteins are transported via the golgi to the plasma membrane [36]. New viral cores assemble in the cytoplasm which then either diffuse to the plasma membrane, associate with the structural proteins, and bud from the membrane [11], or viral cores may fuse with secretory vesicles, forming new virus particles which are released when the secretory vesicle fuses with the plasma membrane [37]. Figure 2 depicts the replication cycle of the Alphaviruses.

Infection of insect cells differs in several ways to infection of mammalian cells. Most notably, several alphaviruses induce apoptosis of infected mammalian cells; however, this is not the case in insect cells [38]. It has been suggested that entry of the virus particles into the insect cells may occur via direct fusion with the plasma membrane rather than receptor mediated endocytosis [33]. Several studies on alphavirus infection of insect cells suggest that release of virus particles from cells may also differ; nucleocapsid cores in the cytoplasm are thought to form virus particles in internal membrane bound vesicles, subsequently being release during exocytosis [37, 39-40].

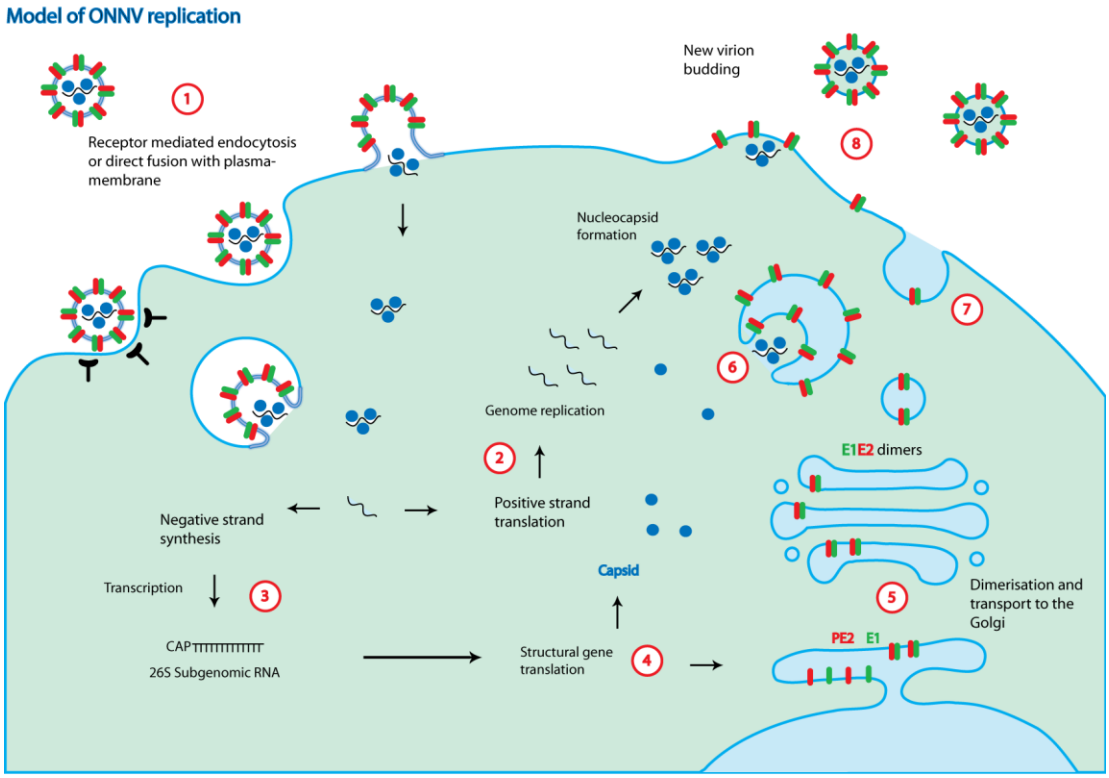


Figure 2. Model of Alphavirus infection and replication in insect cells. **1)** Mature infectious virions attach to the surface of the cell and enter either by receptor mediated endocytosis (clatherin-dependent or clatherin-independent) or by direct fusion with the plasma membrane. After receptor mediated endocytosis, the acidic environment of the endosome induces a conformational change in viral E1 and E2 surface glycoproteins resulting in fusion with the endosomal membrane. The nucleocapsid core is released into the cytoplasm and capsid protein dissociates from the RNA genome. **2)** Translation of the positive strand RNA genome produces the viral genes essential for genome replication. **3)** Transcription (generating the negative strand RNA genome) and **(4)** subsequent translation produces the subgenomic RNA from which the structural genes are translated. **5)** Viral proteins E1 and PE2 are modified in the ER. E1 and E2 form hetero-dimers and are transported via the secretory pathway to the surface of the cell **(7)**. Nucleocapsid cores form in the cytoplasm of the cell. Interactions between the E1/E2 hetero-dimers and capsid protein in the nucleocapsid core result in the budding of new mature virions **(8)**. Alternatively assembled nucleocapsid cores may bud into vesicles containing E1/E2 heterdimers, forming mature virions that are released from the cell when these vesicles fuse with plasma membrane **(6).**

1.2.2.3 Alphaviruses in the laboratory: a model group for molecular biology studies

Several alphaviruses have been made into Alphavirus Transducing systems (ATSs). ATSs use a complementary deoxyribonucleic acid (cDNA) copy of the viral genome to generate a plasmid containing the full length cDNA copy of the viral genome flanked with a promoter sequence, such as T7 or SP6. The first nucleotide of the viral sequence is modified to allow for the addition of a cap analogue [19]. The advantage of using a cDNA copy of the viral genome for genetic manipulation and amplification is the reduction in error rate gained from using a DNA polymerase compared to an RDRP. The plasmid is lineraised using a unique restriction enzyme site, and is subject to *in vitro* transcription. The resulting infectious clones of the virus are capped and then either electroporated or transfected into cultured cells (see figure 3)[19]. ATSs typically have very similar tissue tropism and replication rates as wild-type viruses [41-42]. They offer the advantage of being able to produce 'fresh' virus without serial passage of wild-type viruses, which results in quick adaptation to the conditions of culture and alterations in the original sequence of the virus. ATSs have been generated for SINV, CHICKV, ONNV, SFV, RRV and VEEV [19, 41-47]. The plasmid containing the cDNA infectious clone of the virus can be easily manipulated genetically. Two manipulations are common, one is the insertion of a gene of interest (GOI) and the second is the insertion of an RNAi cassette that can be used to silence GOIs. Both require the insertion of a multiple cloning site, usually under the control of a duplicated viral subgenomic promoter [19]. There are many examples of ATSs that have been used to introduce genes, including: a SINV ATS expressing scorpion toxin [43]; a SINV ATS expressing an anti-circumsporozoite protein antibody which successfully prevents *Plasmodium gallinaceum* sporozoites from invading the salivary glands of infected mosquitoes [44]; GFP expressing chimeric viruses used to study viral components essential for different vector/virus combinations [45]. Additionally ATSs have been used to silence GOIs, including a SINV ATS expressing dsRNA against the luciferase reporter gene [46], and a SINV ATS expressing dsRNA against DENV virus successfully preventing DENV infection in SINV infected mosquitoes [47]. Brault *et al 2004* [42] have generated several ATSs encoding

infectious clones of ONNV, including p5'dsONNVic-Foy, containing a full length cDNA clone of the SG650 strain of ONNV, a multiple cloning site and a duplicated viral subgenomic promoter. This plasmid has been used to generate infectious clones that express enhanced GFP (p5'dsONNVic-eGFP). Studies in this thesis have predominantly been carried out using this infectious clone of ONNV.

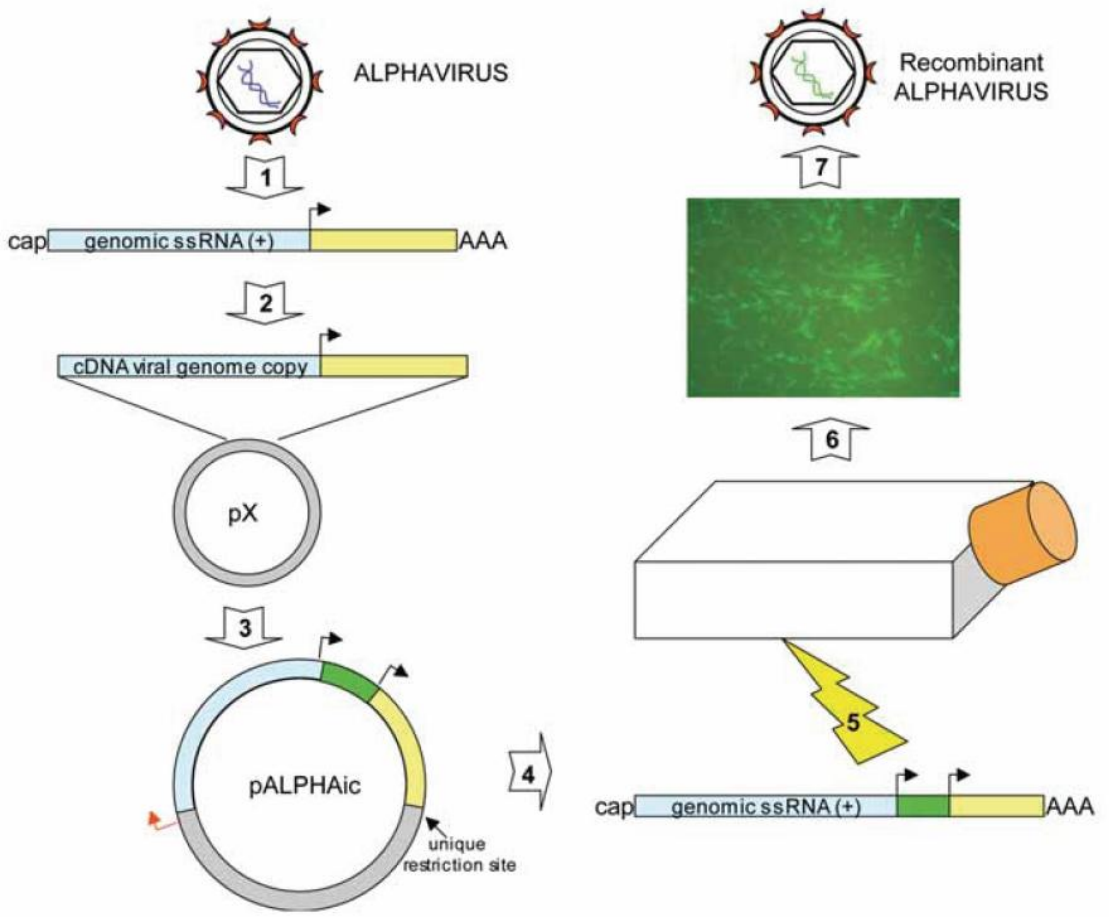


Figure 3. Generation of ATSs and production of infectious clones. Taken from Foy et al 2006 [19]. The full sequence of an Alphavirus is used to generate a full length cDNA copy of the viral genome (non-structural ORF is represented by the blue box, the structural ORF is represented by the yellow box). The full length cDNA copy is cloned into a plasmid (pX) and a unique restriction site is introduced. Additional markers, genes of interest or RNAi probes can be cloned into the viral genome (green box). Plasmid is linearised using the unique restriction site. *In vitro* transcription is performed and the resulting RNA is capped. Capped RNA is either electroporated or transfected into cells where the RNA clone replicates, eventually releasing recombinant infectious clones of the Alphavirus, which will go on to infect and replicate in other cells [19].

1.3 Mosquito innate immunity

Insects are exposed to a wide variety of pathogens both from within their natural habitats, and through their life cycles. In order to combat a barrage of microbial attack, insects have evolved a multifaceted innate immune response including physical barriers to infection (the thick and strong chitin cuticle) and a number of cellular and humoral innate immune effector mechanisms [48]. Invading pathogens are subject to attack from potent anti-microbial peptides (AMPs), the majority of which function by attacking the membrane surface of the pathogen causing lysis. Pathogens are recognised by pattern recognition receptors (PRRs) and are subsequently lysed by complement-like factors or phagocytosed by hemocytes - the insect equivalent of white blood cells. Additionally pathogens can be melanised (captured within a network of cross-linked proteins) or undergo cellular encapsulation by hemocytes. Recognition of pathogens also triggers immune signalling cascades, resulting in the up-regulation of immune genes, including anti-microbial peptides. The immune response of insects can be both systemic and/or localised [48].

1.3.1 Vector immune responses to Arboviruses

Insect immunity is a research field largely developed in the last two decades providing advanced understanding of anti-bacterial, anti-fungal and, more recently, anti-parasitic immune responses [49]. Considering the immense health and economic burden caused by viruses transmitted by insect vectors [3], it is perhaps surprising that anti-viral immunity has not received similar attention. The majority of studies on anti-viral immunity in insects have been conducted in the model insect *Drosophila melanogaster.* To date three immunity pathways have been implicated in *D. melanogaster* anti-viral responses (reviewed in [50-53]). The Toll pathway, known for its involvement in responses to Gram-positive bacteria and fungi [54], is activated by viral infection, and *Toll* or *Dif* loss-of-function mutants exhibit increased susceptibility to the infection [55]. It has been proposed that the function of the Toll pathway during a viral infection is to signal to the fat body and hemocytes, thus

enhancing the innate immune response. In response to a viral infection, specific genes are also up-regulated via the JAK/STAT (Janus kinase/signal transducers and activation of transcription) pathway, such as vir-1 (viral induced RNA 1)[56]. Loss-of-function of JAK/STAT signalling leads to increased susceptibility to infection. Finally, the RNA interference (RNAi) pathway has been shown to modulate the replication of RNA viruses in *D. melanogaster.* Loss-of-function of several components of the RNAi pathway including Dicer-2, Argonaute-2 (Ago2) and Armitage increases the susceptibility of flies to arboviruses [57-61].

The power of genetics and the generated extensive knowledgebase in *Drosophila* were invaluable in establishing the foundations for insect anti-viral immunity research. However, the biology of arboviruses is tightly linked to the physiology of haematophagous arthropods, and as such research in model organisms may not be fully relevant to the transmission of viruses and the associated vector defence. The immune responses of mosquitoes to malaria parasites can serve as a very good example. Several *A. gambiae* genes have been identified to date, which confer resistance to the rodent malaria parasite, *P. berghei*. However, the majority of these genes have no orthologs in *Drosophila*, e.g. Leucine rich repeat immune protein 1 (LRIM1) and APL1C [62-64], the thioester-containing protein 1 (TEP1) and numerous clip-domain serine proteases (CLIPs) and their inactive homologs, serpins (SRPNs) and C-type lectins (CTLs), which are implicated in parasite melanisation [65-67] Therefore, a forward research approach is required to effectively study the vector responses to arboviruses, utilising findings in *Drosophila* as guidance. Such early studies are reviewed below, which ought to be intensified in the future.

1.3.2 Responses of the JAK/STAT pathway

The JAK/STAT pathway was first identified in *Drosophila* via its role in embryo segmentation [68]. It begins with the circulating ligand Unpaired (Upd) that binds to the transmembrane receptor Domeless (Dome). Upd-bound Dome forms dimers activating the tyrosine kinase Hopskotch (or JAK), which in turn phosphorylates the transcription factor STAT92E leading to its nuclear translocation and up-regulation of

STAT-regulated genes [69]. In mammals, the homologous signal transduction cascade regulates the expression of over 30 cytokines and growth factor signals important during viral infection [70].

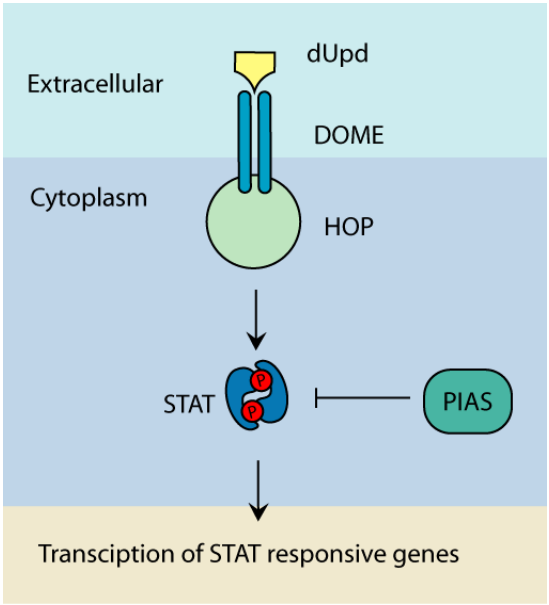


Figure 4. The JAK/STAT pathway in *A. gambiae*. The pathway begins with a ligand (currently not identified in *A. gambiae*, dUpd refers ligand of the *D. melanogaster* JAK/STAT pathway) that binds to the receptor DOME. Bound DOME activates the kinase Hopskotch (HOP), which in turn phosphorylates and activates the transcription factor STAT. Phosphorylated STAT translocates to the nucleus and switches on expression of STAT responsive genes. PIAS is a negative regulator of the JAK/STAT pathway.

In mosquitoes, orthologues of Dome, Hopskotch, and STAT92E have been identified. The ligand of the pathway, UPD, has yet to be discovered, likely due to the variable nature of receptor ligands, their small size, and the little sequence level homology. Figure 4 outlines the JAK/STAT pathway in *A. gambiae*. Recent evidence suggests that the JAK/STAT may have a conserved function in mosquitoes.

In a study characterising two mosquito STATs it was observed that infection of an *Ae. aegypti* C6/36 cell line with JEV led to decreased STAT activity in nucleus extracts. Experiments using a phosphatase inhibitor suggested that in the JEV-infected cells,

STAT phosphorylation is inhibited, probably through the induction of cellular phosphatase(s) or inactivation of JAK or other tyrosine kinase(s) by viral products [71]. The inhibition of the JAK/STAT signalling suggests that this pathway exerts enough evolutionary pressure on the virus to result in a possible inhibitory mechanism during infection.

A study by Xi *et al 2008* [72] investigated the genome-wide transcriptional responses of *Ae. aegypti* midguts and carcasses 10 days after natural DENV infection. It showed that a large proportion of the differentially regulated genes (~1/3) were known or putative immune genes belonging to multiple immune pathways and effector mechanisms, with a bias towards genes involved in the Toll pathway and the JAK/STAT pathways. In the carcass of DENV infected mosquitoes, DOME was up-regulated as well as an ortholog of the *D. melanogaster* SOCS36E. The latter cytokine is known to be a downstream target of the JAK/STAT pathway in *D. melanogaster* and functions to suppress further activation of the pathway [73]. A hypothetical protein, with a JAK/STAT related ortholog in *D. melanogaster* was down-regulated in both infected midgut and carcass tissues. Additionally a metabotropic glutamate receptor was up-regulated in infected carcasses; an ortholog of *D. melanogaster* mXr that was shown to have an inhibitory role in the JAK/STAT pathway in a genome wide RNAi screen [74]. The increase in transcripts of DOME and a gene putatively downstream of the JAK/STAT pathway indicate that viral infection triggers JAK/STAT signalling in the carcass of the mosquito. The increase in transcripts of two negative regulators of the pathway indicates that by 10dpi, negative feedback of the JAK/STAT pathway is in effect. Souza-Neto *et al 2009* [75] subsequently showed that *Ae. aegypti* can be made more, or less susceptible to DENV infection by silencing a positive (DOME) and negative (PIAS) regulator of the JAK/STAT pathway, thus demonstrating that the JAK/STAT pathway regulates DENV infection in *Ae. aegypti* mosquitoes. In order to dissect the downstream targets of the JAK/STAT pathway that have anti-viral effects, the authors carried out microarray analysis of PIAS depleted mosquitoes, and compared the differentially regulated genes to those regulated by DENV infection in *Ae. aegypti* mosquitoes. They found 18 genes that

were transcriptionally regulated in the same direction by DENV infection, and JAK/STAT activation. A selection of these genes were silenced in *Ae. aegypti* mosquitoes to see if they had an anti-viral function, of which two genes tested so far are demonstrated to be anti-DENV factors [75]. One of these genes has a transmembrane domain, possibly suggesting a receptor function. The other gene shares homology with the *A. gambiae* cuticular proteins (CPF) 1 and 2. How these genes function to reduce viral infection in the mosquito is not known.

In a study of SFV in an *Ae. albopictus* derived cell line (U4.4), activation of the JAK/STAT pathway, as well as the Toll and IMD pathways, was investigated [76]. It was shown that viral infection of cells did not induce any STAT dependent immune signalling. However putative activation of STAT (and IMD) signalling, by heat treated bacteria, prior to infection reduced viral gene expression from both the genomic and subgemonic promotors by around half (as demonstrated using recombinant viruses encoding luciferase), and also inhibited genome replication by 40-60% (as demonstrated by real time quantitative PCR (qrt-PCR)). It is not clear from this study which signalling pathway induced by treatment with heat-inactivated bacteria is responsible for the decrease in viral gene expression and replication. Numerous immune mechanisms have a role in targeting gram negative bacteria, including the IMD pathway regulated AMPs, TEP1 and CLT1 in conjuction with CTLMA2[49]. However, it is possible that STAT signalling is responsible for the reduction in viral replication. In summary, to date there is evidence that the JAK/STAT pathway regulates the expression of anti-viral immune effectors in *Ae. aegypti*, however no evidence in *A. gambiae* has been published.

1.3.3 RNAi pathway

The RNAi pathway is triggered by double stranded (ds) RNA molecules of various lengths. It begins with the cleavage of dsRNA, either by Dicer 1 (endogenous dsRNA) or by Dicer 2 (other dsRNA) (figure 5) complexed with R2D2[77] . Dicer 1 and Dicer 2 belong to the RNase III gene family, and contain an N-terminal helicase domain,

two RNase III like domains and a C-terminal RNA binding domain [59]. Processed dsRNA can take one of several forms – short interfereing (si) RNAs, repeat associated short RNA's (rasiRNAs) and micro RNA's (miRNAs). dsRNA generated by virus replication or by hybridisation of overlapping transcrips (e.g transposons) leads to the production of rasiRNAs and siRNAs, which generally target homologous mRNA for degradation [78]. Endogenous transcripts that contain homologous regions forming dsRNA hairpains lead to the production of miRNAs, which generally target homologous mRNA for translational silencing [78]. RNAi was quickly recognised as a mechanism to prevent or slow the replication of RNA viruses; however it also has an important role in the regulation of gene expression [78].

After processing, the small RNA fragments are incorporated into complexes (RNA induced silencing complex (RISC) or mi ribonucleoprotein (miRNP) complex). miRNAs are incorporated into miRNP and function in gene regulation whereas siRNAs are incorporated into RISC which targets transposon, and virus, replication. RISC contains several proteins; Ago2, which targets and destroys homologous RNAs; tudor-SN (TSN), which cleaves non-specific ssRNA, cleaves hyper-edited dsRNA substrates and has an undefined role in anti-viral defence; Vasa Intronic Gene (Vig), which binds RNA but has a poorly defined role in anti-viral defence and endogenous gene regulation [79] and other unidentified proteins [77]. Figure 5 summarises the model of the RNAi pathway in *D. melanogaster*.

The first evidence of RNAi mediated anti-viral mechanisms in mosquitoes came from Keene *et al 2005* [80]. In this study the role of RNAi in response to ONNV infection in *A. gambiae* mosquitoes was investigated. A recombinant virus that expresses GFP (ONNVic-eGFP) was used during investigations as a marker of infection. Initially it was demonstrated that RNAi was capable of modulating virus replication through dsRNA mediated silencing of the viral nsP3 gene, which efficiently reduced viral replication and titres in *A. gambiae* mosquitoes. Several components of the RNAi pathway, belonging to the Argonaute family, were silenced during infection by co-injection of dsRNA homologous to these genes and ONNVic-eGFP. AGO2 silenced

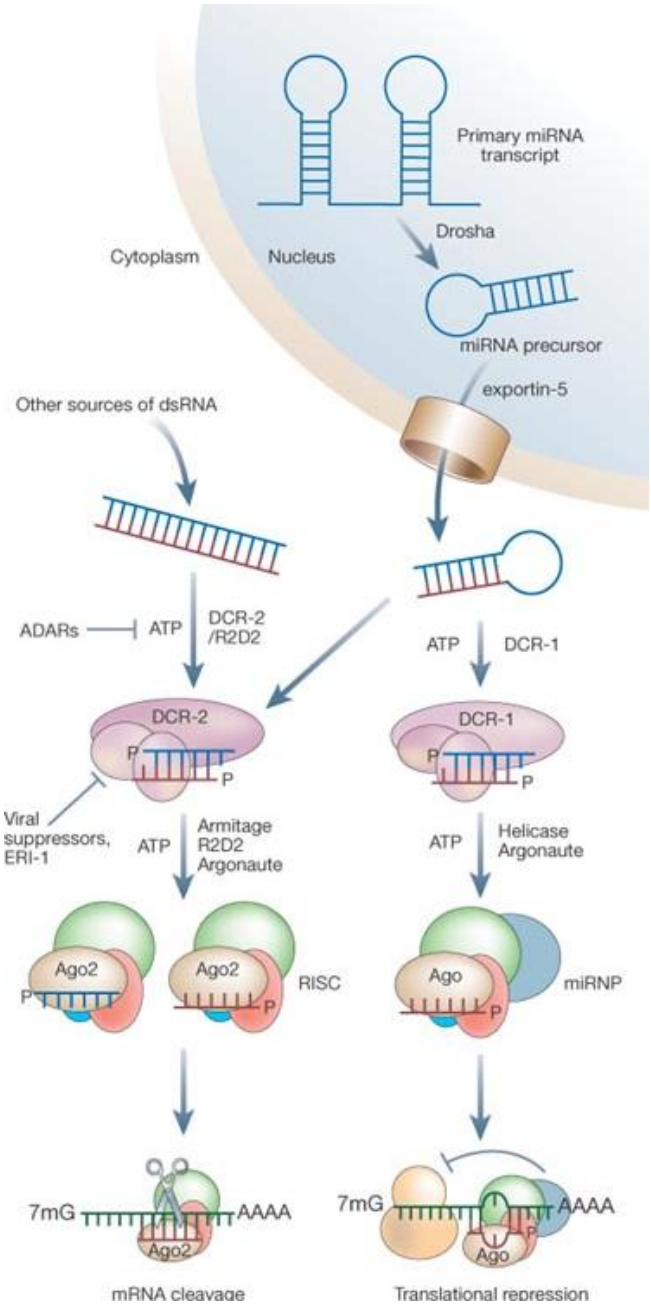


Figure 5. The RNAi pathway. Primary miRNA transcripts are processed to miRNA precursors in the nucleus by the RNase-III-like enzyme Drosha. The miRNA precursor is subsequently exported to the cytoplasm by means of the export receptor exportin-5. The miRNA precursor is further processed by Dicer to siRNA-duplex-like intermediates. The duplex is unwound while assembling into miRNP/RISC. Mature miRNAs bind to Ago proteins, which mediate translational repression or cleavage of target mRNAs. Other sources of long dsRNA in the cytoplasm of a cell are viral RNAs, artificially introduced dsRNA, dsRNAs generated by RdRPs, and genomic sense and antisense transcripts. Like miRNA precursors, long dsRNA is processed by the RNase III enzyme Dicer into 21–23 nucleotide dsRNA intermediates. Assisted by the RNA helicase Armitage and R2D2, the single-stranded siRNA-containing RISC is formed. The stability of the dsRNA and its recognition by Dicer can be regulated by specific ADARs and the exonuclease ERI-1. Taken from Meister and Tuschl [78].

mosquitoes showed significantly increased GFP expression and viral titres. AGO2 is a protein involved in the incorporation of dsRNA into the RISC complex during the RNAi pathway [81]. It is hypothesised that mosquitoes not expressing the AGO2 gene are unable to incorporate viral dsRNA into the RISC complex and are therefore unable to target viral dsRNA for degradation and reduce the rate of viral replication. In *A. gambiae* there are several other members in the Ago family – AGO1, AGO3, AGO4 and AGO5. Their role in RNAi was also investigated by Keene *et al* [80]. It was found that only AGO3 dsRNA treated mosquitoes demonstrated increased viral dissemination and viral titre. Campbell *et al 2008* [82] investigated the role of RNAi against two different recombinant forms of SINV in *Ae. aegypti* mosquitoes. They demonstrated that silencing of three proteins involved in the RNAi pathway, AGO2, DCR2 and to a lesser extent TSN, resulted in higher titres of a moderately infective SINV strain, indicating that AGO2 and DCR2 have clear roles in inhibiting viral replication in *Ae. aegypti* mosquitoes, although silencing of these genes did not have an impact on mosquito mortality after viral infection. Importantly Campbell *et al* [82] demonstrate the presence of small viral RNA fragments in infected mosquitoes indicating that RNAi is active specifically against viral RNA. Interestingly siRNA's could be seen using the moderately infective strain; however a more infective strain showed a reduced build up of siRNA's. This suggests that RNAi activity is limiting infection of one strain but not the other, pointing to a difference in RNAi susceptibility or the presence of an RNAi inhibitor in the more infective strain. They also observed more siRNAs generated from the positive sense RNA strand than the negative sense, indicating that secondary structure in the + RNA strand may be the target of RNAi as opposed to +/- dsRNA replication intermediates, although this is speculative. From investigations into effects of infectious blood meals on RNAi component transcript levels, Campbell *et al* [82] suggest that TSN may act as a sensor for the RNAi pathway.

RNAi mediated gene knock down of several key components of the RNAi pathway in *Ae. aegypti* also leads to increased DENV viral titres [83]. In particular Dicer-2, the enzyme that cleaves dsRNA, was shown to decrease the extrinsic incubation period

of the virus, and increase viral titres in the midgut, carcass and saliva of infected mosquitoes [83]. Mosquitoes infected with DENV showed staining with a dsRNA specific antibody in infected midguts, and virus specific 21-22nt length siRNAs were observed in infected carcasses indicating that viral infection leads to the production of dsRNA and the generation of siRNA's during RNAi [83].

Myles *et al* [84] carried out a study looking at the siRNA generated by Alphavirus infection and its role in the outcome of infection in several vector-virus combinations. Initially focusing on SINV infection of *Ae. aegypti*, and using libraries of siRNAs generated from SINV infected and uninfected mosquitoes they demonstrated a peak at 21 nucleotides of viral specific siRNAs (viRNAs). They also used a recombinant SINV containing the suppressor of RNAi B2 protein from Flock House Virus (FHV) (SIN-B2). This recombinant virus had very little accumulation of viRNAs. Mapping the viRNAs to the SINV genome demonstrated an asymmetric distribution across the genome, indicating 'hotspots' for viRNA generation, and also revealed that viRNAs were more often generated from the positive strand than the negative. In a series of experiments using SIN-B2 and a recombinant virus containing a mutant of B2 that is unable to bind to RNA (SIN-B2mut), the authors demonstrate that the presence of the functional suppressor of RNAi leads to a build up of genomic and subgenomic RNA compared to the mutant, and a reduction in viRNA accumulation compared to the mutant [84]. SIN-B2 was also shown to increase mortality in infected mosquitoes compared to the mutant. Similar results were observed by Cirimotich *et al* 2009 [85] where a recombinant SINV containing the FHV B2 protein caused higher mortality and increased viral titres in *Ae. aegypti* mosquitoes. Additionally, decreases in viRNA accumulation and increases in genomic and subgemonic RNAs were also observed using the recombinant virus by Cirimotich et al [85]. Myles *et al 2008* [84] also demonstrated that a recombinant ONNV containing the repressor of RNAi B2 protein from Nodamura Virus (NoV) causes increased mortality when intrathoracically inoculated in *A. gambiae* mosquitoes, which led to increased prevalence of infection after infectious bloodmeals. These data demonstrate the ability of RNAi in the absence of a

repressor to regulate viral infection loads, which is critical in maintaining the balance between pathogenicity and persistent infection. These data also demonstrate that introducing an inhibitor of the RNAi pathway and allowing viral infection to increase leads to mortality and morbidity that is not usually seen in Arbovirus infection of invertebrate hosts.

RNAi in plants is systemic i.e. the RNAi signal spreads from a site of infection to distal parts of the plant, and so can reduce or prevent viral infection in other parts of the plant [86]. Plants typically use RNA dependent RNA polmerases (RDRPs) to amplify the dsRNA signal [86]. Insects however lack the RDRP's necessary to do this, but there is some evidence of RNAi signals spreading in insects [87]. Limited evidence of localised spreading of the RNAi signal has been observed in an *Ae. albopictus* derived cell line [88]. In a series of experiments using the Renilla/Luciferase (RLuc) reporter system, siRNA's against RLuc and several modified SFV viruses expressing RLuc siRNA and GFP, the authors initially found that SFV cannot interfere with existing RISC complexes and cannot inhibit RNAi mosquito cells. They went on to show that by mixing cells transfected with a) the reporter gene and b) siRNA against the reporter gene, the siRNA 'signal' could be transferred to a neighbouring cell i.e. siRNA from one cell can silence the reporter gene in a neighbouring cell, provided that cells were plated at high density with many cell to cell contacts. They further demonstrated the spread of siRNA using fluorescin-labelled siRNA, and found that infecting cells with truncated viruses (capable of infecting cells, but unable to make progeny virus) and mixing them with cells infected with a virus containing a reporter gene, resulted in significant reduction in expression of the reporter gene, an observation that was reversed when a RNAi inhibitor was also expressed by the reporter virus [88]. Whether these effects are the result of genuine cell to cell communication, or are a by-product of the physical techniques used, such as scraping cells, which can cause significant damage to cell membranes, remains to be demonstrated convincingly.

Another virus shown to be effected by RNAi is YFV. Stable mammalian cell lines expressing siRNAs specific to the YFV have shown an inhibition of YFV infection

ranging from 11-97% depending on the region of the genome targeted by the siRNA [89]. Additionally the authors demonstrate that pre-treating mice with plasmids expressing these siRNAs prior to YFV infection reduces both morbidity of infection and mortality [89]. This is one of a small number of studies that support a role for RNAi in anti-viral defence in mammalian systems. Currently, however, it is though not to play a significant role, with the interferon response being the primary anti-viral response in mammals [84, 90].

These studies are the first steps to elucidating and understanding the RNAi anti-viral pathway in mosquitoes. As Keene *et al* [80] point out; different mosquitoes may vary in their RNAi response to viruses. *A. gambiae* is known to have a strong RNAi response to dsRNA, indicating that they may have a particularly robust RNAi response. Keene *et al* [80] suggest that this could explain why *A. gambiae* is a poor vector of arboviruses. Although it has been shown by Campbell *et al* [82] that *Ae. aegypti* does use RNAi against SINV infection, the effects of silencing RNAi components in *Aedes* on viral infection seems weaker; gene silencing was shown to be transient, with gene expression and viral titre returning to normal 7 days after dsRNA treatment. This differs from the similar study by Keene *et al*, in which it was demonstrated that viral titres were significantly higher in dsRNA treated mosquitoes six days after treatment. Silencing RNAi in *A. gambiae* also appeared to increase prevalence of viral infection to a greater extent than in *Ae. aegypti*. It has been noted that ONNV will readily infect *Ae. aegypti* mosquitoes, but conversely SINV cannot infect *A. gambiae* mosquitoes [80]. Some of these observations may simply be a result of the different vector-virus combinations (for example the presence or absence of a receptor required for infection or a virally encoded suppressor of RNAi). They do, however, support the idea that the vector competence of mosquitoes may be in part dictated by RNAi responses to viral infection. It has been shown from the several studies using repressor of RNAi proteins that the ability of the mosquito to generate viRNAs, and thus regulate the viral load of infection, is critical in maintaining a balance between low level infection and pathogenicity or mortality, and thus is critical in defining the vectorial capacity of the mosquito.

1.3.4 Toll pathway

The Toll pathway was first identified as being involved in the control of dorsoventral polarity in early embryogenesis of *D. melanogaster* [91]. In *Drosophila* it appears to have evolved a dual role, with one in nine Toll receptors being used in immune signalling rather than in developmental signalling [92]. In *Drosophila* the Toll immune signalling pathway begins with the recognition of a pathogen by a PRR. Gram negative binding protein 1 (GNBP1) and peptidoglycan recognition protein SA (PGRP-SA) are believed to be the PRR's involved in the Toll response to Gram-positive bacteria, possibly forming a complex that activates a proteolytic cascade resulting in the cleavage of a cytokine – Spatzle [93]. The cleaved form of Spatzle then binds directly to the Toll receptor and triggers an intracellular signalling cascade [94].

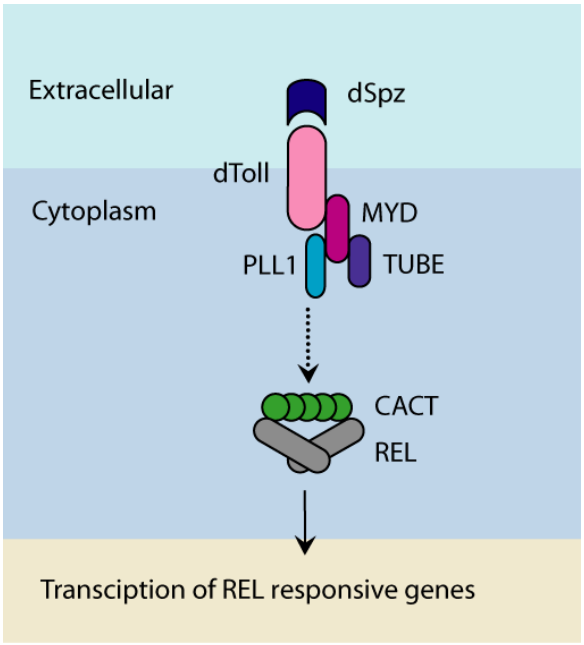


Figure 6. The Toll pathway in *A. gambiae*. Cleaved Spatzle binds to the transmembrane receptor Toll, resulting in the recruitment and activation of a cascade of signalling molecules, including MYD, TUBE, and PLL1. CACT, the inhibitor of REL, is degraded, freeing the NF-kB-like transcription factor REL1 to translocate to the nucleus and promote transcription of REL responsive genes. Genes with unidentified orthologs in *A. gambiae* are given the *D. melanogaster* gene name with the pre-fix d.

The direct binding of Spatzle alters the conformation of the Toll receptor resulting in intracellular recruitment of several proteins – MYD88, Tube and Pelle. Signalling downstream of these molecules results in the disassociation of the NF-κB like Dif from its I-κB like inhibitor, Cactus, which is degraded, possibly by Pelle [95]. Once Cactus has been degraded, Dif is free to translocate into the nucleus, where groups of genes, including immune genes such as Drosomycin, are up-regulated [96]. In *A. gambiae* the NF-κB protein in the Toll pathway is REL1. Figure 6 outlines the Toll pathway in *A. gambiae* mosquitoes.

In a study by Sanders *et al 2005* [97] DNA microarray analysis of 2170 *Aedes* genes at three time points was carried out during infection with SINV. It was observed that viral infection leads to the up- or down-regulation of several known and putative innate immune genes. These include the *Ae. aegypti* dif gene (Aedif, an NF-kB like protein involved in the signal transduction of the Toll pathway), which was up-regulated early in infection; four chitin binding proteins similar to plant and horseshoe crab AMPs; several serine proteases and zinc-dependent proteases, as well as notable changes in some signalling molecules (such as membrane receptors, aGTPase, kinases, phosphatases, and transcriptional regulators).

Sanders *et al* [97] speculate that the early up-regulation of Aedif, and a subsequent return to normal transcript levels at 4 and 8 days post infection (dpi), indicates that viral infection leads to activation of the Toll pathway, which is later inhibited. They speculate that the down-regulation of an ubiquitin ligase, Cullin-1, may lead to the inhibition of the Toll pathway via decreased degradation of the dif inhibitor, Cactus. Inhibition of unbiquitin ligase complexes and thus activation of the Aedif may be a mechanism by which SINV avoids Toll-mediated anti-viral responses; similar mechanisms have been observed in HIV-1 and Tobacco mosaic virus (TMV) infections [97].

As previously mentioned in respect to the JAK/STAT pathway, a microarray based study by Xi et al [72] of DENV infected *Ae. aegypti* mosquitoes showed a strong bias towards genes involved in the Toll and JAK/SAT pathway: Spatzle, Toll, Rel1A and Cactus were differentially regulated at 10 dpi. A comparison was performed between DENV induced gene expression and genes differentially regulated by two immune signalling pathways in *Ae. aegypti*. Cactus, a negative regulator of the Toll pathway, and Casper, a negative regulator of the IMD pathway, were knocked down (KD) using RNAi, and gene regulation was analysed using microarrays. A small overlap between DENV induced and Casper KD induced genes was observed (9%), however 41% of DENV induced genes overlapped with Cactus KD induced genes. Xi et al [72] went on to investigate the effect of activation or silencing of the Toll pathway on DENV infection in *Ae. aegypti* mosquitoes. Silencing of Cactus, the negative regulator of the Toll pathway, prior to DENV infection resulted in a 4-fold decrease of virus in mosquito midguts 7 dpi. Silencing of Myd88, a positive regulator of the Toll pathway, prior to infection resulted in a 2.7 fold increase in virus. Silencing of Caspar, a negative regulator of the IMD pathway, had no effect on viral infection. These results suggest that viral infection leads to the activation of the Toll pathway, but not the IMD pathway, and that the Toll pathway regulates downstream effectors that, through unknown mechanisms, suppress viral infection. Dome, the receptor involved in JAK/STAT signalling, was similarly regulated by DENV infection and Cactus KD. The authors suggest that the JAK/STAT pathway could be activated by the same signalling cascades that activate the Toll pathway, or that it could be activated indirectly by the Toll pathway itself, as evidence suggests in *D. Melanogaster* [72]. Xi et al also observed that removing the natural bacterial flora in mosquito midguts by antibiotic treatment increased viral infection. It is proposed that the basal activation of immune genes by bacteria leads to the up-regulation of anti-viral immune genes, possibly by Toll pathway activation as indicated by the up-regulation of several Toll regulated immune genes in septic compared to aseptic mosquitoes.

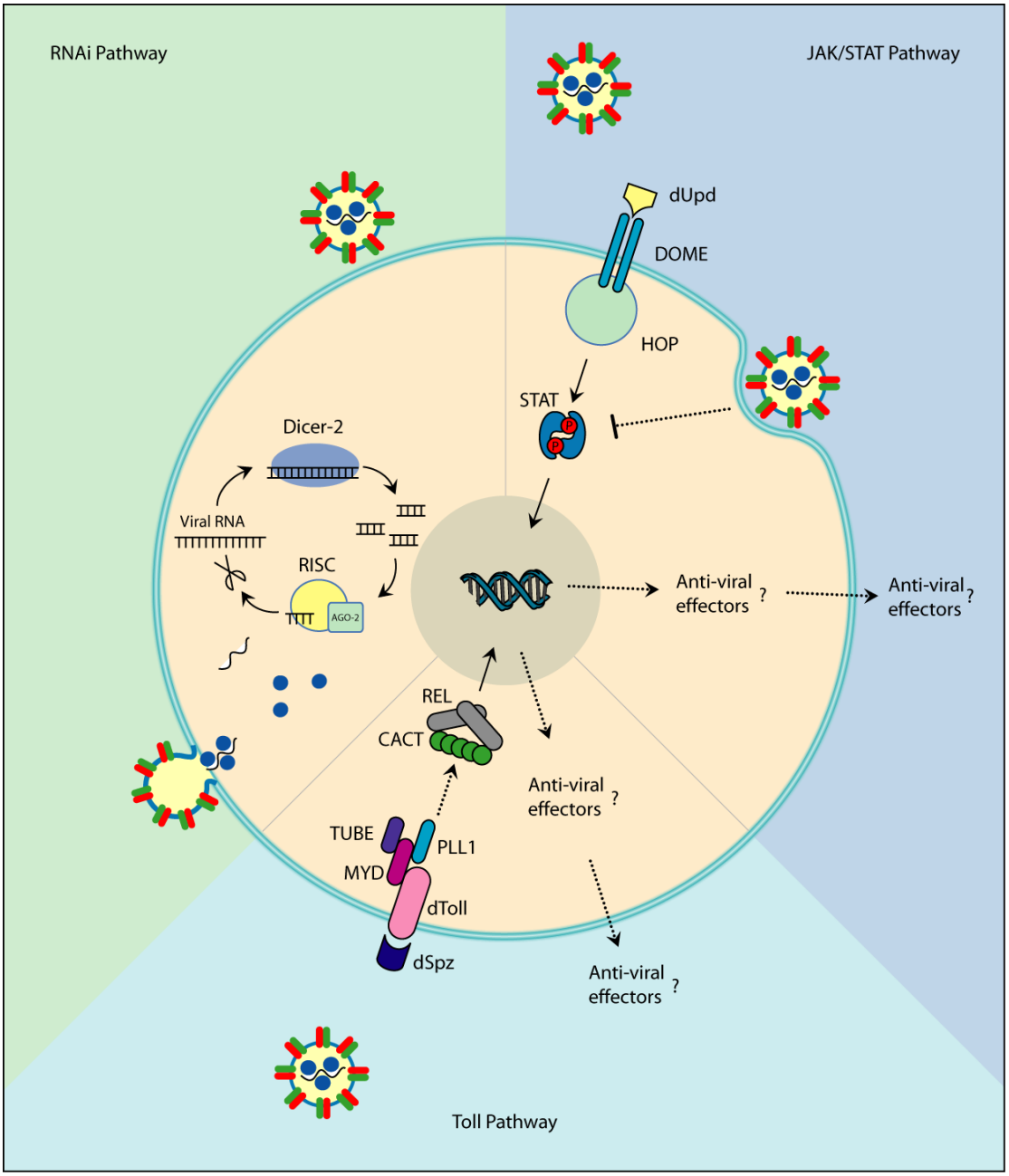


Figure 7. Three conserved pathways involved in anti-viral immunity in mosquitoes. The RNAi pathway recognises virally derived dsRNA formed during replication within host cells. dsRNA is processed by the enzyme Dicer-2 siRNA is loaded into the RISC complex, which targets complimentary RNA for destruction. The Toll pathway regulates the transcription of currently unknown anti-viral effectors. The JAK/STAT pathway regulated the transcription of unknown anti-viral effectors, and is inhiobited by JEV infection in *Ae. aegypti*. Dotted arrows represent unknown interactions. *D. melanogaster* genes have been named where *A. gambiae* orthologs have not been identified (dSpz, dToll, dUpd).

1.4 Summary

Over the last few years evidence has emerged that provides the first insights into invertebrate anti-viral immunology. Three major pathways have been implicated in anti-viral responses, the RNAi pathway, the JAK/STAT pathway and the Toll pathway (figure 7). These three pathways appear to have a conserved role in anti-viral immunity from flies to mosquitoes. Although we are beginning to understand which molecular pathways may be involved in the regulation of viral infection, there are many questions that remain to be answered, both in terms of the biology of the replication cycle of different viruses in mosquito vectors and the molecular mechanisms of anti-viral immunity. Anti-viral downstream effectors of the Toll and JAK/STAT pathways are currently not known and differences in RNAi susceptibility from mosquito to mosquito, and virus to virus is not well understood. With increases in the geographical distribution of vector mosquitoes and in international travel and transport, the probability increases that viral diseases will continue to affect new parts of the world. The elucidation of anti-viral mechanisms may provide new targets for combating these viral diseases.

2 Aims and objectives

The broad aim of this project was to explore the mechanisms that *A. gambiae* employs to cope with viral infections. The main objective was to identify genes involved in anti-viral immunity in *A. gambiae* mosquitoes through literature research into orthologous systems and transcriptional profiling of infected mosquitoes. A secondary objective of this project was to develop a cell based assay for rapid identification of genes involved in viral infection in *A. gambiae*, with a view to carrying out a genome wide RNAi screen.

The working hypothesis was that mosquitoes do posses anti-viral mechanisms capable of limiting or modulating viral infection. Based on observations from the model system *D. melanogaster* and limited evidence in *A. gambiae* outlined in the introduction, several different pathways are thought to play a role in response to viral infection, namely the JAK/STAT and RNAi pathways. The approach was to build upon the existing knowledge of anti-viral immunity mechanisms in invertebrates and couple this with a forward genetics methodology to study the mosquito responses to viral infection. This study sought to increase our understanding of the interactions between the unique Anopheles-ONNV vector-virus combination and provide insights into the *A. gambiae* immune response to viral infection as a critical factor in the determination of its capacity to transmit viral disease.

3 Materials and methods

3.1 Maintenance of mosquitoes

Adult mosquitoes were maintained as described in detail by Sinden and co-workers [98]. Three strains of *A. gambiae* were used in experiments; G3 (colonised from The Gambia in the 1950's), Yaounde (colonised from the Yaounde area, Cameroon, in 1988), N'gousso (recently colonised from wild mosquitoes collected in Cameroon). *A. arabiensis* and *A.quadrinalatus* SKUQUA strain (colonised from the Skukuza area, Kruger National Park, South Africa in December 2005) were also used. In brief, mosquitoes were reared and maintained at 28^{o}C, 65-70% relative humidity with a 12 hour light/dark cycle. Adult mosquitoes were fed on sterile filtered and autoclaved 10% fructose solution and used for experimental purposes when 1 or 2 days old.

3.2 Maintenance of mosquito derived cell lines

Seven *A. gambiae* derived cell lines were used during experiments. All cell lines were derived from minced neonate larvae within one hour of hatching [99]. Strains used to generate cell lines were Suakoko 2La (cell lines SuaE1, Sua5.1, Sua4.0), L35 (L35 cell line) and 4a r/r (cell lines 4a3A and 4a3B). Cell lines are primary cultures that have not been cloned; they contain a heterogeneous population of cells derived from whole minced neonates. Cells were maintained at 27^{o}C in 'complete insect media'; Schneiders medium (Gibco, UK) supplemented with 10% Fetal Calf Serum (FCS) (Sigma) heat inactivated at 56^{o}C for 1 hour and 0.01mg/ml penicillin/streptomycin (Invitrogen, UK). Cells were split by scraping cells or vigorously shaking the dish to loosen cells, and diluting 1:20 in fresh complete media every 4-7 days.

3.3 Maintenance of mammalian derived cell lines

African green monkey kidney cells (VERO) and Baby Hamster Kidney cells (BHK) were maintained at 37^{o}C in 'complete mammalian media': minimum essential media (MEM) (Invitrogen) supplemented with 10% FCS (Sigma) heat inactivated at 65^{0}C for

30 minutes, 0.01mg/ml Penicillin/Streptomycin (Invitrogen), 0.292mg/ml L-glutamine (Invitrogen), 0.01mM non-essential amino acids (aa's) and 0.25ug/ml fungizone (Invitrogen). Cells were split by trypsanisation with 0.5mg/ml Trypsin/0.2mg/ml Ethylenadiaminetetraacetic acid (EDTA) (Invitrogen), washing in phosphate buffered saline (PBS) and diluting 1:15 with complete mammalian media.

3.4 Production, propagation and titration of 5'ONNVic-eGFP

3.4.1 Production of 5'ONNVic-eGFP

5'ONNVic-eGFP plasmid was kindly provided by Dr Brian Foy, Colorado State University. The plasmid was transformed into DH5α *E.coli* (Invitrogen, UK) according to the manufacturers instructions. A single colony was picked and used to inoculate a 5ml culture of lysogeny broth (LB), and grown at 37^{o}C, shaking at 225 revolutions per minute (rpm) for 8-16 hours. 250µl of the 5ml culture were used to inoculate a 250ml culture of LB, grown at 37^{o}C, shaking at 225rpm for 8-16 hours. Plasmid DNA was extracted using the Qiagen Maxiprep Kit (Qiagen, UK) according to the manufacturers instructions.

Five micrograms of the plasmid was linearised using Not1 restriction enzyme (New England Biolabs) in a total volume of 100µl for 2 hours at 37^{o}C. 200mg/ml of Proteinase K were added to the reaction and incubated at 37^{o}C over night. The linearised plasmid was phenol-chloroform extracted using ribonuclease (RNAse) free plasticware and reagents; linearised plasmid was brought to the total volume of 450µl with water, an equal volume of cold saturated phenol/chloroform/iso-amyl alcohol (PCI). The sample was vortexed briefly, placed on ice for 5 minutes and centrifuged at 4^{o}C, 18000 g, for 3 minutes. The top aqueous layer was removed and retained. An equal volume of chloroform was added, the sample was vortexted briefly, placed on ice for 5 minutes and centrifuged at 4^{o}C, 18000 g, for 3 minutes. The top aqueous layer was removed and retained. The DNA was precipitated; 60µl

of 3M Sodium Acetate (pH5.2) was added followed by 1ml of ice-cold ethanol. DNA was precipitated at -80^{o}C for a minimum of 1 hour. DNA was pelleted by centrifugation at 4^{o}C, 18000 g for 15 minutes. Supernatant was removed and the pellet was air dried for around 10 minutes. The pellet was re-suspended in 25µl of water. The linearised plasmid was transcribed into RNA using the T7 megascript Kit (Ambion, UK); 1µg of plasmid was transcribed in a 50µl reaction containing 2.5µl of each deoxyribonucleotide triphosphate (dNTP), 5µl of 10X buffer, 2µl T7 RNA polymerase and 1mM A Cap analogue (m7G(5')ppp(5')A) (Ambion), at 39^{o}C for 1 hour.

Transcribed RNA was cleaned up using the RNeasy mini kit (Qiagen) according to the manufacturers instructions. RNA concentration and purity was ascertained using a Nanodrop. 2µg of RNA were transfected into a confluent culture of VERO cells in a T75 flask using the Transmessenger transfection reagent (Qiagen) according to the manufacturers instructions. At 24 hours post transfection cells were screened for GFP expression. If the transfection was successful, cells were monitored for GFP expression and cell death caused by viral infection. Cells were harvested at around 72-96 hours post transfection: cells were scraped, stored in 200µl aliquots at -80^{o}C and designated as Passage 1 Vero 5'ONNVic-eGFP (P1V 5'ONNVic-eGFP).

3.4.2 Propagation of 5'ONNVic-eGFP

For experimental use, a large volume of Passage 2 virus (P1/P2V 5'ONNVic-eGFP) was generated; one 200 µl aliquot of P1V 5'ONNVic-eGFP was used to inoculate a culture of confluent VERO cells in 20ml media in a T175 flask (Nunc). Cells were checked every 12 hours for GFP expression and at 72 hours post infection (hpi) cells were scraped, filtered through a 0.45µm filter and stored at -80^{o}C in 250µl aliquots. Alternatively P1V/P2L35 5'ONNVic-eGFP was generated as above, inoculating *A. gambiae* L35 cells with P1V 5'ONNVic-eGFP.

3.4.3 Standard plaque assay to determine viral titre

Plaque assay was used to determine the titre of P1V, P1/P2V and P1V/P2L35 5′ONNVic-eGFP stocks. VERO cells were plated in 24 well plates (Nunc) and allowed to reach 100% confluency. Samples to be titred were allowed to thaw on ice. 2% agar was autoclaved prior to use, melted on the day of use, and placed in a 40^0C waterbath to cool but not set. Samples to be titrated were serially diluted 10^1-10^8 in Dulbecco's Minimum Essential Media (DMEM) (Invitrogen). Media was removed from cells and 2 replicates of 150µl of each dilution were added to a single well of the 24 well plate. Plates were incubated for 1 hour at 37°C to allow for viral attachment to cells, plates were rocked every 10 minutes to ensure even distribution of samples. During this hour, equal volumes of molten 2% agar, and a 2X nutrient solution (per litre of nutrient solution, 862ml 1X Earles Basic Salts solution, 66ml yeast extract/Lactalbumen solution (sterile filtered), 40mls FCS, 30mls 7.5% sodium bicarbonate, 2mls 10mg/ml Gentamycin) were mixed and kept in a 40°C waterbath to prevent setting. After 1 hour incubation, 1ml of agar/nutrient solution was pipetted into each well of the 24 well plates, and allowed to cool for 1-2 hours at room temperature (RT) before returning to 37°C. After 4 days incubation at 37°C, 200µl of 5mg/ml Thiozolyl Blue Tetrazolium Bromide (MTT) (Sigma) in PBS was added to each well, and allowed to develop for at least 4 hours. The number of plaques (seen as white plaques of dead cells against the purple colour of live cells) in each well was counted, and the plaque forming units (PFU)/ml calculated using the well in which the highest number of individual plaques could be counted. Each sample was titred in duplicate and values were averaged.

3.5 Infection of mosquitoes and mosquito cells with 5′ONNVic-eGFP

3.5.1 Infection of *A. gambiae* cell lines with 5′ONNVic-eGFP

Confluent cultures of *A. gambiae* cells lines (Sua4.0, Sua5.1*, SuaE1, 4a2, 4a3A, 4a3B and L35) were grown in 96 well/24 well/12 well/6 well plates or T25 flasks. The

number of cells (counted using a haemocytometer) used to seed each well were 5 x 10^4; 5 x 10^5; 1 x 10^6; 2 x 10^6 or 6 x 10^6 respectively. An aliquot of previously titrated virus was thawed and diluted with complete insect media to an appropriate Multiplicity of Infection (MOI; referring to the number of infectious viral particles relative to the number of cells thus an MOI of 1 denotes 1 viral particle for every cell in a culture). Diluted virus was added directly to the cell cultures. Cultures were rocked gently for 30 minutes before being placed at 27^{o}C. Infected GFP expressing cells were viewed using a fluorescent microscope.

3.5.2 Intrathoracic inoculation of *A. gambiae* mosquitoes with ONNVic-eGFP

Newly emerged female mosquitoes (*A. gambiae* G3/Yaounde/N'gousso strains, *A. quadriannulatus* and *A. arabiensis*) were inoculated with the required dilution of P1/P2V or P1V/P2L35 5'ONNVic-eGFP in MEM (Invitrogen), using a pulled capillary glass needle and a Nanoject (Drummond Scientific). Inoculated mosquitoes were kept in cohorts of 50 and maintained as described by [98]. Incoculated mosquitoes were double-contained to prevent escape.

3.5.3 *Per Os* infection of *A. gambiae* mosquitoes with 5'ONNVic-eGFP

Cohorts of ~50 mosquitoes were starved of 10% sugar for 4-12 hours prior to blood feeding. Infectious blood meals were prepared by mixing the required volume of P1/P2V 5'ONNVic-eGFP with a 1:1 mixture of compacted human red blood cells (RBCs) and human AB serum (GemCell©). Compacted human RBCs were prepared by adding freshly drawn blood to 1/10th volume heparin sodium salt (1.66 mg/ml in PBS, sterile filtered) and centrifuging at 600g for 5 minutes in a 50ml Falcon tube. Compacted RBCs were pipetted when required from the bottom of the 50ml Falcon tube. RBCs were kept for up to two weeks at 4^{o}C. Infectious blood meals were delivered using membrane feeders; plastic membrane feeders containing a compartment for flow of warmed water (via connection to a water bath set at 38^{o}C)

and parafilm stretched across to form a second compartment at the bottom of the membrane feeder, into which the infectious blood meal is introduced using a 24 Gauge blunt needle. Mosquitoes were placed under the warmed blood meal and were allowed to feed to occlusion (30mins-1hour). Infectious titre per mosquito was calculated based on the assumption that each mosquito will ingest approximately 2µl of blood. Unfed mosquitoes were removed 24-48h post blood feeding, when the blood meal can be clearly seen through the cuticle of the mosquito.

3.6 Antibiotic treatment of *A. gambiae* mosquitoes

Newly emerged G3 and N'gousso strain mosquitoes were fed on 10% fructose solution containing Gentamycin, Penicillin/Streptomycin (Invitrogen) for 5 days. To confirm that anti-biotic treatment was successful, mosquitoes were homogenised in LB and plated onto LB agar plates and incubated at 27^{o}C overnight.

3.7 dsRNA preparation for gene knock down

Primers were designed (Table 2) for 200-600bp sections of genes of interest, with a T7 promotor sequence (GAATTAATACGACTCACTATAGGGAGA) added to their 5' ends. Polymerase chain reaction (PCR) was carried out using cDNA derived from *A. gambiae* mosquitoes and PCR products were sequenced to confirm correct amplification for each probe. PCR amplicons were used to synthesise dsRNA using the T7 MEGAscript kit (Ambion) according to the manufacturers instructions. Concentration of dsRNA was adjusted to 3µg/µl and stored at -80°C until use.

Table 2. Primers for RNAi probes. Primers were designed using the E-RNAi webservice from the German Cancer Research Centre (http://www.dkfz.de/signaling/e-rnai3/)

Gene	Primer ID	Gene Name	F primer	R primer
n/a	LacZ RNAi F/R	LacZ	GAATTAATACGACTCACTATAGGGAGAATCCGACGGGTTGTTACT	GAATTAATACGACTCACTATAGGG CACCACGCTCATCGATAATTT
	Hop(2) RNAi F/R	HOP	GAATTAATACGACTCACTATAGGGAGAGCCAAGGAGCTGGTTATCAA	GAATTAATACGACTCACTATAGGGAGAAGAGCAGATCGTGCTTGGTT
	AgAGO2 RNAi F/R	Ago2	GAATTAATACGACTCACTATAGGGAGAGCATGAGCACGCTCAACAAC	GAATTAATACGACTCACTATAGGGAGAGTTCGAGTCGTCGTACAGCA
n/a	nsp3 RNAi F/R	nsP3	GAATTAATACGACTCACTATAGGGAGAACCGGTGTGTACTCAGGAGG	GAATTAATACGACTCACTATAGGGAGACATAGGCACGGGACTGTTTT
AGAP002836	Dicer1 RNAi F/r	Dicer1	GAATTAATACGACTCACTATAGGGAGATGCTAAGCTTTGGCTGGAAT	GAATTAATACGACTCACTATAGGGAGATTTCGTTCGACCATGTACCA
AGAP006941	STAT1 RNAi F/R	STAT1	GAATTAATACGACTCACTATAGGGAGAGAAAATCAACCACCGCAAGT	GAATTAATACGACTCACTATAGGGAGACAGCTCTTCCTGTTCCAAGG
AGAP003508	STAT2 RNAi F/R	STAT2	GAATTAATACGACTCACTATAGGGAGACATGAACAACATCGGCAATC	GAATTAATACGACTCACTATAGGGAGATGCATATTCTCTGCCGTGAG
AGAP005031	PIAS RNAi F/R	PIAS	GAATTAATACGACTCACTATAGGGAGAAATCCAATTCCCACCAACAA	GAATTAATACGACTCACTATAGGGAGAGGACAGTTCCATGTTGGCTT
	REL1 RNAi F/R	REL1	GAATTAATACGACTCACTATTAGGGAGAATCAACAGCACGACGATGAG	GAATTAATACGACTCACTATTAGGGAGATCGAAAAAGCGCACCTTAATT
	CACT RNAi F/R	CACTUS	GAATTAATACGACTCACTATTAGGGAGAGTCCGCTCTACACATCAGCA	GAATTAATACGACTCACTATTAGGGAGACCGTTCGGGTTAATGATGAC
	REL2 RNAi F/R	REL2	GAATTAATACGACTCACTATTAGGGAGAAATCCGACGCAAAGATACG	GAATTAATACGACTCACTATTAGGGAGAGACCGCAATGTGAAGGATG
AGAP012352	ML1 RNAi F/R	ML1	GAATTAATACGACTCACTATAGGGAGAGAAATGTCCCGGTGAAGAGA	GAATTAATACGACTCACTATAGGGAGACCCACCAGCGTTGTTTTAGT
AGAP002848	ML9 RNAi F/R	ML9	GAATTAATACGACTCACTATAGGGAGAACGGTCCATAGCAAGGATTG	GAATTAATACGACTCACTATAGGGAGAGTCAGCGGACAGGAAGTGTT
AGAP012529	GALE8 RNAi F/R	GALE8	GAATTAATACGACTCACTATAGGGAGAGGTCTAGGCATTTACCGCAA	GAATTAATACGACTCACTATAGGGAGAGAGCCGTCCTTATTCTGTGG
AGAP007343	LYSC2 RNAi F/R	LYSC2	GAATTAATACGACTCACTATAGGGAGAAAGAAATTGTTGCCGGATTG	GAATTAATACGACTCACTATAGGGAGAGATGACGACAGGCTACAGCA
AGAP007385	LYSC4 RNAi F/R	LYSC4	GAATTAATACGACTCACTATAGGGAGAGAAGACGGTGAATCGGGTAA	GAATTAATACGACTCACTATAGGGAGAGTCGTTCAGAAAGTCCTCGC

AGAP005717	LYSC6 RNAi F/R	LYSC6	GAATTAATACGACTCACTATAGGGAGATGACATCTACTGGTGCTCGC	GAATTAATACGACTCACTATAGGGAGAAACTCACTCCACAAGCCCAC
AGAP000694	CEC3 RNAi F/R	CEC3	GAATTAATACGACTCACTATAGGGAGAGAGATCTCTTCCCGTGTGGA	GAATTAATACGACTCACTATAGGGAGAGCGGTGACCTCTTTCAGTCT
AGAP008654	TEP12 RNAi F/R	TEP12	GAATTAATACGACTCACTATAGGGAGAACAAGCTCTAACCTTCGCCA	GAATTAATACGACTCACTATAGGGAGAAGCACTTTGTTGCCTTGCTT
AGAP006941	helicase 2 RNAi F/R	Helicase 2	GAATTAATACGACTCACTATAGGGAGATCTCATCCCACGATCATTCA	GAATTAATACGACTCACTATAGGGAGAACCACTATGTCGACCTTCGG
AGAP003508	helicase 3 RNAi F/R	Helicase 3	GAATTAATACGACTCACTATAGGGAGAAAGGGGGAGAAAGAGATGGA	GAATTAATACGACTCACTATAGGGAGAGGCGCACAGCAGAATATGTA
AGAP005620	DPT RNAi F/R	Diptericin	GAATTAATACGACTCACTATAGGGAGATGCCCTGACAGTTGCATTTA	GAATTAATACGACTCACTATAGGGAGAGTAACTAGCCGAACCGTCCA
AGAP003776	3776 RNAi F/R	3776	GAATTAATACGACTCACTATAGGGAGACGCCATCTGAAGCAACTGTA	GAATTAATACGACTCACTATAGGGAGAGACACCCTAGTTCTTTGCCG

3.8 Generation of LacZ derived ds/ssRNA

A 500bp fragment of the E.coli LacZ gene was cloned into the PLL10 vector by a previous member of the laboratory (S. Pinto, unpublished data), where the LacZ sequence is flanked by T7 promotor sequences and multiple restriction enzyme cutting sites between the T7 promotor sequence and the LacZ sequence. dsRNA was generated using the plasmid as a template for reverse transcription with the T7 MEGAscipt kit according to the manufacturers instructions. ssRNA was generated by using plasmid linearised with Xbal1, cutting the plasmid between the LacZ and the 3' T7 promotor sequence, as a template for reverse transcription using the T7 MEGAscript according to the manufacturers instructions. Linearised plasmid was inspected by gel electrophoresis prior to use in reverse transcription reactions to ensure linearization.

3.9 Transfection of cell lines with dsRNA

3.9.1 Bathing method

Cells were seeded into 6/12/24 or 96 well dishes and allowed to settle for 2-24 hours. Cells were treated with the desired amount of dsRNA complementary to a gene of interest, mixed with complete insect media. Conditioned media was removed from cells and replace with dsRNA containing media. 48h later dsRNA containing media was removed, cells were washed with PBS and complete media was replaced.

3.9.2 Effectene transfection

Cells were seeded into 6/12/24/96 well dishes and allowed to settle for 2-24 hours. Cells were transfected with dsRNA or DNA according to the manufacturers instructions using the equivalent ratios for a 24 well plate: 0.2µg dsRNA/DNA, 1.6µl Enhancer and 5µl of Effectene reagent.

3.9.3 Transmessenger transfection

Cells were seeded into a T25 flask and were transfected with capped ONNV RNA according to the manufacturers instructions using the following ratios: 2µg RNA, 4µl Enhancer and 8µl Transmessenger reagent.

3.10 dsRNA treatment and infection of mosquitoes with 5'ONNViv-eGFP

Mosquitoes were treated with dsRNA and infected with 5'ONNVic-eGFP using three different protocols (see table 3). In all three protocols mosquitoes were inoculated with 307ng of dsRNA. In brief, mosquitoes were either A) inoculated with dsRNA on day 1, inoculated with 5'ONNVic-eGFP at day 4 and 10 mosquitoes were collected at 1, 3 and 6 dpi; B) inoculated with dsRNA on day 1, inoculated with 5'ONVic-eGFP on

day 4 and 30-50 mosquitoes were collected at 7 dpi; or C) mosquitoes were inoculated with dsRNA and 5′ONNVic-eGFP on day 1 (5′ONNVic-eGFP and dsRNA were mixed 1:1 and each mosquito was injected twice) and 30-50 mosquitoes were collected 7 dpi. Equal numbers of mosquitoes were collected for each gene KD.

Table 3. Methods for gene KD and ONNV infection in *A. gambiae* mosquitoes.

Method	dsRNA (Per mosquito)	P1/P2V 5′ONNVic-eGFP (PFU/mosquito)	Day of inoculation	Day of collection
A	307ng	~1640	dsRNA D1, ONNV D4	10 mosquitoes collected 1, 3 and 6 dpi
B	307ng	~1640	dsRNA D1, ONNV D4	~30 mosquitoes collected 7 dpi
C	307ng	~3000	dsRNA and ONNV D1	~30 - 50 mosquitoes collected 7 dpi

3.11 Total RNA extraction from mosquitoes

Total RNA was extracted from whole mosquitoes or cells using TRIzol reagent (Invitrogen, UK). *A. gambiae* whole mosquitoes were homogenised in 500µl TRIzol, 200µl chloroform were added to extract RNA from protein impurities, and RNA was precipitated using 300µl isopropanol. Precipitated RNA was centrifuged at 13000 g for 15 minutes and resuspended in RNAse free water. Total RNA quality was assessed by gel electrophoresis and spectrophotometry using a Nanodrop® (Labtech International). *A. gambiae* cells were lysed in 500µlof TRIzol and RNA was extracted as described for whole mosquitoes.

3.12 Viral RNA extraction

3.12.1 vRNA extraction from cell culture

Conditioned media containing virus was spun at 5000g for 15 minutes to pellet remaining cells. 140µl samples were used for viral RNA extraction using the Qiagen Viral RNA extraction kit following the manufacturers instructions. Samples greater than 140µl in volume were concentrated using YM100 microcon centrifugation tubes: Samples were centrifuged at 15000 g at 12^{o}C for 10-30 minutes (depending on the volume to be concentrated), the reserved supernatant was collected according to the manufacturers instruction and adjusted to 140µl with Drosophila Schneiders medium.

3.12.2 vRNA extraction from whole mosquitoes

Whole mosquitoes were homogenised in 200µl of Drosophila Schneiders medium (Gibco) using a motorised pestle. Homogenates were centrifuged at 3000 g for 30 minutes at 4^{o}C to pellet debris. Supernatant was transferred to a new 1.5ml eppendorf tube and centrifuged at 5000 rpm for a further 30 minutes at 4^{o}C. The supernatant was filtered through a 0.2um filter, and 140µl of the filtrate was used for viral RNA extraction using the Qiagen viral RNA extraction kit according to the manufacturers instructions.

3.13 cDNA Production

cDNA was produced using the Superscript II kit (Invitrogen). For RNA extracted from whole mosquitoes or cell line extracts, 2µg RNA was reverse transcribed according to the manufacturers instructions. For viral RNA extractions from whole mosquitoes, midguts or conditioned media from cell lines, 10µl of extracted viral RNA was reversed transcribed according to the manufacturers instructions.

3.14 Quantitative real time PCR

qrtPCR was performed using the SYBR-Green detection and amplification reagent (Applied Biosystems) according to the manufacturers instructions. Primers for mosquito transcripts were designed using the Primer3 web interface (http://frodo.wi.mit.edu/primer3/). Primers used for qrtPCR are shown in Table 4 . To ascertain transcript abundance of mosquito genes, a relative quantification method was used. For every sample, primers against the transcript of interest, and the housekeeping gene, S7, were used. S7 is constitutively expressed during all stages of the mosquito life cycle. A mixture of all samples to be run on a single plate was used as a reference with S7 primers – all other samples are normalised using the transcript abundance of S7 relative to that of the reference mixture using a standard curve. After normalisation, the relative transcript abundance of the gene of interest is calculated using a standard curve, giving a relative value compared to that of the control sample, thus giving a percentage increase or decrease of transcript abundance compared to the control.

To ascertain the abundance of viral RNA, or viral genome copy number, an absolute quantification method was used. Viral RNA was extracted from a sample with a known PFU, calculated using standard plaque assay, and cDNA was produced as described. A standard curve the sample was generated using neat, 1:5, 1:10. 1:50, 1:100 and 1:500 dilutions of cDNA. This standard curve was used to calculate the viral genome copy number of an unknown sample by mapping the CT value to that of the standard curve, giving the viral genome copy number.

Table 4. Qrt-PCR primers. Primers were designed using Primer3 (http://frodo.wi.mit.edu/primer3/) specifying product size 50-150, primer size 18-25, TM 58-62, primer GC content 30-80. Primers were designed to not overlap with RNAi probes (to avoid amplification of residual dsRNA used for gene silencing) and where possible to span exons.

Gene	Primer ID	Gene name	F	R
n/a	ONNVnsP3-qrt F/R	nsP3	TGA TGA AGC AGT GCC AGT TC	CGA ATT TGC GTA CAT TGG TG
n/a	E2 qrt F/R	E2	CCATACACCTGTGGATGCAG	TACCGTCAGCGTCTCTCCTT
AGAP012352	ML1 qrt F/R	ML1	GTCGCTATTGTGGCATTGTG	AAAGTTTACTACTTCTGCCCAAGC
AGAP012529	GALE8 qrt F	GALE8	CTGCATACAGTCCGCCAAC	GTCCGCGAATCGTAATCTTG
AGAP007385	LYSC4 qrt F/R	LYSC4	GATATCGAGTGTGCGAAGCA	CAGATCGGGCAGTGTCTTTT
AGAP005717	LYSC6 qrt F/R	LYSC6	ACGGTGGCAGTGGCTATTAC	TCAATGTCATCGTCCAGCTC
AGAP005693	5693 qrt F/R	n/a	TCGGTGAGCAACAGTTTGAC	CAGGTCGAGATGGGTGAACT
AGAP010731	CLIPA8 qrt F/R	CLIPA8	GATCGATTCGACGACCAACT	GCAGGTCGACTCGCTTTAAC
AGAP000693	CEC1 qrt F/R	CEC1	TCATCTTTGTCGTGCTGGCAGCTT	TCTTCAGCCGTCCCGCT
AGAP000694	CEC3 qrt F/R	CEC3	GCTACAGCCAGTCGATGGTG	AACACATTGCGTCCGAGCTT
AGAP011294	DEF1 qrt F/R	DEF1	GCCTTTGTGCCGCTC ACT	GCCTTACTGTTGCAGTAACCACC

3.15 Preparation of labelled probes and Microarray Hybridisations

3.15.1 Preparation of labelled probes

Total RNA extracted from *A. gambiae* mosquitoes or cells was amplified and labelled using the Low RNA imput amplification kit (Agilent, UK) according to the manufacturers instructions. In brief: 2µg of total RNA was used in a random primed reverse transcription reaction to generate cDNA. After amplification by conversion to cDNA, cDNA was transcribed to copy messenger RNA (cmRNA) incorporating either Cy-3UTP (for the reference sample) or Cy-5UTP (for the test sample) fluorescent nucleotide analogs. cmRNA quality and labelling efficiency was assessed by spectrophotometry using a Nanodrop® (Labtech International). If cmRNA yield was sufficient and Cy-3UTP or Cy-5UTP labelling was successful, 825ng of RNA was

hybridised to the Agilent 4X44K array in 2X GEx-hybridisation buffer HI-RPM at 60^0C for 17 hours. Hybridised slides were washed with GE wash buffer 1 at RT for one minute and GE wash buffer 2 at 37^oC for one minute, to remove excess labelled cmRNA prior to scanning.

3.15.2 Microarray scanning and data analysis

Microarrays were scanned using a GenePix semiconfocal microarray scanner (AXON Instruments, Foster City, CA) Gene Pix Pro 4.0 or 6.1 were used to record feature signal intensity, to eliminate local backgrounds, for grid alignment and manual inspection of feature quality. Average feature diameter was calculated and features lying outside three standard deviations of the mean were excluded from analysis. The ratio of feature intensity verses local and global backgrounds were calculated and features not exceeding background intensities were excluded from analysis. Features were normalised using Genespring 6.1 (Axon instruments) by locally weighted linear regression methods (Lowess). Feature intensities over the three or four biological replicates were averaged. T-test p-values were calculated, and normalised data was filtered to exclude data with p-values greater than 0.05. Data was further filtered to include only genes showing 2-fold and greater regulation. Candidate genes were selected based on several criteria, including gene ontology, and known roles of orthologous genes.

3.16 Immunofluorescence Assays

Cells were plated into sterile coverslip in 6 well plates (coverslips were dipped in 75% ethanol and flamed before being placed into 6 well plates). Plates were seeded with 2 x 10^6 cells/well and allowed to reach confluency over 1-2 days. Cells were infected with 1 MOI P1/P1V 5'ONNVic-eGFP by adding virus directly to conditioned media. At the appropriate time point after infection, the coverslip was removed from the well plate and washed once gently with PBS. Cells on the coverslip were fixed with 4% PFA in PBS for 15 minutes at RT. Cells were washed twice with PBS. Cells were

permeabilised with 0.5% Triton X 100 in PBS for 20 minutes, followed by two quick washes with PBS. Cells were blocked with 3% Bovine Serum Albumin (BSA), 0.1% Tween 20 in PBS for 2 hours. Cells were washed twice with PBS for 10 minutes. Cells were incubated with a rabbit anti-STAT1, generated by a previous worker in this laboratory[100], at a 1:200 dilution in 3% BSA in PBS at 4^{o}C overnight. Cells were washed twice for 10 minutes with 3% BSA in PBS, and then incubated with secondary antibody - ALEXA FLUOR 568 mouse anti-rabbit (Molecular probes) at 1:1000 dilution in 3% BSA in PBS for 3-4 hours at RT. Cells were washed twice with PBS and mounted on glass slides with Vectorshield© with DAPI (Vectorlabs). GFP, DAPI and anti-STAT staining were imaged using a Leica DMT fluorescence microscope and a Zeiss AxioCam HRc camera with Zeiss Axiovision40 software.

3.17 Coinfection of *A. gambiae* with ONNV and *P. berghei*

3.17.1 Parasite maintenance

P. berghei ANKA clone 259c12 (that constitutively expresses GFP throughout it's lifecycle) was maintained in Theiler's original mice (Harlan, UK) as described in Sinden *et al* 2002 [98]. All animal work was carried out by Dr Tibebu Habetewold and Kasia Sala. Mice were infected by intraperitoneal (IP) injection of 100-200µl of *P. berghei* infected blood. Parasitaemia was calculated from methanol fixed and Giemsa stained thin tail blood smears; blood smears were air dried, dipped in methanol and covered in Giemsa solution (1:5 dilution in Giemsa buffer – 0.7% (w/v) anhydrous KH2PO4, 1% (w/v) anhydrous Na2HPO4) for 15 minutes at RT. For mosquito infections, three days after passage with infected blood mice were terminally anaesthetised with an intramuscular (IM) injection of 0.05ml/10g body weight of Rompun (2% stock solution, Bayer), Ketastet (100mg/ml ketamine, Fort Dodge Animal Health Ltd) and PBS in a 1:2:3 ratio).

3.17.2 Infection of mosquitoes with 5'ONNVic-eGFP and *P. berghei*

Newly emerged adult G3 mosquitoes were intrathoracically inoculated with ~1640 PFU P1/P2V 5'ONNVic-eGFP. Inoculated mosquitoes were maintained at 27^{o}C for 48h. Mosquitoes were starved of sugar for 4-5 hours prior to blood feeding. Mosquitoes were fed on a terminally anaesthetised *P. berghei* infected mouse by placing the mouse over the netted tops of the mosquito pots, alllowing mosquitoes to probe through the net for 30 minutes-1 hour in the dark. Mosquitoes were maintained at 19^{o}C for 72h post blood feeding to allow successful parasite development, and were subsequent maintained at 27^{o}C to allow for optimal viral replication. Unfed mosquitoes were removed between 24 and 48h post blood feeding, when the blood bolus is clearly visible through the abdomen of the mosquito. Seven days post blood feeding, mosquito midguts were dissected and fixed in 4% PFA (midguts were fixed in 4% PFA in PBS for 45-60 minutes, and subsequently washed in PBS three times for 15 minutes). Fixed midguts were mounted in Vectorshield© (Vectorlabs) on glass slides with sealed coverslips. Dissected midguts were kept in the dark wherever possible to prevent bleaching of GFP.

3.17.3 Live oocyst and melanised ookinete counts

Fixed and mounted midguts were observed under 20X magnification using a fluorescent microscope. Live oocysts expressing GFP were counted using fluorescence, and melanised ookinetes were counted using light microscopy.

3.17.4 Statistics

For *P.berghei* oocyst, melanised ookinete counts and ONNVic-eGFP plaque assays tests results were compared to controls using the Mann Whitney U test; statistical significance was accepted when $P<0.0001$ (***), $P<0.001$ (**), $P<0.01$(*). For microarray analysis test results were compared to controls using the students

unpaired T-test. Statistical significance was accepted when $P<0.05$. For *per os* infections, tests midguts were compared to controls using the Chi squared test. Statistical significance was accepted when $P<0.0001$ (***), $P<0.001$ (**), and $P<0.01$ (*). For infection of *A.gambiae* derived cell lines, viral RNA genome copy number in conditioned media was compared to controls using the students unpaired T-test. Significance was accepted when $P<0.0001$ (***), $P<0.001$ (**), $P<0.01$(*).

4 5′ONNVic-eGFP: production, propagation and characterisation in *A. gambiae* mosquitoes and cell lines

4.1 Introduction

A. gambiae-ONNV is an unusual vector-virus combination, with ONNV being the only known arbovirus to be transmitted by Anopheline mosquitoes. This chapter will describe the generation of infectious clones of ONNV that express GFP and the dynamics of ONNV infection in adult *A. gambiae* mosquitoes and in *A. gambiae* derived cell lines. Characterising ONNV infection in *A. gambiae* was important for the design of subsequent microarray studies and gene silencing experiments.

4.2 Generation of infectious clones 5'ONNVic-eGFP

In order to study the mosquito immune responses to viral infection, in collaboration with Colorado State University, Arthropod Infectious Diseases Laboratory (AIDL), an ATS of ONNV was obtained from Dr Brian Foy. The ATS was generated by Brault *et al (*2004)[42], using the genome sequence of the SG650 strain of ONNV isolated from human serum in Uganda during the 1996 outbreak in Africa [29]. Five overlapping fragments of the SG650 genome were generated by high fidelity RT-PCR and were subcloned into pBluescript II SK(+)[42]. The fragments were combined by ligating segments digested with rare cutting restriction enzymes. Enhanced GFP (eGFP) was subsequently added along with a duplicate of the ONNV subgenomic promoter (figure 8)[42], thus when the subgenomic promoter is switched on during infection, GFP is expressed. The first nucleotide of the virus genome was modified to allow for the addition of a cap analogue [19]. Having obtained this ONNV ATS, infectious clones were generated (section 3.4). In brief an RNA copy of the viral genome was transcribed using the T7 promotor in pBluescript II SK(+), from plasmid linearised using the unique restriction site Not1 as a template. An A-nucleotide cap analog was added to the transcription reaction, to generate 5' capped transcripts of the viral genome. The capped transcripts were transfected into VERO cells using the Transmessenger kit (Invitrogen). Transfected cells were monitored for cytolytic infection and GFP expression every 12 hours. GFP expression was observed at 24 hours after transfection: single cells expressing GFP could be seen (figure 9), in some cases the infected cells were already rounding up and dying. Infection spread rapidly

throughout VERO cells, with spreading foci of infection visible at 48 hours post infection (hpi). At 72h extensive cell death was seen, presumably caused by apoptosis as is typical in mammalian cells after infection with alphaviruses. 72 hours after successful transfection, the first passage virus (P1V 5'ONNVic-eGFP) was harvested (by scraping the cell monolayer), aliquoted and stored at -80^{o}C until further use.

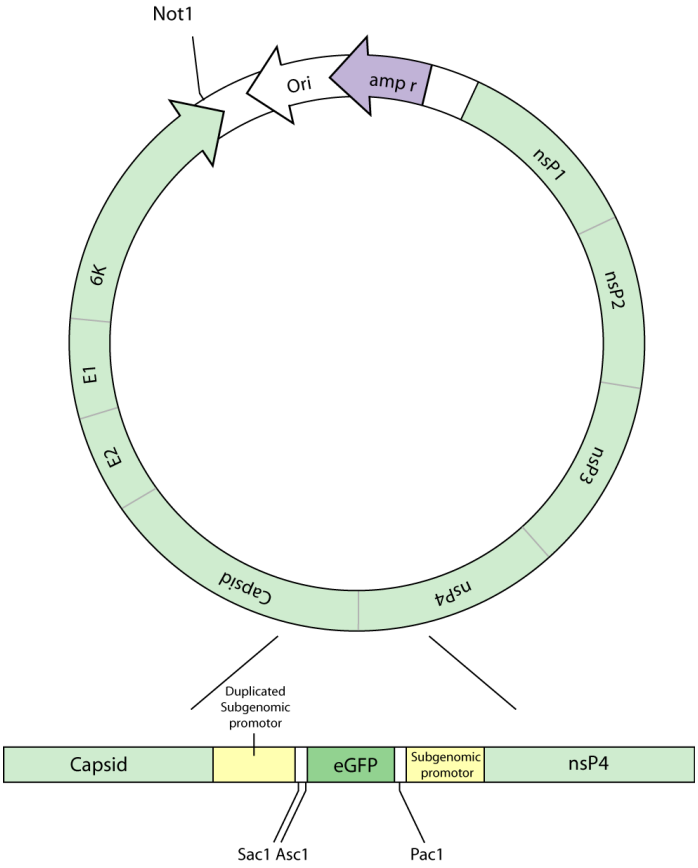


Figure 8. Plasmid map of 5'ONNVic-eGFP; the full length genome of the SG650 strain of ONNV was cloned into pBluescript II SK(+), the viral subgenomic promoter was duplicated and eGFP was placed under control original viral subgenomic promoter by Brault *et al[42]*. The plasmid was kindly given by Dr Brian Foy, Colorado State University, Arthropod Infectious Diseases Laboratory (AIDL).

4.2.1 Propagation and calculation of titre of infectious clones

To obtain large quantities of high titre 5'ONNVic-eGFP for use in experiments, the virus was passaged a second time. Serial passage of 5'ONNVic-eGFP results in the gradual loss of GFP expression, and decreased ability of the virus to infect cells[42]. Thus 5'ONNVic-eGFP was only ever passaged for a second time before use in experiments.

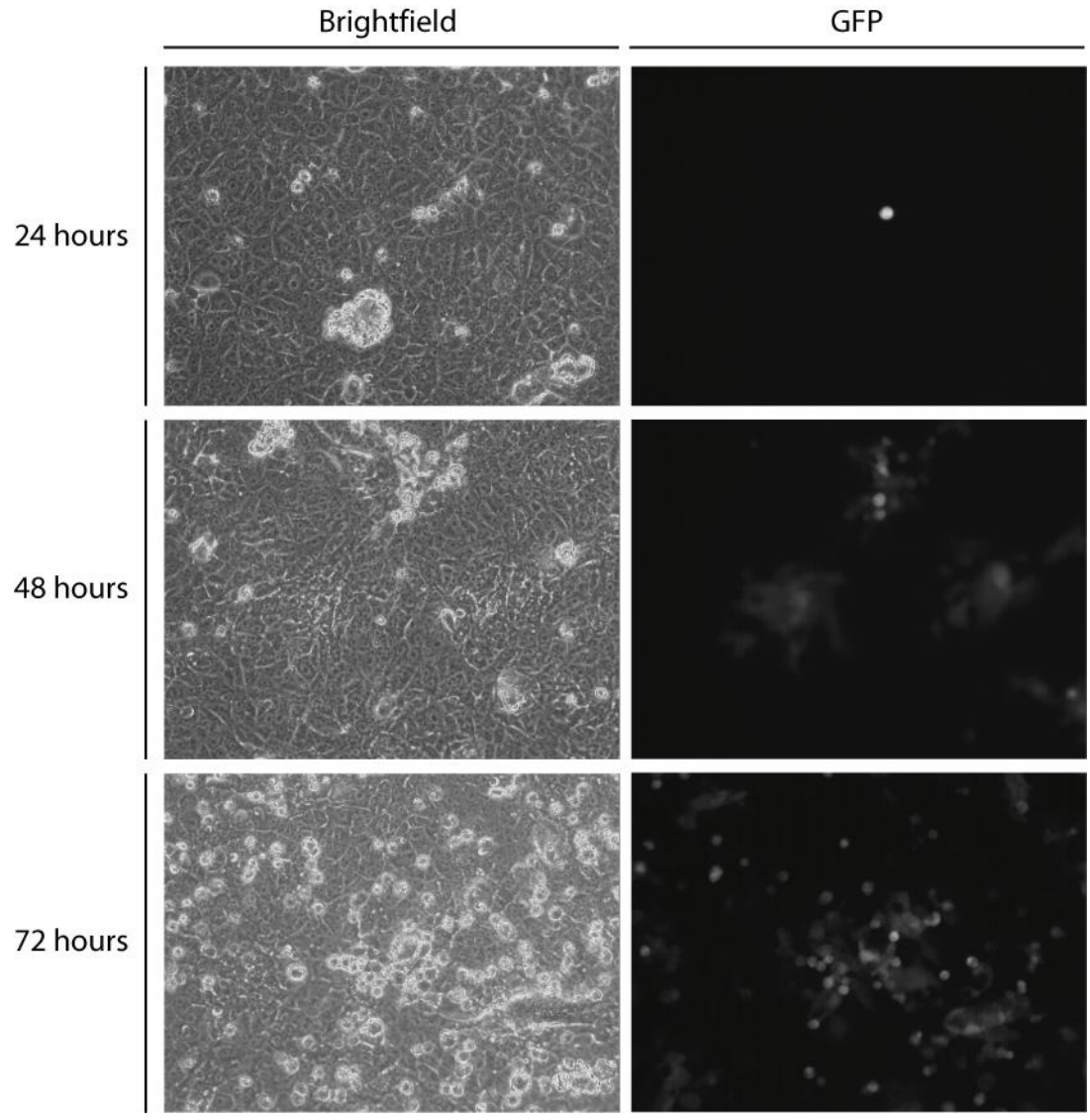


Figure 9. Transfection of VERO cells with 5'ONNVic-eGFP: Confluent cultures of VERO cells were transfected with capped transcripts of 5'ONNVic-eGFP. Successfully transfected cells expressed GFP (white) at 24h. At 48h, 5'ONNVic-eGFP had spread to neighbouring cells. At 72h 5'ONNVic-eGFP infection had spread across the whole monolayer and infected cells had rounded up and died.

To obtain second passage virus, 250µl of P1V 5'ONNVic-eGFP was added to a confluent culture of VERO (mammalian) cells or L35 (mosquito) cells in a T175 flask. The virus was allowed to infect cells for 72 hours, when cells were scraped, filtered through a 0.22µm filter, and stored at -80^{o}C in aliquots until required. Second passage virus was thereafter referred to as P1/P2V when passaged through VERO cells for the second time, or P1V/P2L35 when passaged through L35 cells for the second time. The titres of P1/P2V and P1V/P2L35 were ascertained using standard plaque assay in VERO cells (as described in material and methods). Table 5 gives the Plaque forming units (PFU) calculated for all batches of viruses produced and used in this thesis. The highest titre obtained was from virus passaged the second time in mosquito cells (P1V/P2L35).

Table 5. PFU/ml for P1 and P1/P2 5'ONNVic-eGFP stocks calculated using standard plaque assay in VERO cells.

Batch	PFU/ml
P1V batch A (14/6/8)	2.6 X 10^7
P1 V batch B (30/6/8)	5.2 X 10^7
P1/P2V batch A (18/6/8)	7.4 X 10^6
P1/P2V batch B (18/6/8)	8 X 10^ 6
P1/P2V batch C (1/7/8)	7.4 X 10^7
P1/P2V batch D (26/6/9)	5.45 X 10^7
P1/P2V batch E (14/7/9)	1.73 X 10^8
P1V/P2L35 (28/7/8)	5 X 10^8

4.3 5′ONNVic-eGFP infects *A. gambiae* derived cell lines

Seven cell lines derived from *A. gambiae* larvae were tested for susceptibility to 5′ONNVic-eGFP infection. Initially cells were infected at a multiplicity of infection (the number of infectious viral particles or PFU for every cell; MOI) of 0.25 i.e. ~ 0.25 viral particles for every cell in the culture. Two cell lines showed substantial susceptibility to viral infection: L35 derived from neonates of the L35 strain of *A. gambiae* and 4a3A derived from the neonates of the 4a r/r strain of *A. gambiae* (figure 10). The L35 cell line was selected for use in the majority of cell based assays during this project.

Initially infection of L35 cells was optimised. Cells were infected at 0.1, 1 and 5 MOI with and without washing cells in PBS prior to infection. Washing cells had no discernable effect on viral infection (data not shown). Cells infected with 5 MOI showed little infection, and the cells showed signs of cell death (data not shown). Consequently 1 MOI was used to characterise infection in L35 cells.

The dynamics of infection was investigated over a 96h time course. L35 cells were infected at 1MOI by adding virus directly into conditioned media of confluent cultures in 6 well plates, and rocking the cells for 30 minutes. Cells were incubated at 27^{o}C. Conditioned media from cells was collected at 1, 3, 6, 9, 12, 24, 48, 72 and 96hpi. Cells were also collected in TRIzol at the same time points. Viral RNA was extracted from conditioned media using the QiaAmp viral kit. Total RNA was extracted from cells collected in TRIzol. Qrt-PCR was used to ascertain viral genome copy number from conditioned media and total RNA samples. cDNA was generated from 10µl of viral RNA (vRNA) from conditioned media and 1µg of total RNA (tRNA) from cells. Qrt-PCR was carried out using a standard curve consisting of a serial dilution of a sample of a known PFU (ascertained using plaque assay) and the viral genome copy number was calculated. Figure 11A shows the dynamics of viral infection over 96h in conditioned media (i.e. released virus) and within cells (i.e. replicating virus within the cytoplasm of infected cells). During the first 12 hours

after infection no GFP was expressed and no virus was released from infected cells. Between 12-24 hours the first cells expressed GFP and virus was released into conditioned media (Figure 11B). The peak of infection occurred at 72 hpi, and began to drop at 96 hpi.

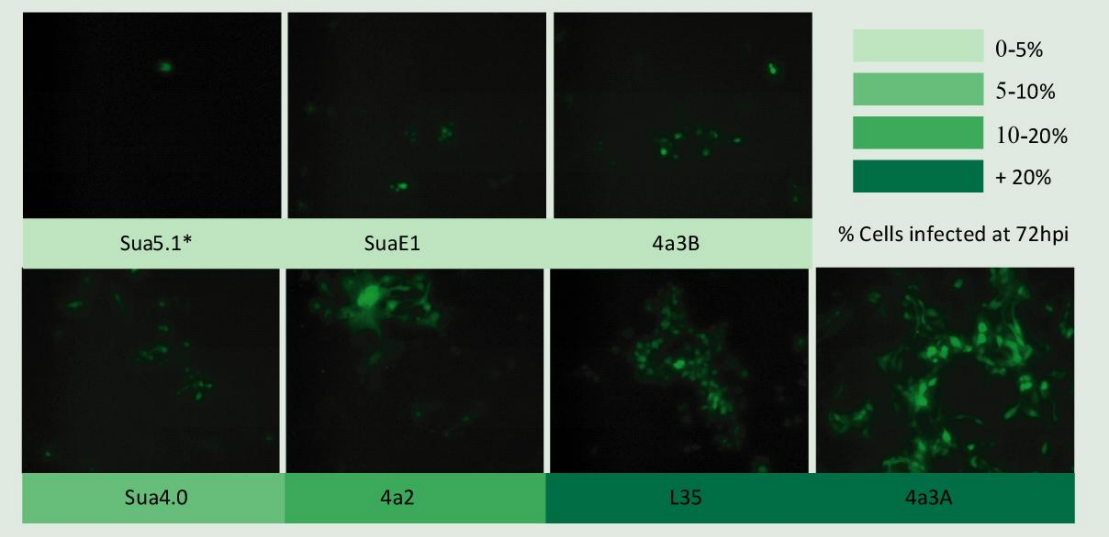


Figure 10. 5'ONNVic-eGFP infection in 7 *A. gambiae* derived cell lines: Cells were grown to confluency in T25 flasks, and were infected with 0.25 MOI P1/P2V 5'ONNVic-eGFP by adding virus directly to the conditioned media and rocking the cells for 30 minutes. Cells were incubated at 27^{o}C. Pictures were taken at 72 hpi. Coloured bars represent the average % of cells expressing GFP counted from three fields of view at 72 hpi.

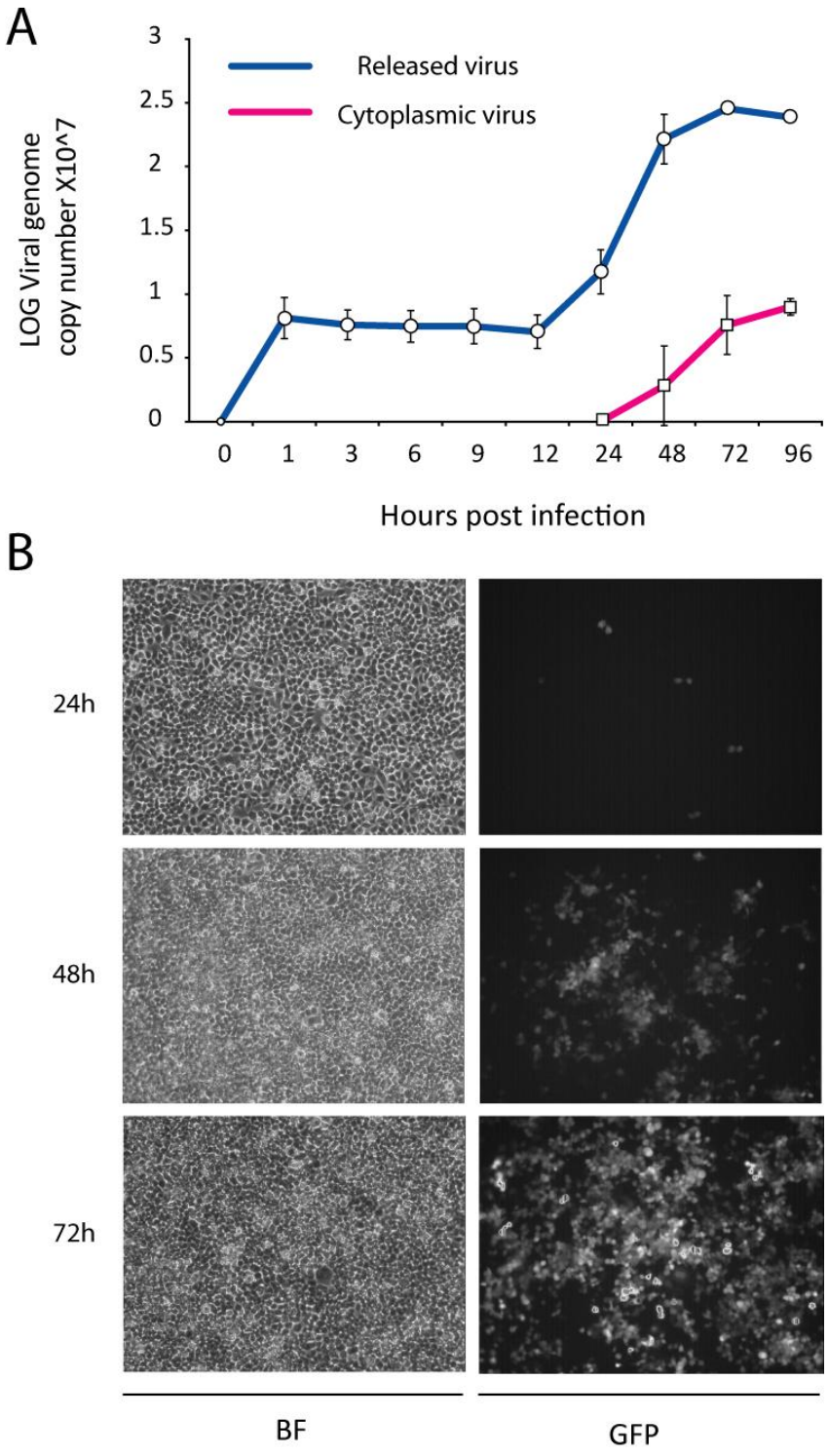


Figure 11. Dynamics of 5'ONNVic-eGFP infection in L35 *A. gambiae* cells; Cells were grown to confluency in 6 well plates, and were infected with P1/P2V 5'ONNVic-eGFP at 1MOI. **A)** LOG_{10} Viral genome copy number/ml in conditioned media (released virus) and total RNA extracted from cells (cytoplasmic virus) over 96 hours was calculated using qrt-PCR. Error bars standard error of three biological replicates. **B)** Brightfield (BF) and GFP expression (white) images of infected cells at 24, 48 and 72 hpi.

4.4 5′ONNVic-eGFP infects *A. gambiae* after intrathoracic innoculation

~1640 PFU of 5′ONNVic-eGFP was injected into the thorax of adult *A. gambiae* mosquitoes. Initially two strains of *A. gambiae* (G3 and Yaounde) were tested using P1/P2V and P1V/P2L35 5′ONNVic-eGFP. Ten innoculated mosquitoes were collected at 1, 3 and 6 days post infection (dpi). Mosquitoes were homogenised in Schneiders *Drosophila* medium, vRNA was extracted using the QiaAmp Viral kit, and qrt-PCR was used to calculate the vial genome copy number per mosquito. Infection in G3 strain of *A. gambiae* with P1/P2V 5′ONNVic-eGFP was far higher than that observed using P1V/P2L35 5′ONNV and infection with either in the Yaounde strain of *A. gambiae* (figure 12). G3 mosquitoes were subsequently used in the majority of experiments using adult mosquitoes.

The dynamics of 5′ONNVic-eGFP infection in G3 mosquitoes was investigated further; a time course following GFP expression and viral genome copy number/mosquitoes was carried out over nine days (figure 13A/B). Mosquitoes were injected with ~1640 PFU of P1/P2V 5′ONNVic-eGFP. GFP expression was checked and 10 inoculated mosquitoes were collected each day for 9 days. VRNA extraction and qrt-PCR was carried out (section 3.12.2. and 3.14). Viral RNA genome copy number peaked at 6 dpi, levelling out at 7 dpi, and decreasing by 9 dpi. GFP expression conversely increased and the highest number of GFP expressing mosquitoes was observed at 9 dpi (figure 13B). GFP expression could be seen often in the head tissues (seen clearly through the ommaditia of the eyes) and occasionally in the thorax through the cuticle (figure 13D). After injection, mosquitoes were dissected at 7 dpi; head squashes, thorax and abdomen preparation and fixed guts were observed for GFP expression. Autofluorescence in the head and thorax preparations meant that GFP could not be distinguished from uninfected tissue (data not shown), however prepared guts showed GFP expression in what appear to be muscle bands or nerves running close to the muscle bands around the anterior- and mid-gut (figure 13C).

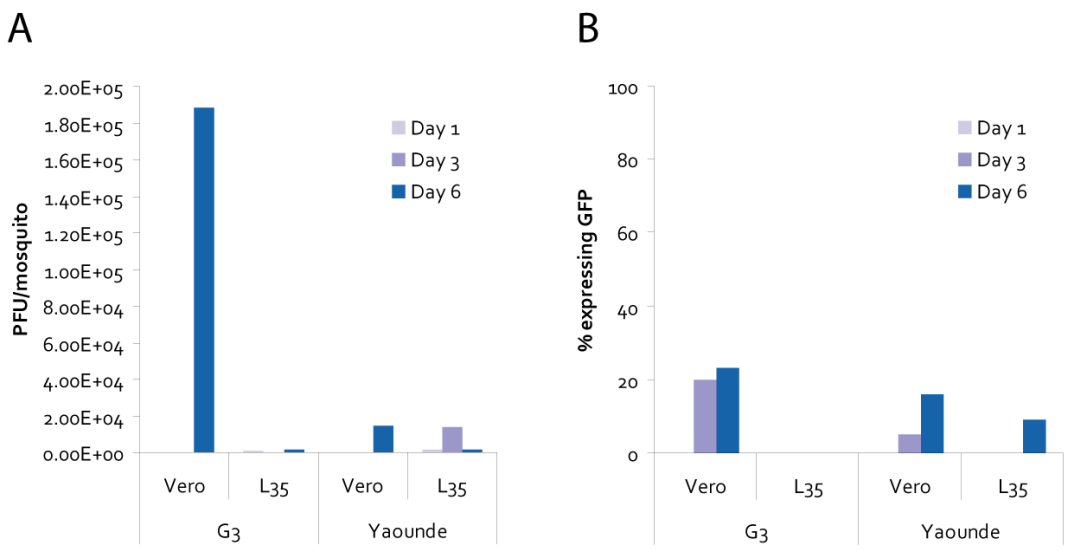


Figure 12. ONNV infects adult *A. gambiae* mosquitoes; G3 and Yaounde *A. gambiae* mosquitoes were intrathoracically inoculated with ~1640 PFU P1/P2V (Vero) or P1V/P2L35 (L35) 5'ONNVic-eGFP. 10 mosquitoes were collected at 1,3 and 6 dpi. **A)** viral genome copy number/mosquito ascertained using qrt-PCR and **B)** GFP expression observed by fluorescent microscopy in inoculated mosquitoes.

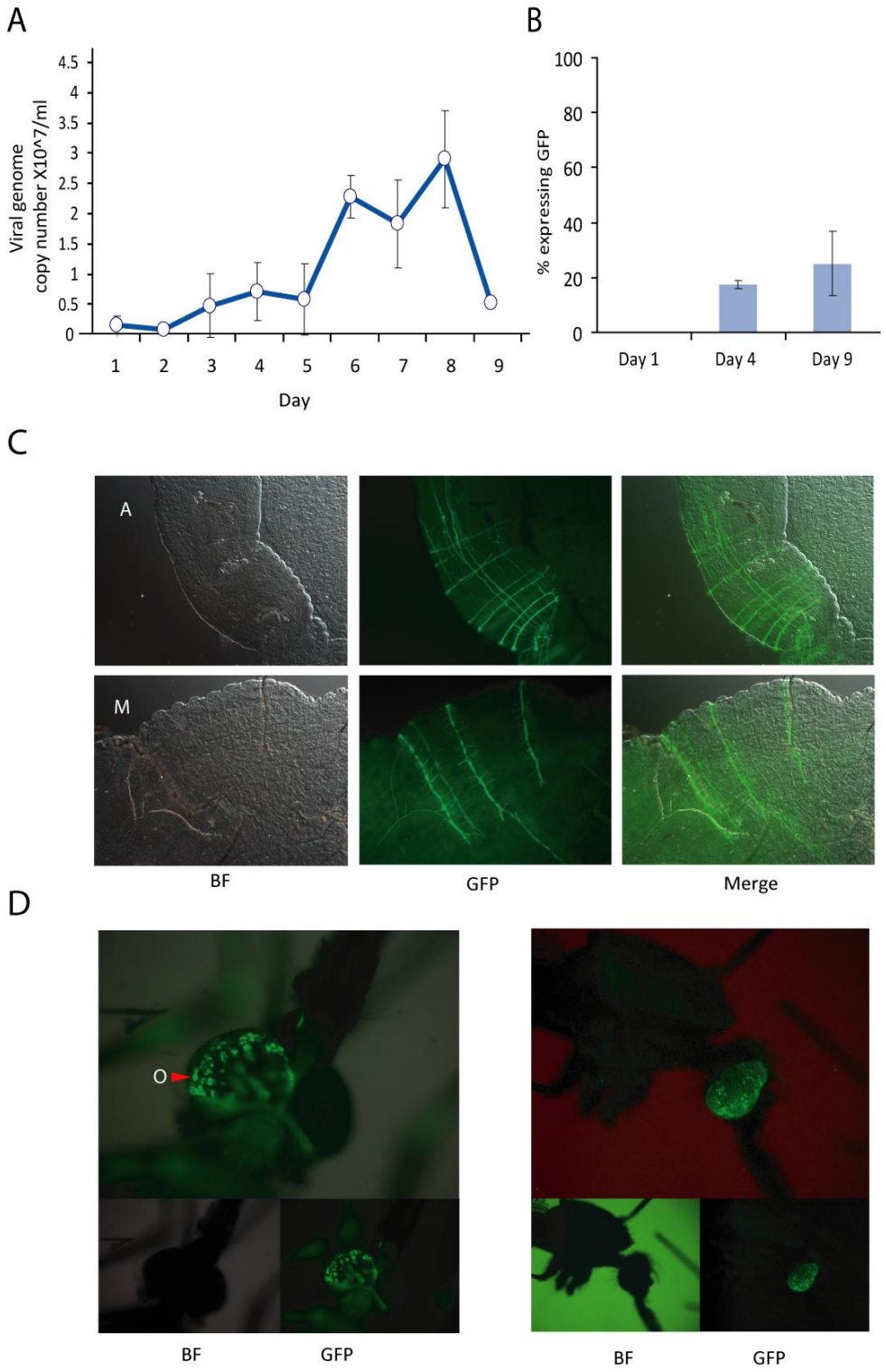


Figure 13. ***A. gambiae*** **G3 mosquitoes intrathoracically inoculated with 5′ONNVic-eGFP**. Cohorts of 50 mosquitoes were inoculated with ~1650 PFU of P1/P2V 5′ONNVic-eGFP. **A)** 10 inoculated mosquitoes were collected daily and qrt-PCR was used to ascertain viral genome copy number/mosquito. **B)** Percent of inoculated mosquitoes showing GFP expression at 1, 4 and 9 dpi. **C)** GFP expression in nerves and/or muscle bands in the anterior- (A) and mid-gut (M) of inoculated mosquitoes and **D)** GFP expression in the head tissues through the ommatidia (O) of inoculated mosquitoes at 9 dpi. Error bars represent SD of 3 biological replicates.

4.5 5′ONNVic-eGFP infects *A. gambiae* mosquitoes after *per os* infection

To see if 5′ONNVic-eGFP can infect *A. gambiae* mosquitoes by the natural *per os* route of infection, cohorts of two strains of *A. gambiae* (G3 and N'gousso) were given infectious bloodmeals via membrane feeders. Compacted human red blood cells (RBCs) were mixed in a 1:1 ratio with human serum. Virus was mixed with the resuspended RBCs; initially the same dose per mosquito that was used during intrathoracic inolculation was added (~1640 PFU), based on the assumption that each mosquito will consume ~ 2µl of blood. Infectious blood was loaded into membrane feeders and mosquitoes were allowed to feed for an hour in the dark. Mosquitoes were maintained at 27^{o}C. Bloodfed mosquitoes were separated from unfed mosquitoes 24-48h after feeding. Bloodfed mosquitoes were dissected at 3 and 6 dpi; midguts were removed from the carcass, taking care to keep the anterior gut and cardia attached, and fixed in 4% paraformaldehyde (PFA) for 30-60 minutes. Fixed guts were washed three times with PBS before mounting on slides in Vectorshield with DAPI. Anterior-, mid-guts and the cardia were observed using a fluorescent microscope to check for GFP expression. Prevalence of infection ranged from ~40-60% in N'gousso mosquitoes and 40-80% in G3 mosquitoes during different replicates of the same experiment (data not shown). The effect of antibiotic treatment on prevalence of viral infection was investigated; mosquitoes were fed on 10% fructose containing gentamycin, penicillin and streptomycin for 5 days prior to feeding on an infectious bloodmeal. Treatment of antibiotics did not effect the prevalence of viral infection in G3 (Figure 14A) or N'gousso (data not shown) mosquitoes. The dependence of prevalence on dose was also investigated; mosquitoes were fed with a single dose (~1640 PFU) or a double dose of virus (~3280 PFU) and prevalence was assessed at 3 and 6 dpi. Infection was seen to be dose dependent, with prevalence increasing around 2.6 fold at day 3, and 1.5 fold at day 6 when a double dose was given (Figure 14B).

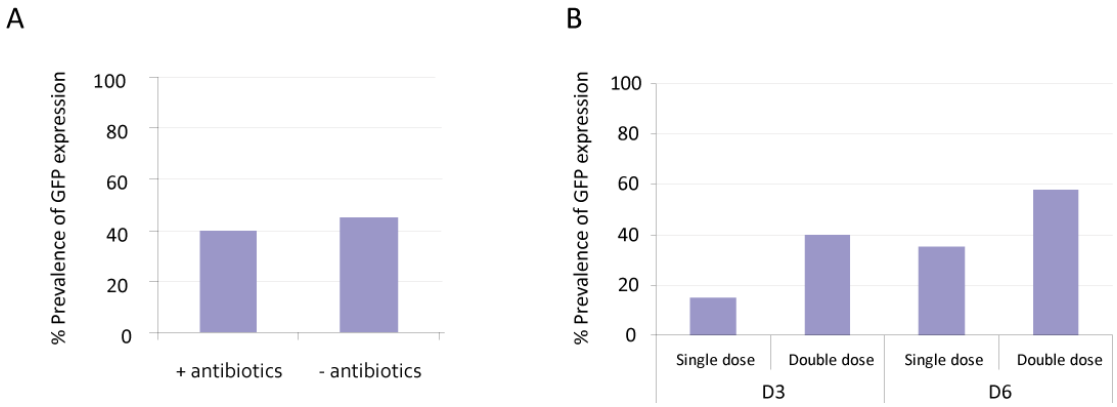


Figure 14. Effect of antibiotics and dose on infection in *A. gambiae* mosquitoes; **A)** Prevalence of GFP in cardia/anterior gut/midguts dissected at 6 dpi; G3 mosquitoes were fed on 10% fructose with (+) or without (-) antibiotics for 3 days prior to blood feeding. Mosquitoes were starved of sugar overnight, and then allowed to feed on an infectious bloodmeal. At 6 dpi midguts with a complete anterior gut and cardia were screened for GFP expression. **B)** Prevalence of infection is dose dependent; mosquitoes were allowed to feed on an infectious bloodmeal containing a single dose (~1640 PFU) or a double dose (~3280 PFU) of virus. At 3 and 6 dpi midguts with a complete anterior gut and cardia were screened for GFP expression.

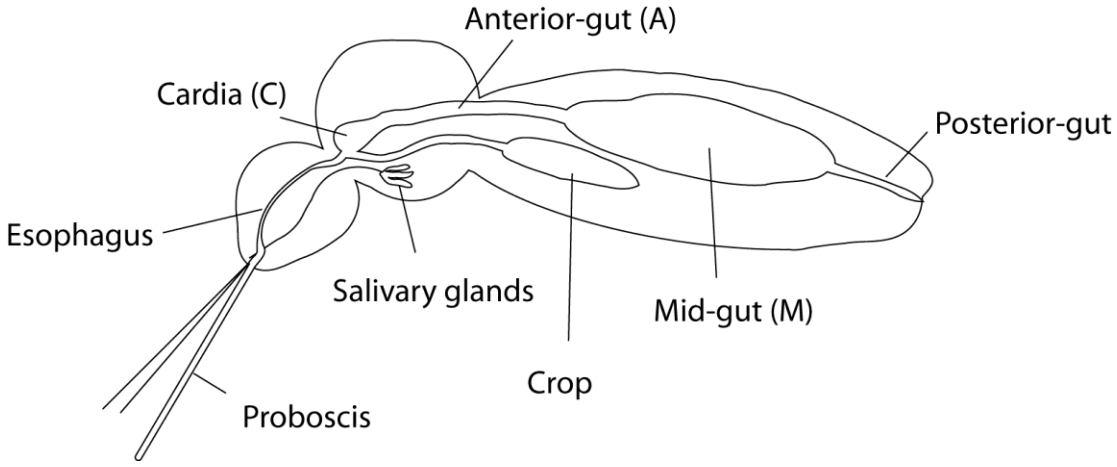


Figure 15. Diagram of the mosquito alimentary canal. Blood meals are taken up through the proboscis, to the midgut (M) via the esophagus, cardia (C) and anterior-gut (A). Sugar and water are taken up into the crop. Adapted from Gusamo *et al* 2007[101].

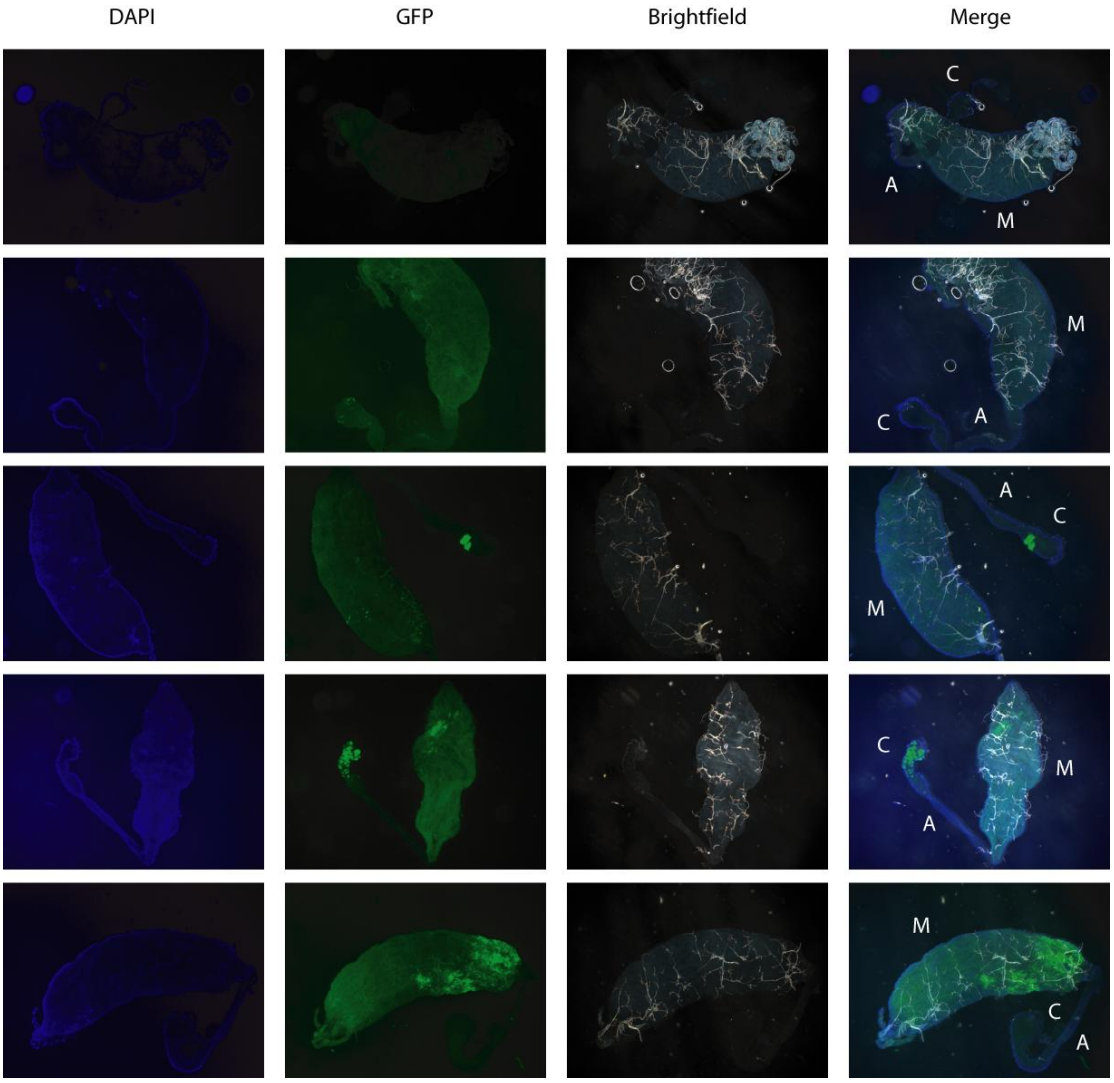


Figure 16. Infection patterns of 5'ONNVic-eGFP in *per os* infected *A. gambiae* mosquitoes at 5 X magification; Infection patterns in *per os* infected mosquitoes; G3 mosquitoes were given infectious bloodmeals of ~5000 PFU/mosquito. Pictures were taken of dissected midguts with an intact cardia and antierior gut at 6 dpi. GFP expression, and thus viral infection, can be seen frequently in the cardia (C), anterior gut (A) and occasionally in the midgut (M).

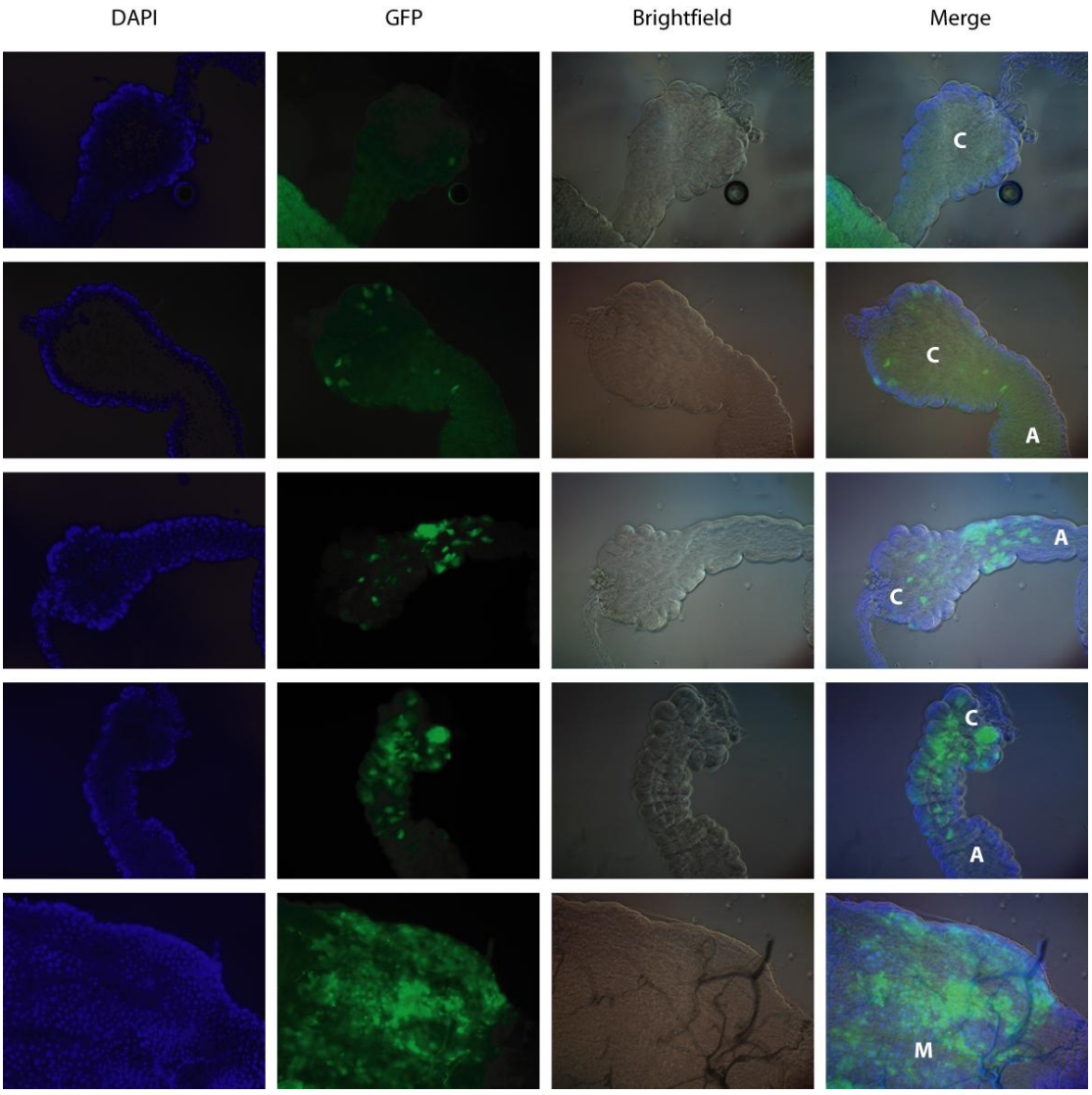


Figure 17. Infection patterns of 5'ONNVic-eGFP in *per os* infected *A. gambiae* mosquitoes at 20 X magification; Infection patterns in *per os* infected mosquitoes; G3 mosquitoes were given infectious bloodmeals of ~5000 PFU/mosquito. Pictures were taken of dissected midguts with an intact cardia and antierior gut at 6 dpi. GFP expression, and thus viral infection, can be seen frequently in the cardia (C)), anterior gut (A) and occasionally in the midgut (M).

To further investigate susceptibility to infection, N'gousso mosquitoes were fed with a high dose (~5000 PFU/mosquito) of P1/P2V 5'ONNVic-eGFP. Prevalence of infection was 48% (259 out of 519 mosquitoes) at D3 post feeding, dropping to 38% (267 out of 697 mosquitoes) by D6 post feeding. Figure 15 outlines the alimentary canal of a typical mosquito indicating the positions of the cardia, the anterior gut and

the midgut. Tissue tropism of infection was restricted mainly to the cardia and anterior gut, with typically small clumps or several single cells expressing GFP (figure 16 and 17). Some individuals displayed widespread infection of the cardia and anterior midgut (10% of blood fed mosquitoes at D3 dropping to 9% by D6) and fewer individuals still demonstrated infection of the midgut, typically small patches, sometimes extending across around half the midgut (7% of blood fed mosquitoes at D3 dropping to 4% by D6).

4.6 Cross species comparison of ONNVic-eGFP infection in *Anopheles gambiae*, *Anopheles arabiensis* and *Anopheles quadriannulatus*.

Three different species of *Anopheles* mosquitoes were infected with P1/P2V 5′ONNVic-eGFP; *A. gambiae*, *A. arabiensis* and *A. quadriannulatus*. Cohorts of 50 mosquitoes were intrathoracically inucolated with ~1640 PFU. Survival rates of infected mosquitoes were recorded (see Figure 18 A/B/C/D) and at day 7, viral titre for the remaining mosquitoes was calculated using qrt-PCR. *A. quadriannulatus* consistently showed substantial death compared to control inoculated mosquitoes, with only around 35% of mosquitoes surviving to 7 dpi. *A. arabiensis* also showed some death associated with viral infection, however this was not consistent between replicates. G3 and N'gousso, two strains of *A. gambiae*, showed no difference in survival rate when infected with 5′ONNVic-eGFP compared to control mosquitoes. Interestingly *A. quadriannulatus*, as well having reduced survival rates to infection compared to *A. gambiae* and *A. arabiensis*, also showed the highest titre of viral infection (Figure 18D).

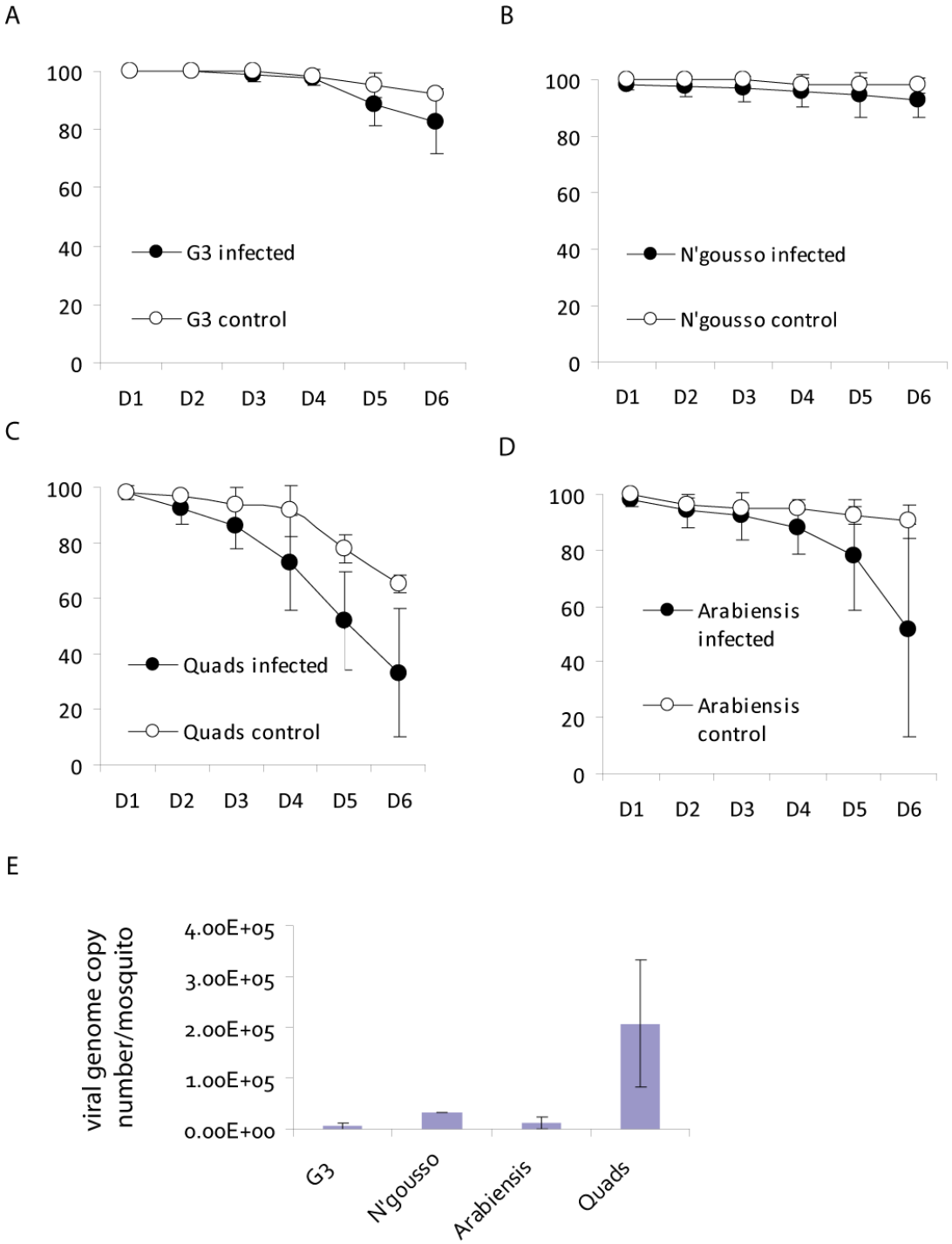


Figure 18. Infection of three different mosquito species; Survival curves for **A)** *A. gambiae* N'gousso strain, **B)** *A. gambiae* G3 strain, **C)** *A. quadriannulatus* and **D)** *A. arabiensis* intrathoracically inoculated with ~ 1640 PFU/mosquito 5'ONNVic-eGFP. **E)** Viral RNA genome copy number/mosquito at 6 dpi calculated by qrt-PCR. Error bars represent SD from two biological replicates.

4.6.1 Summary

This chapter has covered:

- Production of infectious clones of 5′ONNVic-eGFP; GFP expressing clones of ONNV were generated in a mammalian cell line.
- Propagation and titring of infection clones; Infectious clones were propagated for use in experiments and standard plaque assay was used to titre viral stocks.
- Infection of *A. gambiae* derived cell lines; Seven *A. gambiae* derived cell lines were tested for susceptibility to 5′ONNVic-eGFP infection. Infection dynamics in one of those cells lines, L35, was optimised and characterised.
- Infection of adult *A. gambiae* mosquitoes: Intrathoracic inoculation and *per os* infection; Adult *A. gambiae* mosquitoes were tested for susceptibility to 5′ONNVic-eGFP through two infection routes; intrathoracic inoculation directly into the hemolymph of the mosquito, or through infectious blood meals.
- Cross species comparison of 5′ONNVic-eGFP infection; Differences in susceptibility to 5′ONNVic-eGFP between three *Anopheles* spp. was investigated

4.7 Discussion

4.7.1 Generation of infectious clones and infection of mammalian and insect cells

ATSs are invaluable tools for the study of host-pathogen interactions between viruses and their mosquito hosts, allowing large scale production of viruses genetically modified to incorporate markers, genes of interest or RNAi cassettes. In these studies, an ONNV ATS expressing GFP developed by Brault et al [42] was tested for infection *in vivo* and *in vitro*. The advantages and limitations of the GFP marker are discussed below.

Two of the seven *A. gambiae* cell lines tested for infection showed significant susceptibility - the L35 and the 4a3A cell lines. The *A.gambiae* cell lines were derived from newly hatched larvae of several *A.gambiae* strains. These are primary cell cultures that have a heterogeneous population of cells with varying propterties form cell line to cell line. Microarray analysis of cell line responses to pathogen challenge show immuno-responsiveness (Dr G Christophides, personal communication) and as such the cell lines are considered to be a good model for studing immune genes. However differences in gene expression do exist between these cell cultures and may change over time due to the heterogenicity of the cells. It is possible that the 4a3A and L35 cell lines are less immuno-competent and thus are more susceptible to infection that the other cell lines tested. It is known that some cell lines lack the ability to express certain immune genes, for example the 4a3A cell line does not express pro-phenoloxidases (PPOs; exzymes involved in the melanisation cascade) whereas the 4a3B cell line expresses 6 PPOs, despite being derived from the same initial pool of newly hatched larvae [99]. The ability of a cell line to express immune genes may explain the differences in infecrtivity observed, although the presence or absence of a receptor required for infection could also explain this.

In both cells lines, not all cells within an infected culture become GFP positive, even after several days of infection. It is unclear why these cells remain apparently uninfected – perhaps due to the heterogenicity of the cells, some lack a receptor utilised by the virus for infection, or perhaps some cells are more immuno-competent than others and thus prevent or clear infection. Another possibility is that the virus may no longer be expressing GFP during infection. The fast and efficient replication cycle of viruses results in very short generation times compared to many other organisms, allowing genetic mutations to appear, and if advantageous, spread through a population quickly. Additionally RNA virus populations are believe to exist as a plethora of quasispecies, where due to high rates of error during genome replication (likely due to the high error rates of RDRPs and the lack of proof reading enzymes) different mutations occur during each genome replication resulting in a mixed population of viruses [102]. It is believed that the presence of a large and varied population allows for quick adaptation of

arboviruses to the disperate vertebrate and invertebrate hosts during their transmission cycle [102-103]. The addition of genetic material to the genome of the virus impairs the ability to replicate, and this additional material will be under strong pressure to be removed. Eventually this additional material will be lost in the virus population after a certain number of generations (as demonstrated by a decrease in mammalian cells infected with 5'ONNVic-eGFP expressing GFP from 55% in a first passage to less than 5% by a fifth passage [42]). In order to limit the possibility of genetic changes to the virus, such as deletion of the GFP cassette, or mutations in viral genes, the virus was only passaged twice before use in experiments. However, once cells are infected, several generations of viral replication may occur before all cells are infected, during which the GFP cassette may be lost. Thus cells may be infected with an infectious clone of ONNV that no longer expresses GFP. Carrying out immunofluorescence assays (IFAs) using an antibody against one of the highly conserved envelope proteins would allow distinction between uninfected cells and those infected with an infectious clone that no longer expresses GFP.

Infection of VERO cells with 5'ONNVic-eGFP leads to extensive cell death, presumably by apoptosis, which is typical of Alphavirus infection in mammalian cells[33], as well as a variety of other families of viruses, including the herpesviruses, adenoviruses, poxviruses, baculoviruses and many others [104]. In stark contrast to infection of mammalian cells, infection of *A. gambiae* derived cell lines resulted in no morphological pathology, although in the literature there have been reports of cytopathic infection caused by arboviruses in mosquitoes [105-109] and in invertebrate cell lines infected with certain arboviruses [39]. It is not clear whether the overall lack of apoptosis in arbovirus infected cells is due to inhibition of apoptosis by the virus, avoidance of detection by host cells, or whether host invertebrate cells do not use apoptosis as a defence mechanism against infection. Indeed, using SIN ATSs expressing inhibitors and activators of apoptosis, it has been demonstrated that triggering apoptosis in infected cells does not reduce the initial acute phase of virus production in the *A.aegypti* derived C6/36 cell line, although subsequent virus production was reduced as the cells died [38]. Current evidence

suggests that apoptosis may be triggered by high levels of Arbovirus infection, but is generally not active during persistent infection[110].

Higher titres of virus are generated during the initial acute stage of infection of invertebrate cells compared to mammalian cells ($5X10^8$ in L35 cells compared to an average of ~$6X10^7$ in mammalian cells). This is most likely due to the rapid cell death that occurs in mammalian cells following infection, and thus fewer cells producing progeny virus during the acute phase of infection.

Infection of *A. gambiae* cells follows the typical pattern of Alphaviruses in invertebrate cells [110], with an initial acute phase of infection between 12 and 72h, where production of infectious clones peaks. After this peak infection levels begin to fall, consistent with the observation that alphaviruses form a lifelong but low level infection in invertebrate cells [11, 110]. A single replication cycle of 5'ONNVic-eGFP takes between 12 and 24h, where the first virus is released from cells. The expression of GFP clearly shows that the infection spreads from a cell to neighbouring cell, forming patches of infection, probably arising from a single infected cell. This indicates that the virus can spread more easily between neighbouring cells, possibly through the contacts formed between the cells such as gap juctions, although this remains to be proven.

Correlation between GFP expression and production of infectious clones is consistent until 96 hours, when GFP expression persists but infectious clone production has slowed. Thus GFP is a good marker in cells for early infection, but it gives you a 'yes or no' answer for infection, and does not necessarily reflect the level of virus being produced in infected cells i.e. is a qualitative but not a quantitative marker of infection. If components from the subgenomic promoter i.e. the structural proteins and in this case GFP, are the limiting factors in production of virus, then we might expect cumulative GFP expression to be roughly proportional to the amount of new virus being produced. If the structural proteins are produced in excess, and it is the number of replicated genomes that is the limiting factor, then the GFP may not correspond to the amount of new virus being produced.

Additionally GFP will remain within the cytoplasm of an infected cell after viral replication has slowed or ceased, until the protein is degraded. Considering that the genome is the likely target of RNAi through dsRNA replication intermediates or secondary structure, we might expect the limiting factor to be the number of replicated genomes available to produce new clones. Qrt-PCR analysis of viral RNA genome copy number gives a more accurate and quantitative measure of viral infection in *A. gambiae* cells. Although the GFP marker cannot be used as a measure of viral production within cells, GFP can still be used to identify genes involved in viral infection, such as genes required for invasion.

4.7.2 Infection of adult *A. gambiae* mosquitoes – intrathoracic innoculation

Initially two different strains of *A. gambiae* were infected with virus passaged through mammalian (VERO) or insect (L35) cells. The highly laboratory adapted strain of *A. gambiae*, G3, showed the high levels of infection compared to a relatively recently colonised strain, Yaounde. Infectivity of mosquitoes was severely impaired by passaging virus through *A. gambiae* cells prior to infection, consistent with previous observations (Dr B Foy, personal communication). The reason for such a large difference in infectivity is not known. One major difference in post translational processing of virus envelope proteins between invertebrate and vertebrate cells is glycosylation. In SINV, both envelope protein E1 and E2 have two glycosylation sites [111]. During SINV infection of insect cells, all glycosylation sites have high-mannose sugars due to the absence of n-acetylglucosaminyl-, galactosyl-, and sialyltransferases [111]. In vertebrate cells a combination of complex oligosacaccharides and high-manose sugars are used, depending on the accessibility of the site to processing enzymes in the golgi apparatus [111]. Mammalian-cell derived and mosquito-cell derived RRV has been shown to differentially induce type I interferons in mammalian cells. Mosquito-cell derived viruses are poor inducers of IFN alpha/beta, whereas mammalian cell derived viruses are potent inducers of IFN alpha/beta [112]. This differential induction is attributable to the differences in N-linked glycosylation resulting from passage through mosquito cells [112-113].

Whether the complex oligosaccharides present in mammalian-cell derived virus are required for binding to a receptor and thus infection, or whether the high-mannose sugars from mosquito-cell derived virus trigger anti-viral immune responses in the mosquito and so prevent or clear infection is not clear, but may explain the differences in infectivity observed.

The dynamics of 5'ONNVic-eGFP infection of adult G3 mosquitoes follows a similar trend as that seen in the L35 cell line, with an initial acute phase of infection peaking at days 6-8, followed by a decline in production of virus by day 9. The cause of the switch from initial acute phase of infection where large amounts of virus are released from infected cells, to the persistent phase of infection, where small amounts of virus are released from infected cells, or a small number of cells release virus, is not known. This switch could be mediated either by the virus, in order to limit damage to the vector and establish a lifelong infection ensuring transmission to as many other hosts as possible, or could be mediated by the host i.e. immune reactions that limit viral infection may be switched on. In cultured *A.albopictus* cells infected with SINV or SFV, a polypeptide factor is secreted that downregulates production of virus [11, 114]. This factor is mosquito and virus specific i.e. it only effects mosquito cells, and the factor secreted in response to SINV is ineffective for SFV infection and vice versa [11]. Uninfected cells treated with the secreted factor arrest cell division, but subsequently recover and are resistant to virus infection [114]. They also constitutively secrete the factor, and another protein that is associated with lysosomes [115]. It is possible that in mosquitoes, where virus is typically downregulated at 7 dpi, a similar mechanism is in place [11]. In the case of SFV, an alternative model suggests that the virus is responsible for the switch, through the mechanism it employs to prevent 'superinfection' of cells [116]. Early after infection nsP4 and P123, formed from cis-cleavage of P1234, function to synthesise minus strand RNA, later further trans-cleavage forms nsP1, P23 and nsP4, capable of both plus and minus strand synthesis. A final cleavage to form the 4 nsP's abolishes the ability to generate minus strand RNA, and due to the presence of high concentrations of trans-acting protease at this stage, new P123 is not generated [11, 116]. At this stage the virus can continue to produce progeny from already

established replicase complexes, but no new replicase complexes can be formed. Thus, viral production is dramatically reduced, preventing 'superinfection' of cells [116]. It is suggested that eventually the non-structural proteins and trans-acting protease will be degraded, removing the inhibition of minus strand synthesis, and hense the cell is once again susceptible to infection, explaining how low levels of infection continue persistently in infected vertebrates [116]. Whether either of these mechanisms is present in ONNV infection of *A. gambiae* is unknown.

The spread of 5'ONNVic-eGFP infection in intrathoracically inoculated *A. gambiae* mosquitoes is slow compared to other vector-virus combinations. In fact, when ONNV is injected into Culicine mosquitoes it rapidly replicates and spread throughout the mosquito host [80]. It has been suggested that the robust RNAi response observed in *A. gambiae* mosquitoes [117] may contribute to the poor vectorial capacity of these mosquitoes in comparison with other typical vector-virus combinations [80].

Tissue tropism of infection after intrathoracic inoculation was similar to that observed previously using this vector-virus combination [80], with the head tissues most commonly showing GFP expression, however this may be due to the reflective quality of the ommatidia in the eye of the mosquito, reflecting light out as well as into the eye and thus allowing clear observation of GFP. When observing whole mosquitoes, GFP expression through the cuticle was more difficult to see, coupled with the autofluorescent quality of the cuticle itself. As such, it can be difficult to discern whether a mosquito is infected or not based on the presence of GFP expression. The more reliable and accurate quantitative real-time PCR method or standard plaque assay were used subsequently to quantify infection in mosquitoes, giving more accurate results, but also a quantitative result rather than the qualitative results obtained from observing GFP expression alone. Although tissue tropism in intrathoracically inoculated mosquitoes was not observed in detail (as previous studies have done so [42]), interestingly either nerve or muscle cells surrounding the gut were often seen to be infected, as observed in other vector-virus combinations; Rift Valley Fever (RVF) infects skeletal and visceral muscles, SIN infects visceral

muscles in *Ae. albopictus* [118] and *Ae. Aegypti* [119] and WNV infects the contiguous muscles of the posterior and anterior midgut in *Culex* mosquitoes [106, 118]. It is possible that infection of nervous or muscle tissue of the gut may provide a route of dissemination from the gut after a blood meal.

4.7.3 Infection of adult *A. gambiae* mosquitoes – *Per Os* infection

Patterns of viral infection following an artificial blood meal show limited and restricted tropism in comparison to many other virus-vector combinations. Viral infection was almost exclusively seen in the cardia and anterior gut tissues, with occasional midgut infection in a small number of individuals (7% at D3 dropping to only 4% at D6 post blood feeding). This restricted tissue tropism is similar to those observed in ONNV infections of *A. gambiae* in the past [42]. In comparison, midgut infection rates in *A.aegypti* infected with DENV are 87-90% (in three different *A.aegypti* strains)[120] and infection is spread throughout the midgut by 10 dpi despite a lower infectious loads in the blood meal than used for 5'ONNVic-eGFP. SIN infection of *A.aegypti* leads to large patches of infected midgut epithial cells and midgut musculature [119-120]. EEEV infection of *Culiseta melanura* (the virus' enzootic vector) results in rapid infection of the midgut [121]. Several other viruses closely related to ONNV cause broad midgut infection in their natural vectors, including VEEV, SIN in Culex pipiens, and WEEV [42]. Although not carried out in this study, dissemination rates for ONNV infection in *A. gambiae* have also been shown to be low compared to other vector-virus combinations [42]. This indicates that *A. gambiae* should be a poor vector of ONNV, and may not be the natural vector of the disease outside of epidemics. This is supported by the findings of ONNV positive mosquitoes only during epidemics of the disease [23]. *A.funestus* mosquitoes that have been found positive for ONNV infection and are postulated to be the main vector of the virus [25, 122].

4.7.4 Comparison of 5'ONNVic-eGFP infection in three freshwater *Anopheles* species

A comparison of 5'ONNVic-eGFP infection in *A. gambiae* was carried out with two other fresh water Anopheles species; *A. arabiensis* and *A.quadiannulatus*. *A. arabiensis* shares a similar geographical distribution to *A. gambiae*, spread across the whole of central Africa from Senegal through to Sudan and Ethiopia, and further south throughout Kenya, Tanzania, Mosambique, and the Democratic Republic of the Congo [123]. *A. arabiensis* is additionally found through Angola and Zambia, where *A. gambiae* is less prevalent. In contast, *A. quadriannulatus* is restricted to two zones; one in the highlands of Ethiopia (called *A. quadriannulatus* B), and the other in South Africa, Botswana and Zimbabwe (called *A. quadriannulatus* A) where the weather is more temperate than other parts of Africa [123]. It is believed that *A. arabiensis* is the most ancient of the three spp, from which an ancestral *A. quadriannulatus* evolved. *A. quadriannulatus* A, *A. quadriannulatus* B and *A. gambiae* speciated from this ancentral form. *A. gambiae* and *A. arabiensis* are both important vectors of human malaria in sub-Saharan Africa, however *A. quadriannulatus* are not known to be vectors of malaria in natural populations, indeed no malaria parasites have ever been found in field caught female mosquitoes [63]. In laboratory experiments with *A. quadriannulatus* and the rodent model of malaria, *Plasmodium berghei*, invading parasites crossing the midgut are frequently melanised, with a small proportion forming oocysts. This partial refractoriness to infection is heritable and dominant. Silencing key components of the mosquito immune system (LRIM1, APL1C and TEP1) allows parasites to develop normally in *A. quadriannulatus*, effectively turning the mosquito into a competent vector [63]. It is speculated that through repeated exposure to malaria parasites in *A. gambiae* and *A. arabiensis* spp may have led to an evolutionary co-adaptation between mosquito immune responses and the parasite, however *A. quadriannulatus*, through limited geographical range and a more zoophillic lifestyle retained an ancestral refractory phenotype, based on innate immune responses [63]. *A. gambiae* and *A. arabiensis* are relatively refractory to 5'ONNVic-eGFP infection compared to *A. quadriannulatus*, with viral titres at 6 days post infection of 20 times higher in *A. quadriannulatus* compared to *A. arabiensis*, and 6 times higher in *A. quadriannulatus* compared to *A. gambiae*. The distribution of ONNV epidemics reported to date occurred in areas where both *A. gambiae* and *A. arabiensis* are prevalent, however

A. quadriannulatus are scarce. It is possible that the repeated exposure of *A. gambiae* and *A. arabiensis* has resulted in the evolution of immune mechanisms to cope with viral infection, limiting the ability of the virus to replicate, whereas in *A. quadriannulatus* these immune mechanisms have not evolved, allowing the virus to replicate to high levels in mosquito tissues. Current evidence in insects indicates that when high levels of virus are present, exceeding a threshold, apoptosis is induced [110]. This could explain the death induced by viral infection; by 6 dpi, around 60% of infected *A. quadriannulatus* mosquitoes had died. The cause of death was not determined; however studies looking at differences in gene copy number of the whole genome between *A. gambiae* and *A. quadriannulatus* show large differences in apoptotic related genes (Tibebu Habtewold, unpublished data). It would be interesting to investigate further whether apoptosis induced by viral infection may be responsible for the shortened life span observed in *A. quadriannulatus* mosquitoes.

It is also noteworthy that the cell line susceptible to 5′ONNVic-eGFP infection, L35, is derived from a genetically selected Plasmodium refractory line of *A. gambiae* mosquitoes that, similarly to *A. quadriannulatus*, melanise all invading *Plasmodium berghei* parasites as they cross the midgut. There may be a link between genetic refractoriness to one pathogen, and genetic susceptibility to another, perhaps rooted in the innate immune system.

4.7.5 Final comments

In most arbovirus-mosquito vector combinations viral infection leads to widespread infection of multiple tissues, including the midgut, fat body and, importantly for transmission, the salivary glands of the infected mosquito. Although infection of *A. gambiae* with 5′ONNVic-eGFP does lead to infection, tissue tropism is restricted in adults infected both intrathoracially and through infectious blood meals. Infection of a variety of *A. gambiae* cell lines similarly shows limited infection in the majority of cell lines tested. This indicates that *A. gambiae* is a poor vector of ONNV, and probably only functions as a vector of the virus during epidemics of the disease.

The use of GFP as a marker of infection in this model system has limitations, particularly in reference to infection of adult mosquitoes. GFP is a good marker of early infection in the susceptible L35 cell line, however GFP expression does not correspond to generation of new virions after the acute phase of infection. The GFP marker could be used to screen genes for functions in initial infection of cells, however for quantitative data measuring the titre of the virus directly would give more accurate results. In adult mosquitoes an additional problem lies in the innate auto-fluorescence and thickness of the mosquito cuticle. GFP expression is difficult to see and is not an accurate marker for infection. Again, direct titring of the virus provides more accurate quantitative data.

This unusual vector virus combination offers an interesting opportunity to study the interactions between a mosquito and virus that often leads to clearance of viral infection, possibly mediated by the robust immune system of *A. gambiae* mosquitoes. Investigating these interactions may give insight into the genetic determinants of vectorial capacity in mosquitoes.

5 *A. gambiae* transcriptional responses to viral infection, and identification of genes regulating viral infection in adult mosquitoes

5.1 Introduction

In order to identify components of an organism's immune system, there are several approaches that can be adopted. Firstly a literature based approach can be taken, where already characterised genes from model systems; typically *D. melanogaster* for insects, can be hypothesised to have similar functions in closely related systems. Using this approach narrows the window of focus to a few distinct immune signalling pathways with, in this case, known roles in anti-viral immunity in various insects. Although focusing studies to more likely relevant genes or gene families, this approach will miss any species specific adaptation to the pathogens that they will encounter. The haematophagous lifestyle of *A. gambiae* confers exposure to a wide variety pathogens that *D. melanogaster* does not encounter, as such we might expect the immune responses to these disparate groups of pathogens to be quite different. Using a genome wide forward genetics approach to study pathogen-vector interactions can reveal genes with important functions that are not present or utilised in other model systems. A good example of this is the leucine rich repeat immune (LRIM) family of proteins identified in *A. gambiae*. This family of genes was discovered through microarray analysis of *A. gambiae* responses to *Plasmodium* and bacterial infection [124-125]. LRIM1 is a potent inhibitor of *Plasmodium berghei* development in *A. gambiae* mosquitoes [62, 126]. It belongs to a family of similar proteins that are mosquito specific and are not found in *D. melanogaster*. Subsequently it has been seen that other members of this family are also potent antagonists of *P. berghei* development [127-128]. Without using genome wide approaches to study vector-pathogen interactions, such gene families may not be discovered.

Microarray technology allows researchers to study the transcriptional profiles of entire genomes, generating huge datasets containing vast amounts of biological information. Although very powerful, microarray experiments must be carefully planned and tightly controlled to avoid generating false positive data. During this project a full genome microarray using Agilent technology was used. In order to

begin identifying *A. gambiae* genes involved in anti-viral immunity, the global transcriptional responses of mosquitoes to ONNV infection were investigated.

5.2 Transcriptional response to 5'ONNVic-eGFP infection: experimental design

A microarray study carried out by previous members of the laboratory, in collaboration with the University of Notre Dame, transcriptionally profiled *A. gambiae* mosquitoes infected with ONNV using a spotted amplicon microarray encompassing only a small number of genes [129]. With the rapid improvement of microarray technology, and the significantly improved gene predictions from the *A. gambiae* genome on which a genome wide microarray chip was developed, repeating the transcriptional analysis was necessary to gain a global view of *A. gambiae* responses to viral infection. In this study a microarray platform was developed by colleagues in the laboratory using Agilent technology. This microarray platform covered the full genome of *A. gambiae* based on the PEST strain, assembly version AgamP3, gene build Agam P3.3 (http://www.vectorbase.org/Help/AgamP3.3_annotation_metrics). The probes were designed using Agilent's eArray online service. In brief, the transcript sequences of all predicted *A. gambiae* genes were uploaded into Aglient eArray, and 60 base pair (bp) probes with similar TMs were designed. Each probe was duplicated and placed randomly on the array to avoid spatial effects such as drying out/scratches on the array. The predicted immunity gene transcripts were split into 10 sequences and a single probe was designed for each of the 10 sequences, giving 10 unique probes for each immunity gene in addition to the probes designed for every predicted gene. A large number of controls for both Cy-3 and Cy-5 labelled probe hybridisations were included. All probes were randomly spotted on the array. All probe design was carried out by Dr Amanda Jackson and Seth Redmond.

Based on the dynamics of 5'ONNVic-eGFP infection in *A. gambiae* mosquitoes three time points were selected for transcriptional analysis (figure 19): Day 1 post infection (dpi) represents the early stages of infection, when the first cells become infected; 4 dpi represents the exponential phase of infection, where viral infection is

spreading from cell to cell, and tissue to tissue; 9 dpi represent the persistent lower level of infection established by alphaviruses in their mosquito host.

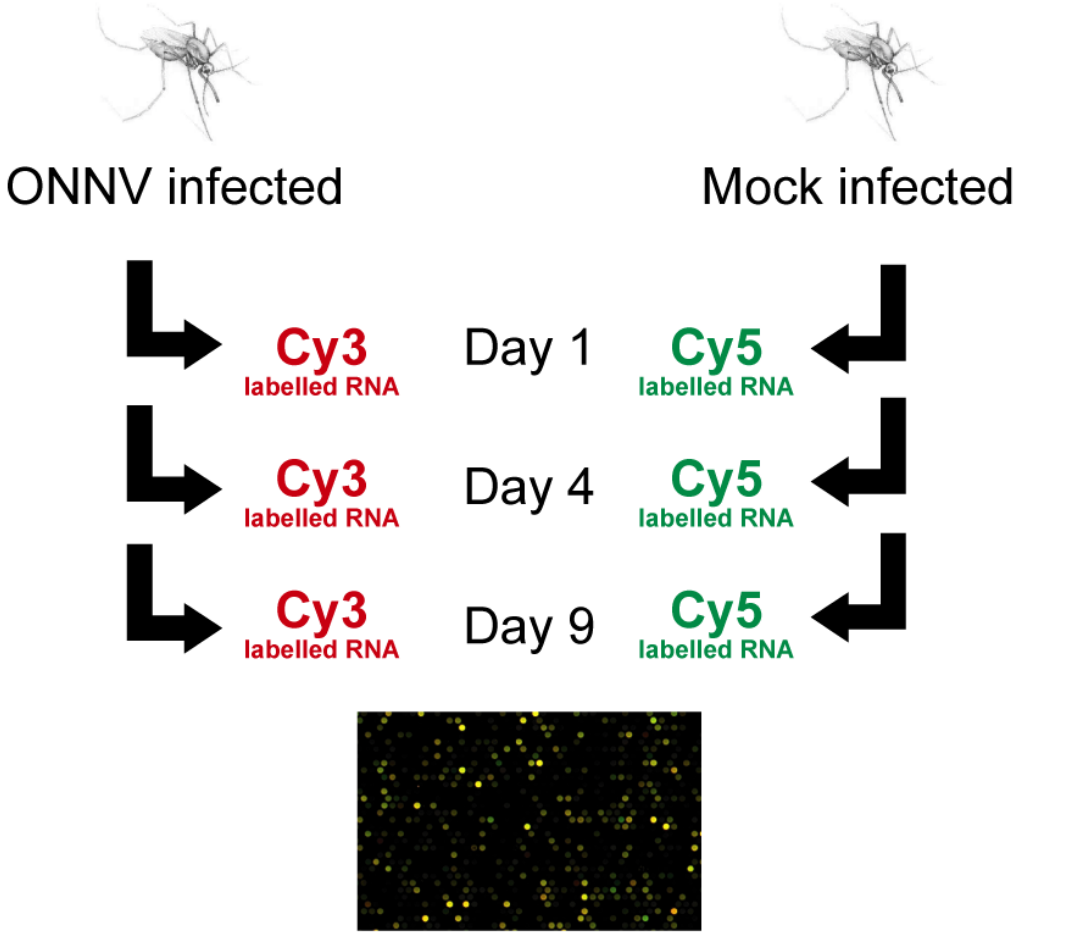


Figure 19. Experimental design for transcriptional analysis of 5'ONNVic-eGFP infection. Total RNA from infected (~1640 PFU/mosquito P1/P2V 5'ONNVic-eGFP in 69 nl) and mock infected (injected with 69nl of conditioned media from mock transfected Vero cells) mosquitoes was extracted at 3 time points after infection – 1, 4 and 9 dpi. RNA from infected mosquitoes was labelled with Cy-3 and mock infected with Cy-5 nucleotide analogs. Equal amounts of RNA from each were hybridised on an aglient 4 X 44K array platform.

5.3 Transcriptional response to 5'ONNVic-eGFP infection: experimental procedure

Mosquitoes were intrathoracically inoculated with the virus to ensure dissemination of the virus to all mosquito tissues; mosquitoes were injected with either P1/P2V

5′ONNVic-eGFP (~1640 PFU/mosquito) or with conditioned medium generated by transfecting LacZ RNA into VERO cells, and 'passaging' in an identical manner to passaging virus (P1/P2V LacZ). Three biological replicates were carried out, with each replicate using G3 strain *A. gambiae* that have hatched from the same batch of eggs and reared in identical conditions. Each cohort of mosquitoes was injected at the same age, and kept in identical conditions for the duration of the experiment. Mosquitoes were *not* screened for GFP expression prior to total RNA extraction to prevent selection of susceptible individuals. This microarray study aimed to investigate the immune responses of *A. gambiae* to infection, including those capable of clearing viral infection. At each selected time point, 10 mosquitoes were collected and homogenised in 500µl TRIzol reagent. Total RNA was extracted, and assessed for quality using gel electrophoresis and spectophotromotry (section 3.11). Using the Agilent low RNA imput kit (Agilent), RNA from infected and control mosquitoes was amplified and labelled with Cy-3 and Cy-5 nucleotide analogues respectively. Equal quantities of labelled RNA were mixed and hybridised to the Agilent 4X44K microarray (section 3.15.1). Hybridised slides were washed to remove unbound labelled RNA (section 3.15.2) and were scanned using an GenePix semiconfocal microarray scanner (AXON Instruments, Foster city, CA).

5.4 Analysis of transcriptional responses

Candidate gene lists were generated as follows (summarised in figure 20); Gene Pix Pro 6.1 was used to record feature signal intensity, to eliminate local backgrounds, for grid alignment and manual inspection of feature quality. Average feature diameter was calculated and features lying outside two standard deviations of the mean were excluded from analysis. The ratios of feature intensity versus local and global backgrounds were calculated and features not exceeding background intensities were excluded from analysis. Features were normalised using Genespring 6.1 (AXON Instruments, Foster city, CA) by locally weighted linear regression (Lowess) methods. Feature intensities over the three biological replicates were averaged. T-test p-values were calculated, and normalised data was filtered to exclude data with p-values greater than 0.05. Data was further filtered to include

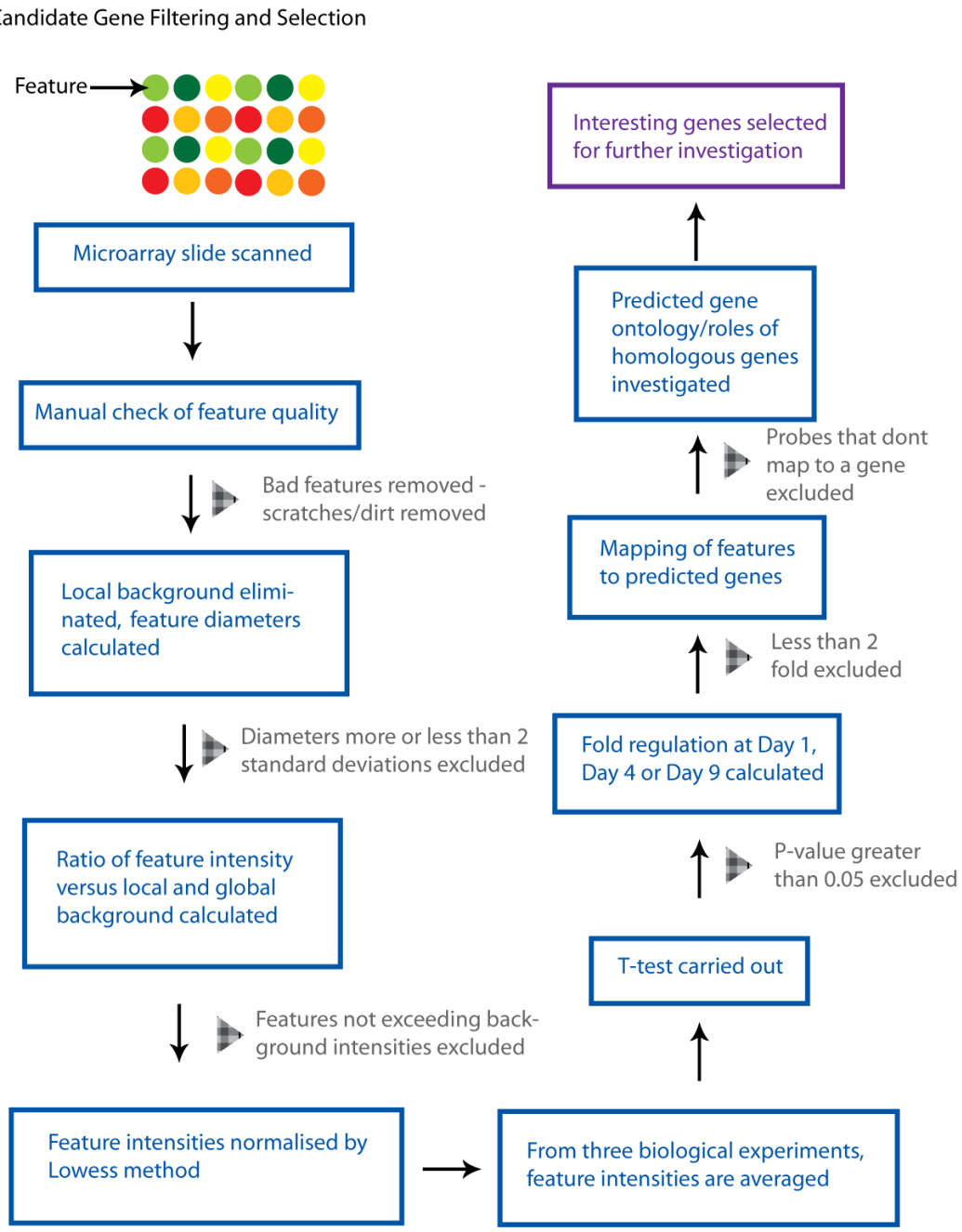


Figure 20. Microarray feature filtering and candidate gene selection process. GenePixPro 6.0 was used to record feature intensities, and remove manually checked features. Local background elimination, feature diameters and filtering of ratio intensities to remove features not exceeding background intensities was performed in Micorsoft Exel. Genespring was used to normalise feature intensities, calculate T-tests, remove data with P values >0.05 and to calculate fold-change ratios. Biomart was used to identify GO terms and orthologous genes in other insects.

only genes showing 2-fold and greater regulation. For immune genes, where multiple probes for each gene are present, fold change ratios that pass all the described filters were averaged and given a single value.

5.5 Genome wide responses to 5′ONNVic-eGFP infection in *A. gambiae* mosquitoes

Transcriptional profiling of infected versus mock infected mosquitoes revealed a large number of genes that are differentially regulated (figure 21). At 1 dpi, 64 genes are differentially regulated; these genes represent the early onset genes, rapidly responding to virus injected into the hemocoel, to virus attaching to, and invading cells. At 4 dpi 214 genes are differentially regulated. This is during the acute phase of infection where viruses are released from infected cells and spread throughout the carcass of the mosquito. By 9 dpi a dramatic drop in the number of differentially regulated genes is observed from 214 to 23. These genes are regulated during the persistent phase of infection. A full list of regulated genes can be found in Table1 Appendix 1.

Clustering the differentially regulated genes by time point allows us to see how the temporal pattern of gene expression changes during infection. Genes can be divided into 6 catagories: early onset genes (those differentially regulated only at D1); early to mid onset genes (those regulated at D1 and D4); mid onset genes (regulated at D4 only); mid to late onset genes (those regulated at D4 and D9); late onset genes (D9 only); broadly responsive genes (those regulated at D1, D4 and D9).

5.5.1 Early onset genes

There are 23 early onset genes that respond to 5′ONNVic-eGFP infection (Table 1 Appendix 1). A cluster of 3 guanine nucleotide binding proteins (G-proteins, with functions in signal transduction) are down-regulated. There are 4 up-regulated genes with putative immune functions; 3 are pathogen recognition/immune

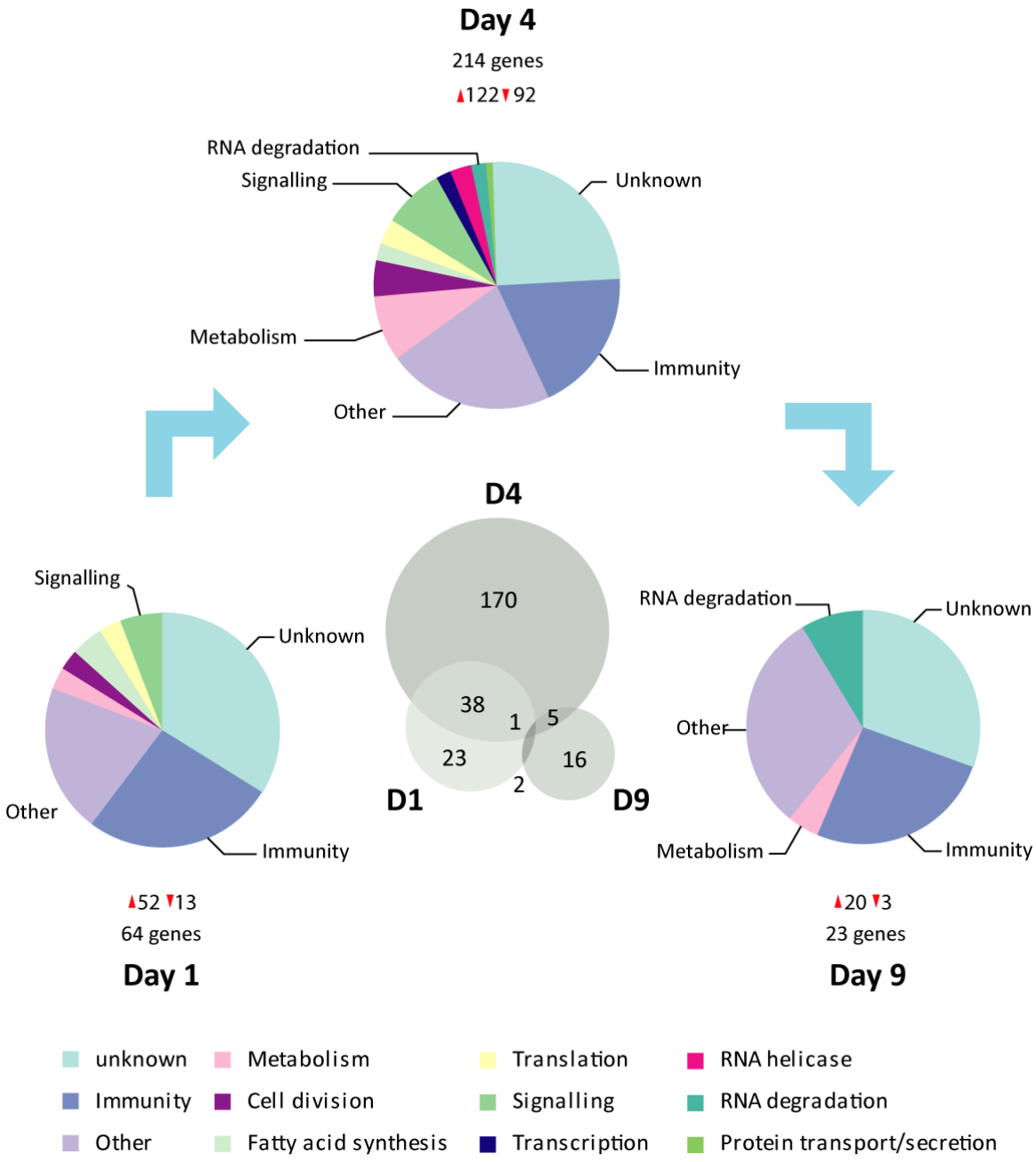


Figure 21. Global transcriptional responses to 5'ONNVic-eGFP infection. The transcriptional responses of *A. gambiae* mosquitoes to 5'ONNVic-eGFP infection were profiled using 4X44K Agilent RNA microarrays. Gene lists include only features that pass strict criteria outlined in Figure 20. Genes included the analysis are 2-fold or greater regulated at a minimum of 1 of the 3 time points, with T-test P values of <0.05. Genes were categorised based on gene ontology and orthologs in other insects.

signalling activation proteins (ML9, LRIM1 and TEP5), and 1 is a peroxidise (GPXH3) with a putative function in pathogen clearance. A further possible immune related gene is down-regulated; the orthologs of AGAP005901 from *D. melanogaster* to humans have putative functions in innate immune signalling, specifically in the regulation of apoptosis[130-133]. Of the 16 up-regulated genes, 9 are genes with no putative function or predicted gene ontology (GO), including a gene cluster (AGAP003773/5/7/8). Other up-regulated genes have diverse functions in metabolism, fatty acid synthesis & translation, and structure (cytoskeletal proteins).

5.5.2 Early to mid onset genes

Thirty eight genes are differentially regulated at both D1 and D4 post infection, the majority of which are up-regulated. Of the 32 up-regulated genes, a significant proportion is immune related: 13 putative immune genes with diverse functions from pathogen recognition, immune signal modulation and pathogen clearance are described in detail later. Nine genes with no known function or predicted gene architecture are up-regulated. Two fatty acid synthesis related proteins are up-regulated (a fatty acid synthase, and a fatty acid elongation protein). Further up-regulated genes include: a laminin B-like gene (the ortholog of an extracellular matrix protein, which regulates growth factor-like signalling pathways in *C.elegans* and humans); two insect odorant binding proteins; genes associated with nitrogen compound metabolism and carboxylesterase type B activity.

Two genes associated with cell division are down-regulated: a putative cell cycle checkpoint kinase and a gene with cell cycle associated roles in *D. melanogaster*. Additionally a matrix metalloproteinase (MMP) is down-regulated, ortholog of MMP1 in *D. melanogaster*, which has diverse functions in cell proliferation, apoptosis and immune defence. Further down-regulated genes include a translation initiation factor-like gene, an actin/microtubule binding protein, and a protein with no predicted functional domains.

5.5.3 Mid onset genes

The mid onset genes are by far the largest cluster regulated by 5'ONNVic-eGFP infection. 170 genes are differentially regulated, and unlike the other clusters where the majority of genes are up-regulated, 52% are down-regulated. Some striking groups of genes with similar functions respond to ONNV infection during the acute phase of infection. Twenty seven putative immunity genes are differentially regulated (described in detail in section 5.5.6). Four genes associated with RNA degradation are down-regulated: Dicer-1; the ortholog of *D. melanogaster* P-element induced wimpy testis (PIWI); a gene involved in RNA metabolism; TSN, a member of the RNAi pathway. Additionally 6 RNA helicases are down-regulated. Eight genes associated with cell division are down-regulated, and 13 genes with roles in cellular signalling are down-regulated: 4 transcription factors (including the general transcription factor TFIIH); 2 genes implicated in WNT signalling in *D. melanogaster (*a negative regulator of the WNT pathway, and a gene thought to phosphorylate dishevelled in the WNT pathway); 2 protein phosphatases (1 tyrosine and 1 serine/threonine specific); a PI-4 kinase (with typically diverse roles in signalling); a GTPase activator; a GTP binding protein; a TGF-beta superfamily protein and an ortholog of a *D. melanogaster* regulator of transcription in the decapentaplegic (DPP) pathway (both of which have diverse putative functions in cell growth and differentiation).

Nine genes associated with translation and transcription are down-regulated including: three regulators of splicing; a tRNA synthesis enzyme; an enzyme that catalyses the addition of tRNA to growing amino acid chains; a translation initiation factor; an ortholog of the RNA polymerase 1 specific transcription initiation factor RRN3; an inhibitor of transcription. A histone acetyltransferase (a promoter of transcription) is up-regulated.

Forty genes have no known putative function and/or no predicted protein domain architecture. Thirty nine genes have diverse functions in many aspects of cell

biology. There are 13 genes associated with metabolism, and as seen in the early to mid onset cluster of genes, 3 more genes associated with fatty acid synthesis.

5.5.4 Mid to late onset genes

Only 4 genes are differentially regulated at both 4 and 9 dpi, demonstrating a large switch in gene regulation between D4, where a large number of genes are regulated, and D9, where a small number of genes are regulated. The change in gene expression differentiation follows the dynamics of viral infection, where by 9 dpi viral infection is persistent but at a low level. A single putative immunity gene (an LRR gene), a thio-reductase, a protein with unknown function and a fumerate lyase are all up-regulated at this stage.

5.5.5 Late onset genes

Fifteen genes are differentially expressed only at the late phase of infection – a dramatic decrease from the 170 expressed at D4. The majority of genes are up-regulated including 4 genes with no putative function and/or predicted gene architecture, 3 immunity genes (described in detail in section 5.5.6), and interestingly 2 genes associated with RNA degradation. The first contains a PIWI and a PAZ domain (similarly to PIWI and several genes associated with RNAi) and the second is a closely related to TSN (a member of the RNAi pathway).

The expression of 3 genes peaked during early and late infection; 2 with no predicted gene architecture, and a putative immunity gene (FREP44). Only 1 gene was found to be up-regulated throughout the time course (Lysozyme C4).

Qrt-PCR was used to independently confirm the differential regulation of 4 genes found to be virally responsive (figure 22).

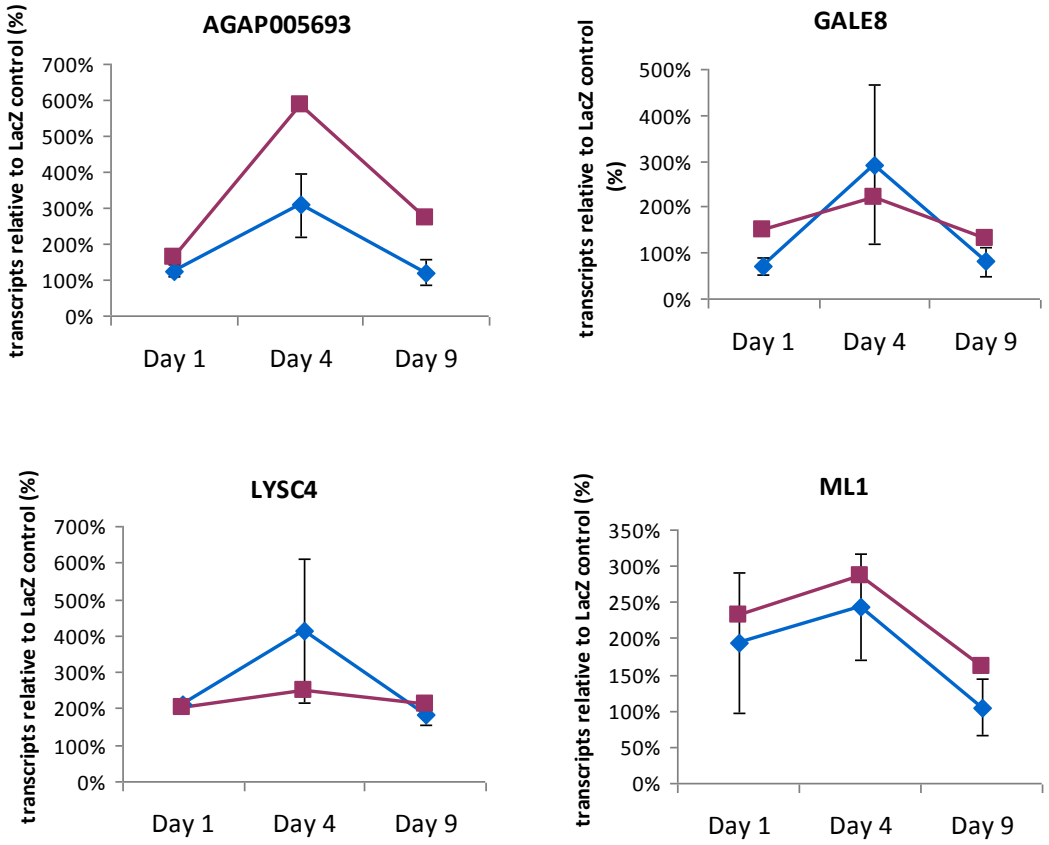


Figure 22. Qrt-PCR confirmation of virally responsive genes. The expression of 4 virally responsive genes ascertained by microarray analysis (red) was confirmed using qrt-PCR (blue). cDNA was generated from RNA extracted from 5'ONNVic-eGFP infected and mock infected (LacZ control) mosquitoes. Transcript levels are expressed as a % of those observed in the LacZ control. Error bars represent standard deviation of 3 biological replicates.

5.5.6 Immune genes regulated by 5'ONNVic-eGFP infection

In addition to the probes designed against all predicted genes in the *A. gambiae* genome, the array platform used during this project encompassed an additional 10 probes for each predicted immunity gene. During analysis, data was rejected according to the filtering criteria outlined in figure 20. The majority of this data is rejected due to high levels of variation between biological replicates. As a result, data from all 10 probes corresponding to each immunity frequency did not pass the filtering criteria, giving data for only a few of these 10 probes. For each immunity gene where at least 1 probe passed the filtering criteria, in order to add confidence

that the probe was showing true differential regulation, the whole probe set was plotted including data that did not pass the T-test filter (i.e. data that had passed all filtering criteria except filtering on P value) (Appendix 1, figure S1). These plots gave two valuable insights into the data that have passed all the filtering criteria: firstly the trend of expression over the whole time course was seen; secondly, using the multiple probes generated for immunity genes, the trend of multiple probes corresponding to the same gene was visualised. The plots, although containing data that is statistically inaccurate, gave an increase or decrease in confidence of the data that passed the filtering criteria.

Four examples are given in figure 23. ML1 expression shows a trend of up-regulation at 1 dpi, followed by a peak of expression at 4 dpi reducing by 9 dpi. Four of the 10 probes designed for this gene did not generate data, most likely due to a combination of poor sequence from which the probes were designed and/or low intensity data. Five of the 6 probes show the same trend of expression, demonstrating that there is an overall increase in ML1 expression during early to mid infection. GALE8 expression shows a peak at 4 dpi. For this gene, nine probes have generated data. Agreement between the probes is better than that for ML1 expression, and the data demonstrate that GALE8 expression is up-regulated during mid infection. LYSC4 expression is up-regulated throughout the time course. For this gene only 5 of the 10 probes generate data, however the trend of expression is similar for these 5 probes, showing increased expression throughout the time course, peaking at 4 dpi. For LYSC6, only 2 of the 10 probes generate data. Both probes follow the same trend, with expression peaking at 4 dpi, however the change in expression level is lower for one probe than the other, reducing the confidence that this gene is up-regulated by ONNV infection. These four examples show how the inclusion of statistically in-accurate data can still be used to indicate where false positive data has been included in analysis. Appendix 1 figure S1 shows plots for all the immunity associated genes that were identified as being differentially regulated using the filtering criteria used for this study. Using these plots, several genes were removed as probable false positives; FREP10 (1 probe shows up-regulation at 4dpi and 3 probes show no differential regulation); SRPN5 (1 probe shows up-regulation

at 4 dpi and 5 show no differential regulation); CLIPB9 (2 probes show up-regulation at 4 dpi and 4 probes show no differential regulation). Two further genes, HPX6 and CTL5, show discrepancies between the probes sets, however it is not clear whether these genes are differentially regulation or not.

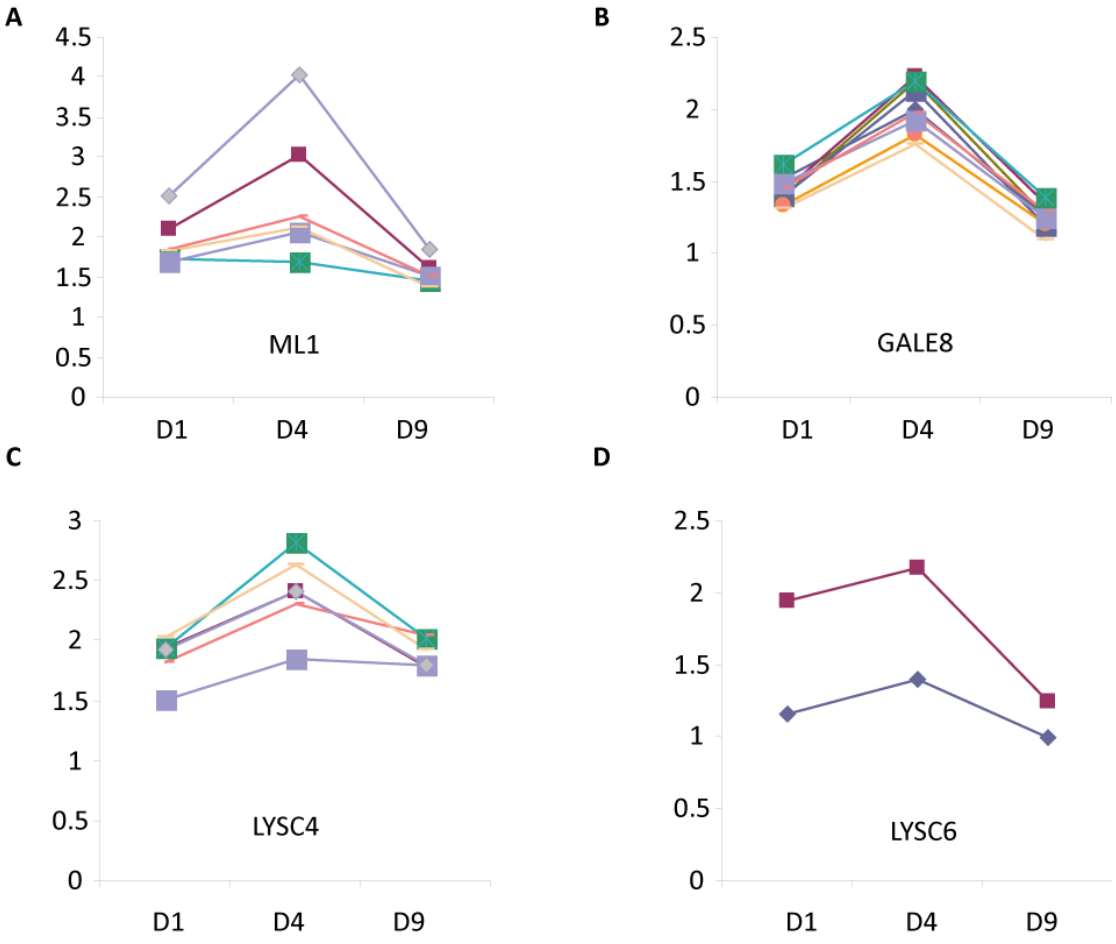


Figure 23. Relative gene expression plots of immunity genes. **A)** ML1 expression **B)** GALE8 expression **C)** LYSC4 expression and **D)** LYSC6 expression in ONNV infected mosquitoes expressed as fold change ratios compared to a mock infected control. Data has passed all the filtering criteria outlined in figure 20, excluding filtering on P value.

Plotting the data in this way also indicated whether the genes had been clustered into the appropriate temporal clusters. The majority of genes are clustered correctly, with a few exceptions. An example is LRIM1, which is shown to be up-regulated during early infection using the current filtering criteria. Data for LRIM1 expression at 9 dpi is rejected during the filtering due to inconsistency between replicates. If all the data is plotted including the highly variable results for 9dpi (see figure 24), the indication is that LRIM1 up-regulation is not only an early response to infection, but may also increase in expression during late infection. This observation would require confirmation using qrt-PCR.

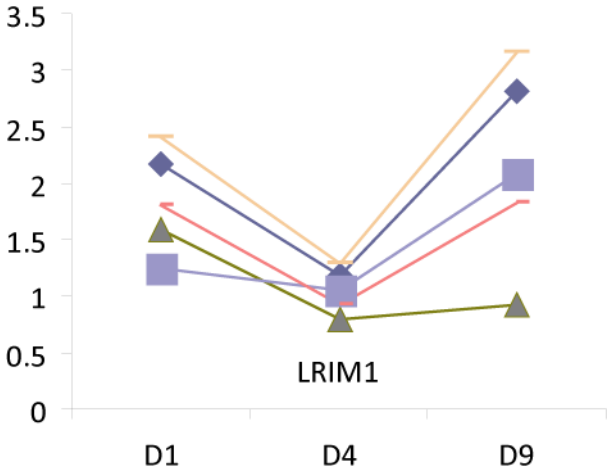


Figure 24. Gene expression plot for LRIM1. LRIM1 expression in ONNV infected mosquitoes plotted as a fold change ratio compared to a mock infected control. Data has passed all the filtering criteria outlined in figure 20, excluding filtering on P value. As such the data is not statistically accurate.

Infection of *A. gambiae* with 5'ONNVic-eGFP led to the differential regulation of 45 genes with putative functions in immunity. Clustering these genes based on their putative function within the 6 temporal clusters of gene expression (Table 6) shows a broad range of genes with function in all aspects of immune signalling including pathogen recognition, complement, immune signalling pathway components, immune signal modulation and effector genes.

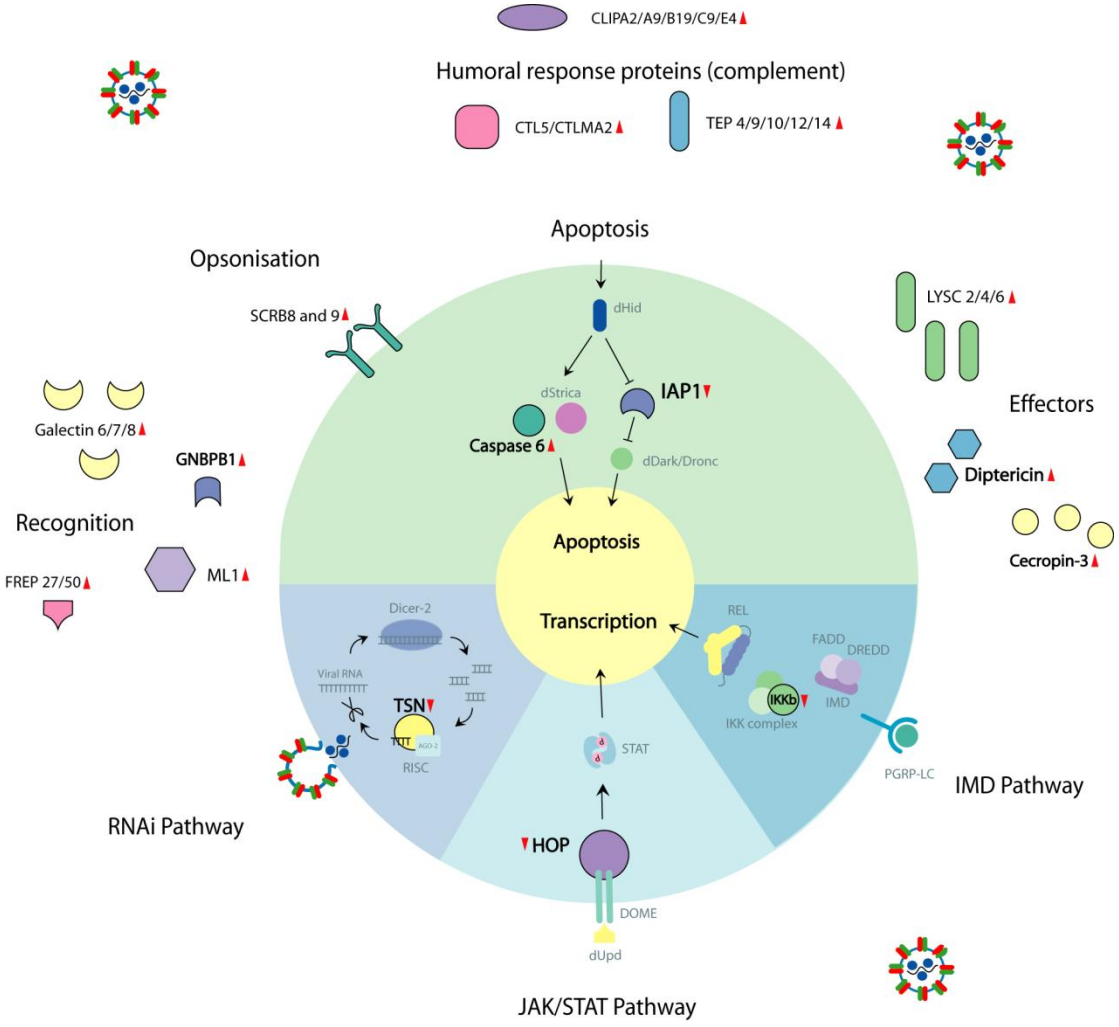


Figure 25. Immunity genes differentially regulated by 5'ONNVic-eGFP at 4 dpi. Genes differentially regulated by ONNV at 4 dpi are shown in black bold text, with black outlined cartoons. Other genes in signal transduction pathways are shown in grey text. Four major signalling pathways are represented; the RNAi pathway, the JAK/STAT pathway, the IMD pathway and the apoptotic pathway. Differentially regulated genes are grouped based on cellular location (intracellular/secreted) and putative function (recognition, humoral response proteins, opsoniation and effectors). Red arrows indicate the direction of differential regulation.

Table 6 Putative immunity genes regulated by 5'ONNVic-eGFP organised by function. Putative immune genes are grouped based on temporal expression, and functional category (Recognition, Complement, Signal modulation, Effector, Apotosis, IMD pathway, JAK/STAT pathway, RNAi pathway).

			Fold change ratio			
Gene	Functional group	Name	Day 1	Day 4	Day 9	putative function
Early onset						
AGAP010814	Complement	TEP5	2.604			Thioester containing protein, complement
AGAP006348	Complement	LRIM1	2.284			LRIM protein, complement [126]
AGAP004248	Effector	GPXH3	2.267			peroxidase, GPX sub familiy
AGAP002848	Recognition	ML9	2.006			MD2-like lipid recognition
AGAP007039	Recognition	LRIM4	2.676			LRIM protein, unknown function
Early to Mid onset						
AGAP010812	Complement	TEP4	2.876	2.48	1.611	Thioester containing protein, complement [134]
AGAP008654	Complement	TEP12	2.153	2.36		Thioester containing protein, complement
AGAP010819	Complement	TEP10	3	2.10		Thioester containing protein, complement
AGAP010830	Complement	TEP9	2	2.06		Thioester containing protein, complement
AGAP005620	Effector	DPT	2.024	2.92		Anti-microbial prptide
AGAP011790	Melanisation	CLIPA2	2.089	2.08		Clip domain serine protease, inhibitor of melanisation
AGAP012352	Recognition	ML1	2.31	2.70		MD2-like lipid recognition
AGAP009556	Recognition	FREP50	2.153	2.71		Fibrinogen like, function unknown
AGAP007457	Recognition	LRIM7	2.1	2.31		LRIM protein, function unknown
AGAP004455	Recognition	GNBPB1	2.099	2.45		GRAM NEGATIVE BINDING PROTEIN SUBGROUP B
AGAP004845	Opsonisation	SRCB8	2.007	3.13		Scavenger receptor, cell adhesion,
AGAP010968	Signal modulation	CLIPA9	2.466	2.64		Clip domain serine protease, function unknown
AGAP003247	Signal modulation	CLIPB19	2.136	2.17		Clip domain serine protease, function unknown
Early and late onset						
AGAP005848	Recognition	FREP44	2.051		2.59	Fibrinogen-like, function unknown
Mid onset						
AGAP004920	Apoptosis	CASPS6		2.05		Caspase, promoter of apotosis
AGAP007294	Apoptosis	IAP1		0.48		Inhibitor of apoptosis
AGAP008368	Complement	TEP14		2.35		Thioester containing protein, complement

AGAP005717	Effector	LYSC6		2.10		Lysozyme, lysis of pathogens
AGAP007343	Effector	LYSC2		2.08		Lysozymes, lysis of pathogens
AGAP000694	Effector	CEC3		2.04		Anti-microbial peptide
AGAP004036	Effector	HPX7		0.49		Peroxidase
AGAP004038	Effector	HPX8		0.48		Peroxidase
AGAP009166	IMD pathway	IKK1		0.49		IMD pathway component
AGAP008354	JAK/STAT pathway	HOP		0.41		Janus kinase of JAK/STAT pathway
AGAP005334	Melanisation	CTLMA2		2.10		C-type lectin, inhibitor of melanisation
AGAP004846	Opsonisation	SRCB9		2.21		Scavenger receptor, cell adhesion,
AGAP007455	Recognition	LRIM10		2.96		LRIM protein, function unknown
AGAP004806	Recognition	GALE6		2.20		galectin, sugar binding
AGAP004807	Recognition	GALE7		2.16		galectin, sugar binding
AGAP010774	Recognition	FREP27		2.13		Fibrinogen-like, function unknown
AGAP012529	Recognition	GALE8		2.08		galectin, sugar binding
N/A	Signal modulation	CLIPC9		2.59		Clip domain serine protease, function unknown
N/A	Signal modulation	CLIPE4		2.33		Clip domain serine protease, function unknown
AGAP000443	Signal modulation	CTL5		2.04		C-type lectin, function unknown
AGAP005672	RNAi	TSN	0.61	0.34		RNA degradation
Mid to late onset						
AGAP005693	Recognition	LRIM17		3.90	2.763	LRIM protein, function unknown
AGAP003246	Signal modulation	CLIPB2		2.05	2.176	Clip domain serine protease, function unknown
Late onset						
AGAP012037	Signal modulation	CLIPB20			2.262	Clip domain serine protease
AGAP003502	Effector	HPX6			2.551	Peroxidase
Broadly responsive						
AGAP007385	Effector	LYSC4	2.026	2.37	2.037	Lysozyme, lysis of pathogens

Genes from 3 well characterised immune signalling pathways were differentially regulated; the RNAi pathway, the IMD pathway, and the JAK/STAT pathway, as well as genes with putative roles in apoptosis. Figure 25 shows putative immune genes differentially regulated at 4 dpi. There are 5 up-regulated early onset immunity genes showing a strong bias toward recognition of pathogens and complement

associated genes; 2 genes have putative roles in pathogen recognition (an LRR protein (LRIM4) and an MD2-like receptor (ML9)); 2 genes are complement proteins (LRIM1 and a thioester containing protein – TEP5). The remaining gene (GPXH3, a glutathione peroxidise-like enzyme) has putative functions in immune effector mechanisms. The early to mid onset genes are also biased towards to genes with putative roles in recognition and complement, with 4 (of 13) genes being recognition receptors or molecules; ML1, GNBPB1 (a Gram negative bacteria binding protein), FREP50 and LRIM7; and 4 being complement thioester-containing proteins: TEP4/9/10/12. In addition to pathogen recognition and complement-associated genes, 2 CLIPs with likely functions in the regulation of immune responses, a possible oposonisation associated gene (scavenger receptor (SCRB8)) and an anti-microbial peptide Diptericin (DPT) are also up-regulated.

The largest number of immune-related virally responsive genes are regulated at 4 dpi. Although genes associated with recognition are still differentially regulated, there is a shift towards genes involved with the regulation of immune signalling and effector mechanisms of the immune system at this stage. Several pathogen recognition genes are up-regulated, including 3 carbohydrate binding galectins (GALE6/7/8) and 2 fibrinogen-like proteins (FREPS) that have a putative function in pathogen recognition although the mechanism of recognition remains unknown. Four genes with roles in immune signal modulation are up-regulated: CLIPE4/C9 (Clip domain serine proteases that putatively regulate immune signalling cascades); CTLMA2 and CTL5 (C-type lectins that modulate immune signalling). Three effector molecules are up-regulated: 2 lysozymes (LYSC2/6) that hydrolyse the cell wall of bacteria and Cecropin 3 (CEC3), an AMP that has been shown to be regulated by the IMD pathway in mosquitoes[135].

Two apoptotic genes are differentially regulated at D4 - Inhibitor of apoptosis 1 (IAP1) is down-regulated, and Caspase 6 (CASPS6) is up-regulated. Three other genes are down-regulated: 2 peroxidises HPX7/8; Hopskotch (the Janus kinase from the JAK/STAT pathway).

A dramatic reduction in the number of mid to late and late onset genes occurs, with only 1 mid to late onset immunity gene, and 3 late onset immunity genes. An LRR (putative recognition protein) gene is up-regulated at D4 and D9. HXP6 and CLIPB2/20 are up-regulated at D9 only. One single gene is differentially regulated throughout the time course; Lysozyme C4 (LYSC4).

5.6 Development of a qrt-PCR-based screen for genes with anti-viral properties in *A. gambiae*

In order to identify genes that may have a role in anti-viral immunity in *A. gambiae* mosquitoes, an RNAi and qrt-PCR based assay was developed to asses the impact gene KD has on viral infection and replication. Standard RNAi methods were used to silence GOIs[117], in brief dsRNA complimentary to a GOI was injected into 1-2 day old adult *A. gambiae* mosquitoes. The mosquito RNAi pathway uses the exogenous dsRNA as a template to destroy mRNA corresponding to the same gene, and thus silence gene expression. Three different methods were tested to optimise the protocol; initially cohorts of mosquitoes were injected with dsRNA on day 1, and subsequently injected with 5′ONNVic-eGFP on day 4. Pools of 10 mosquitoes, at three time points, were collected and homogenised in 140µl media using a motorised pestle, debris was centrifuged, and viral RNA was extracted from the resulting supernatant. Viral RNA was used as a template for cDNA production and qrt-PCR was performed, using primers designed against the viral E2 protein, to ascertain the viral genome copy number per mosquito. Two genes were tested using this protocol; the known viral antagonist in *A. gambiae* AGO2 and HOP (the janus kinase of the JAK/STAT pathway, known to be anti-viral in *Ae. aegypti* mosquitoes), with interesting, yet inconsistent results. Figure 26A shows the corresponding viral titres from these experiments with very high levels of variation from replicate to replicate. To account for variation in viral titre between the experiments, the data were normalised to the internal LacZ (non-specific) dsRNA control. This allows the trend to be visualised disregarding overall higher or lower levels of infection in one replicate compared to another. Figure 26B shows the normalised numbers where levels of variation are still very high, and as such the data may not be reliable,

however it appears that both HOP and AGO2 KD results in increase viral titre in *A. gambiae* mosquitoes

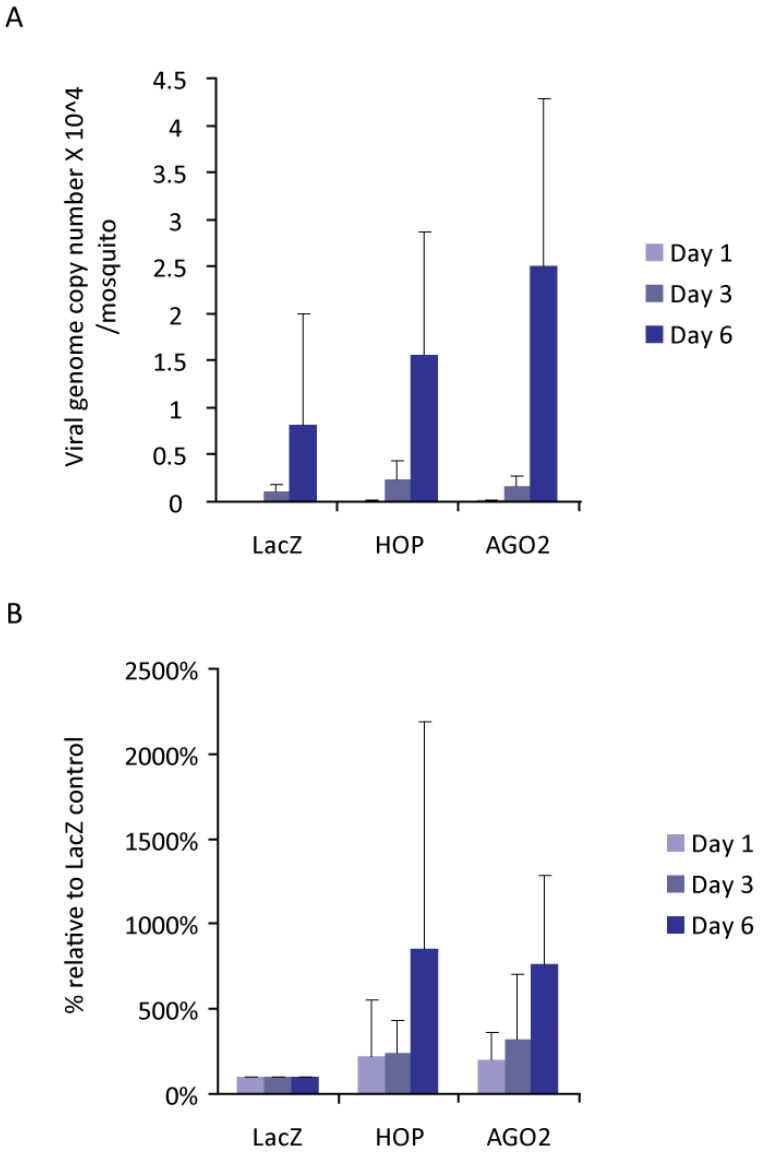


Figure 26. Viral genome copy number/mosquito for gene silenced *A. gambiae*; A) mosquitoes were injected with dsRNA complementary to a gene of interest on day 1, and further injected with 5'ONNVic-eGFP on day 4. Pools of 10 mosquitoes were collected at 1, 3 and 6 dpi, homogenised, viral RNA was extracted, cDNA was generated and qrt-PCR using primers against the viral E2 protein was carried out to calculate viral genome copy number/mosquito. **B)** Numbers were normalised relative to the internal LacZ (non-specific) dsRNA control and are displayed as a % of the LacZ titre for each time point. Error bars represent SD of 3 biological replicates.
To try and limit the variability from replicate to replicate, the number of mosquitoes

pooled was increased from 10 to 30, assuming that a larger population size would decrease the variation observed. Increasing the sample size from 10 to 30 mosquitoes had little effect on the variability observed from replicate to replicate (data not shown). A third method was tested, where mosquitoes were injected with dsRNA and ~3000 PFU P1/P2V 5'ONNVic-eGFP concurrently, and 30-50 mosquitoes were collected at day 7 post infection. Consistency between replicates was improved using this method, however variation remained quite high.

Considering the variation that appeared to be inherent to the *A. gambiae*-ONNV system, the third qrt-PCR method described above was used to initially screen a selection of genes, after which any interesting phenotypes were investigated using standard plaque assays, allowing the infection load of individual mosquitoes to be calculated. Nineteen genes were selected for screening from a) the transcriptional analysis of infected mosquitoes b) genes known to have anti-viral roles in other insects and c) the four well known immune signalling pathways in *A. gambiae* mosquitoes. Additionally a viral non-structural protein (nsP3) was included in the screen. Table 7 summarises the 19 genes selected for screening. Figure 27 shows the viral genome copy number/mosquito normalised to the LacZ control for each gene KD. Importantly 2 controls included in the screen (the known viral antagonist AGO2, and the viral nsP3 gene) behave as expected, with AGO2 KD resulting in increased viral titre, and nsP3 KD resulting in markedly decreased viral titre.

5.7 Four genes found to be viral antagonists in *A. gambiae* adult mosquitoes

Five genes were selected from the screen of 19 for further investigation (see figure 27, shaded in light blue). In order to observe the dynamics of 5'ONNVic-eGFP within a population of highly variable individuals, standard plaque assays were carried out, where the viral titre of individual mosquitoes can be calculated, allowing for statistical analysis of larger population numbers. Around 30 mosquitoes were

Table 7. Genes selected for qrt-PCR screen in *A. gambiae* mosquitoes. Candidate genes were selected from transcriptional analysis of ONNV infection in *A. gambiae*, and also included genes with known anti-viral roles in other insects and genes from the four well characterised immune signalling pathways in *A. gambiae*.

Gene Name	Regulated By ONNV	Putative function	Category	Orthologs with known phenotype
LacZ		Non-specific control	Control	
nsP3		Viral gene	Control	
Ago2		RISC complex formation	Control/RNAi pathway	Viral antagonist in *A. gambiae* [80]
Dicer 1		Enzyme	RNAi pathway	
Hopskotch	0.41	Janus kinase	JAK/STAT pathway	Viral antagonist in *Ae. aegypti* [75]
STAT1		Transcription factor	JAK/STAT pathway	
STAT2		Transcription factor	JAK/STAT pathway	
PIAS		Negative regulator of JAK signalling	JAK/STAT pathway	Viral agonist in *Ae. aegypti* [75]
REL1		Transcription factor	Toll pathway	Viral antagonist in *Ae. aegypti* [72]
CACTUS		Inhibitor of REL1	Toll pathway	
REL2		Transcription factor	IMD pathway	
ML1	2.7	Lipid recognition receptor	Pathogen recognition	
ML9	2.0	Lipid recognition receptor	Pathogen recognition	
GALE8	2.08	Sugar recognition receptor	Pathogen recognition	
LYSC2	2.1	Lysis of pathogens	Immune effector	
LYSC4	2.3	Lysis of pathogens	Immune effector	
LYSC6	2.1	Lysis of pathogens	Immune effector	
CEC3	2.04	Pathogen clearance	Immune effector	
TEP12	2.36	Pathogen clearance	Immune effector	
Helicase 2	0.4	Unwinds dsRNA	Helicase	
Helicase 3	0.3	Unwinds dsRNA	Helicase	

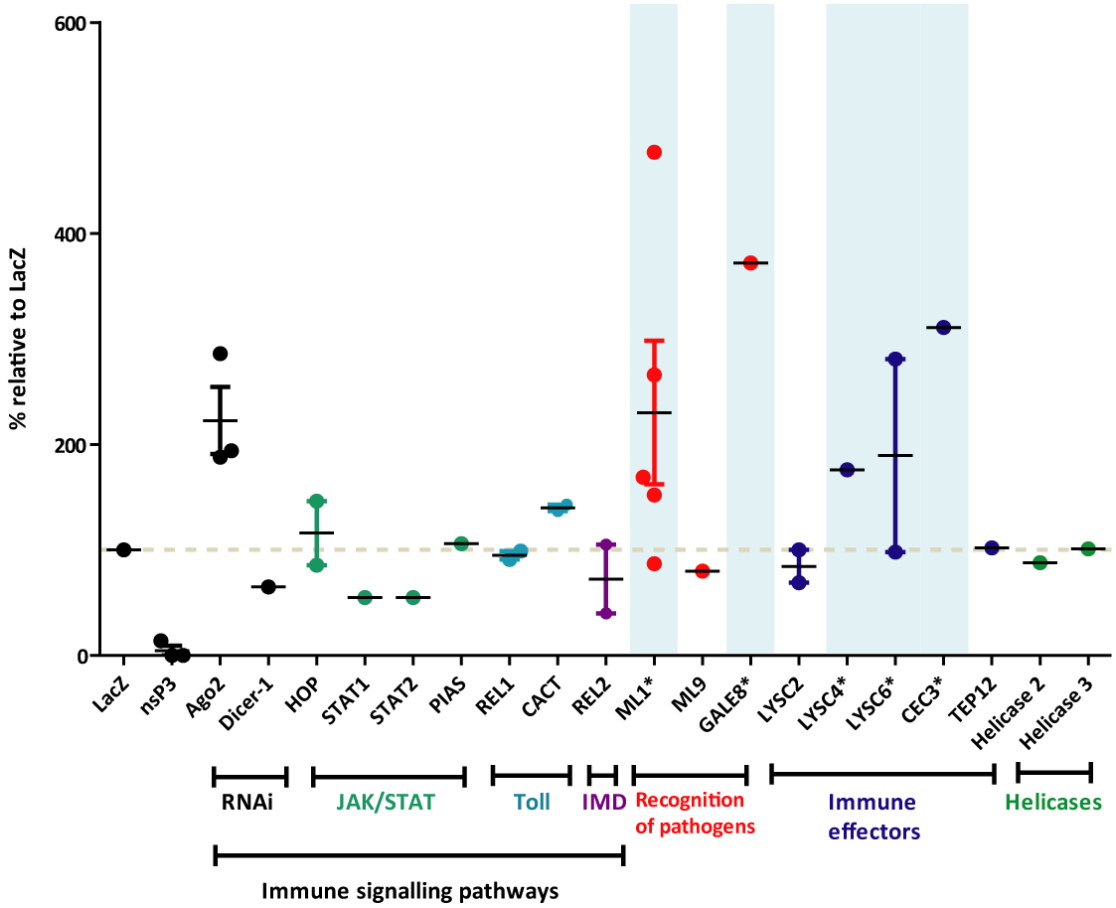


Figure 27. RNAi and qrt-PCR based screening of 19 genes for anti-viral function in *A. gambiae* mosquitoes. Mosquitoes were injected with dsRNA and ~3000PFU/mosquito 5'ONNVie-eGFP concurrently, 30 mosquitoes were collected at 7 dpi, and qrt-PCR was used to ascertain viral genome copy number/mosquito. Values were normalised relative to the LacZ (non-specific) control and are given as a percentage relative to the LacZ titre. nsP3, a viral gene was included in the screen. Genes are divided into functional categories based on which immune signalling pathway they belong to, or their putative function. Genes with an asterix* and light blue shading were selected for further characterisation.

injected with dsRNA complimentary to LacZ, ML1, CEC3, GALE8, LYSC4 or LYSC6 and 5'ONNVic-eGFP concurrently and were collected at 7 dpi. Individual mosquitoes were homogenised in MEM, debris was centrifuged, and the resulting supernatant was sterile filtered. A serial dilution of each filtered supernatant was placed onto confluent VERO cells and immobilised with a layer of nutrient agar, forcing virus to infect only the neighbouring cells to the first cell infected, and forming plaques of dead cells. After 5 days cells were live stained, and the dead plaques were counted. Two biological replicates were carried out. Plaque forming units/mosquito were

calculated and subjected to Mann Whitney Testing for significant difference to the LacZ control. Four of the 5 genes tested showed significant increases in viral titre upon gene knock down (see figure 28). CEC3 showed an increase in viral titre in one replicate, however this was not consistent in the second replicate and so it was excluded from analysis. ML1 knock down resulted in a 5.7 fold increase in viral titre (*** P<0.0001), the highest of the four genes. LYSC4 KD resulted in a 4.9 fold increase (*** P<0.0001), LYSC6 KD resulted in a 2.8 fold increase (** P<0.001) and GALE8 KD resulted in a 3.9 fold increase (* P=0.0163) in viral titre at 7 dpi. Qrt-PCR was used to confirm gene KD in mosquitoes; ML transcripts were reduced to ~9%, GALE8 transcripts were reduced to ~20% and LYSC4 transcripts were reduced to ~12% compared to LacZ controls at 7 days post dsRNA treatment. LYSC6 KD efficiency was not ascertained due to in-efficient primers.

5.7.1 Effect of silencing viral antagonists and agonists on *per os* infected *A. gambiae*

The effect of silencing two viral antagonists (ML1 and AGO2) on *per os* infections with 5'ONNVic-eGFP was investigated. Newly emerged female G3 mosquitoes were injected with dsRNA complementary to the bacterial LacZ gene, AGO2, ML1 and nsP3. 48h later mosquitoes were fed a blood-meal containing compacted RBCs, human serum and 5'ONNVic-eGFP. Assuming each mosquito ingested ~2µl of blood, each mosquito was infected with ~1640 PFU. At 3 dpi mosquito midguts were dissected and fixed in 4% PFA. GFP expression was observed using a fluorescent microscope; midguts were scored negative for GFP (uninfected), few GFP positive cells (low infection), patches of GFP positive cells throughout the cardia (medium infection), many GFP cells/most of the cardia and spreading into the anterior gut (high infection). Figure 29A shows the prevalence of GFP expression, and thus infection at 3 dpi. Chi squared tests show that silencing the viral nsP3 gene significantly reduces the prevalence of infection (P value<0.001). Silencing AGO2 or ML1 does not significantly change the prevalence of infection; however the distribution of infection is affected. Figure 29 shows the distribution of infection in

B) ML1 KD, C) AGO2 KD and D) nsP3 KD mosquitoes at 3 dpi. Silencing ML1 shifts the distribution towards medium-high infection, with significantly fewer midguts showing low infection levels (Chi Squared test P<0.001). Silencing nsP3 shifts the distribution towards uninfected/low infection with significantly larger numbers of uninfected midguts and significantly fewer highly infected midguts (Chi Squared test P<0.001). AGO2 KD shows a similar infection distribution to the LacZ control.

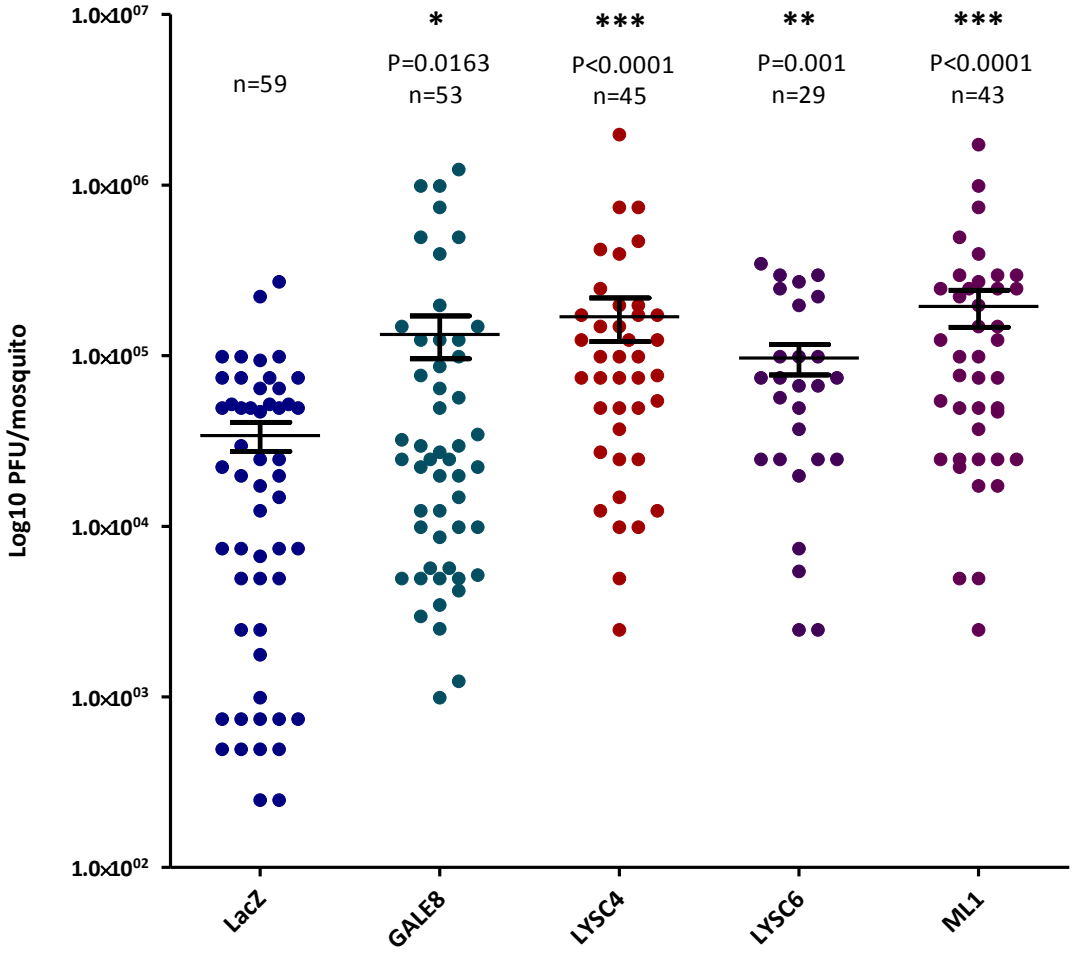


Figure 28. Four antagonists of 5'ONNVic-eGFP in *A. gambiae* mosquitoes. Mosquitoes were injected with dsRNA and ~3000 PFU/mosquito P1/P2V 5'ONNVic-eGFP concurrently. Individual mosquitoes were collected and subject the plaque assay at 7 dpi. Data was Log10 transformed. Two independent biological replicates were carried out. Error bars represent standard deviation. P values indicate significance from Man Whitney Testing of each gene KD compared to the LacZ control.

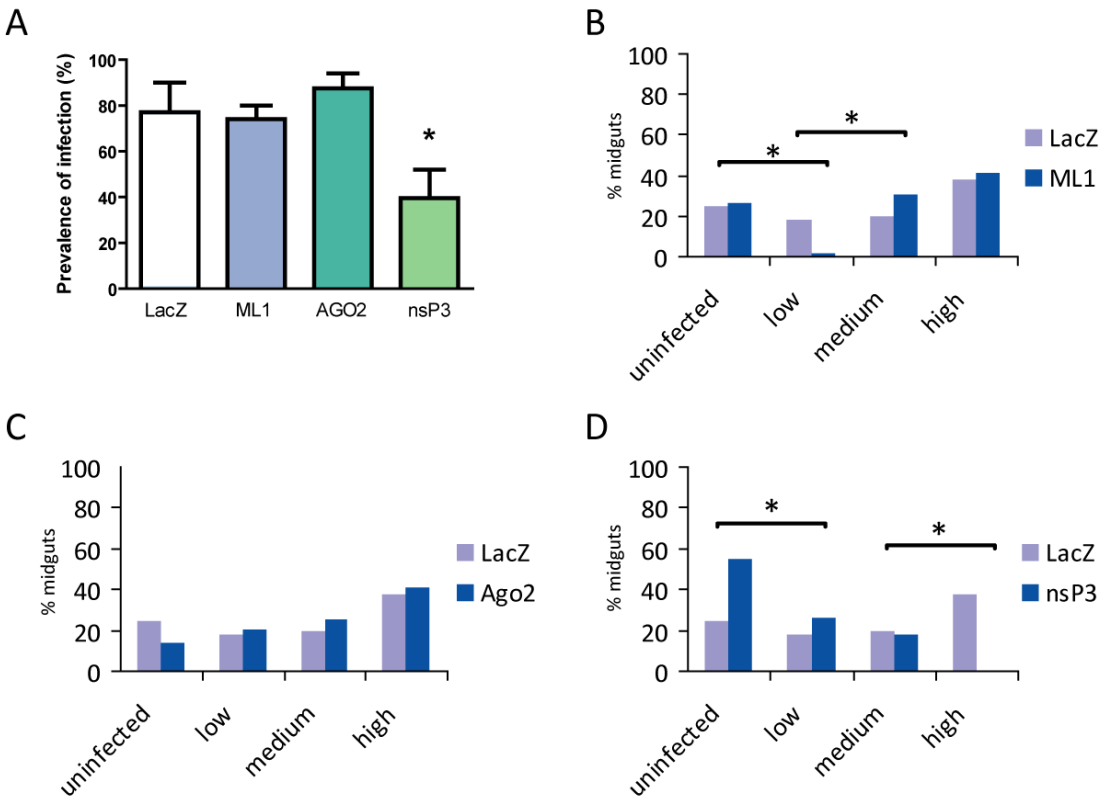


Figure 29. Effect of gene silencing on *per os* infection of *A. gambiae* **A**) % prevalence of *per os* infected *A. gambiae* after gene KD. G3 mosquitoes were injected with dsRNA against LacZ, ML1, AGO2 and nsP3 and given an infectious blood-meal of ~ 1640 PFU/mosquito 48h later. Dissected midguts were screened for GFP expression. Error bars represent standard deviation of 2 biological replicates. ***** indicates statistically significant difference in prevalence compared to the LacZ control using the Chi Squared test (where P<0.001). Infection distribution of **B)** ML1 **C)** AGO2 and **D)** nsP3 gene silenced *per os* infected G3 mosquitoes. Midguts were scored for no infection (uninfected), low infection (low), medium infection (medium) and high infection (high). ***** indicates statistically significant difference in distribution of infection between LacZ control and gene KDs using the Chi squared test (where P<0.001).

5.8 Co-infection of 5'ONNVic-eGFP and *P. berghei* in *A. gambiae* mosquitoes

5'ONNVic-eGFP is transmitted (at least during epidemics of the disease) by the important malaria vector *A. gambiae*. Epidemics of ONNV during the early 1960s and in 1996 occurred in areas of Africa where malaria is prevalent. Due to the lifelong infection formed by the alphaviruses in their mosquito hosts, it is likely that *A. gambiae* were exposed to both pathogens simultaneously during these epidemics. A study published shorted after the outbreak in the early 1960s observed a decrease in malaria cases being reported in areas where ONNV infections were recorded. In

order to investigate whether viral infection has any effect on parasite development in *A. gambiae*, co-infections of the two pathogens were carried out. Newly emerged female G3 mosquitoes were injected with ~1640 PFU 5'ONNVic-eGFP or were mock infected. 48h later mosquitoes were fed from a mouse infected with the rodent malaria, *P. berghei*. Unfed mosquitoes were removed, and seven days post blood-meal midguts were dissected and fixed in 4% PFA. Parasite oocysts on the midgut (live oocysts), and melanised ookinetes were counted. Additionally midguts were scored for 5'ONNVic-eGFP infection of the midgut musculature or nerves. Figure 30A shows the oocyst and melanised ookinete distribution in virally infected (+ virus) and mock infected (- virus) *A. gambiae* mosquitoes.

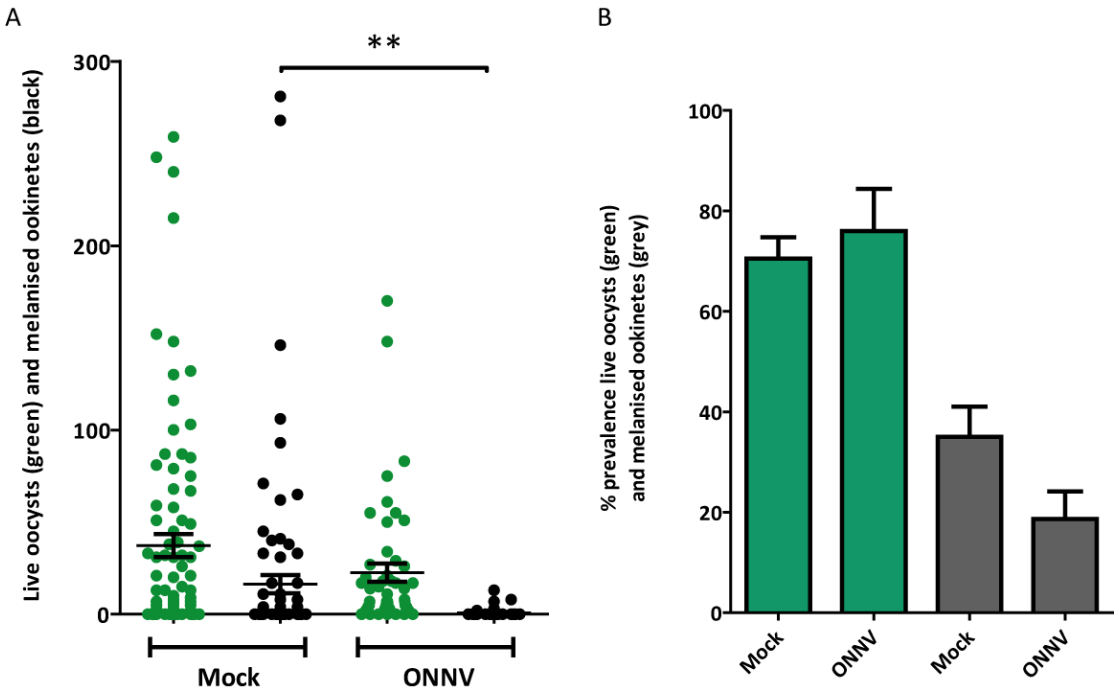


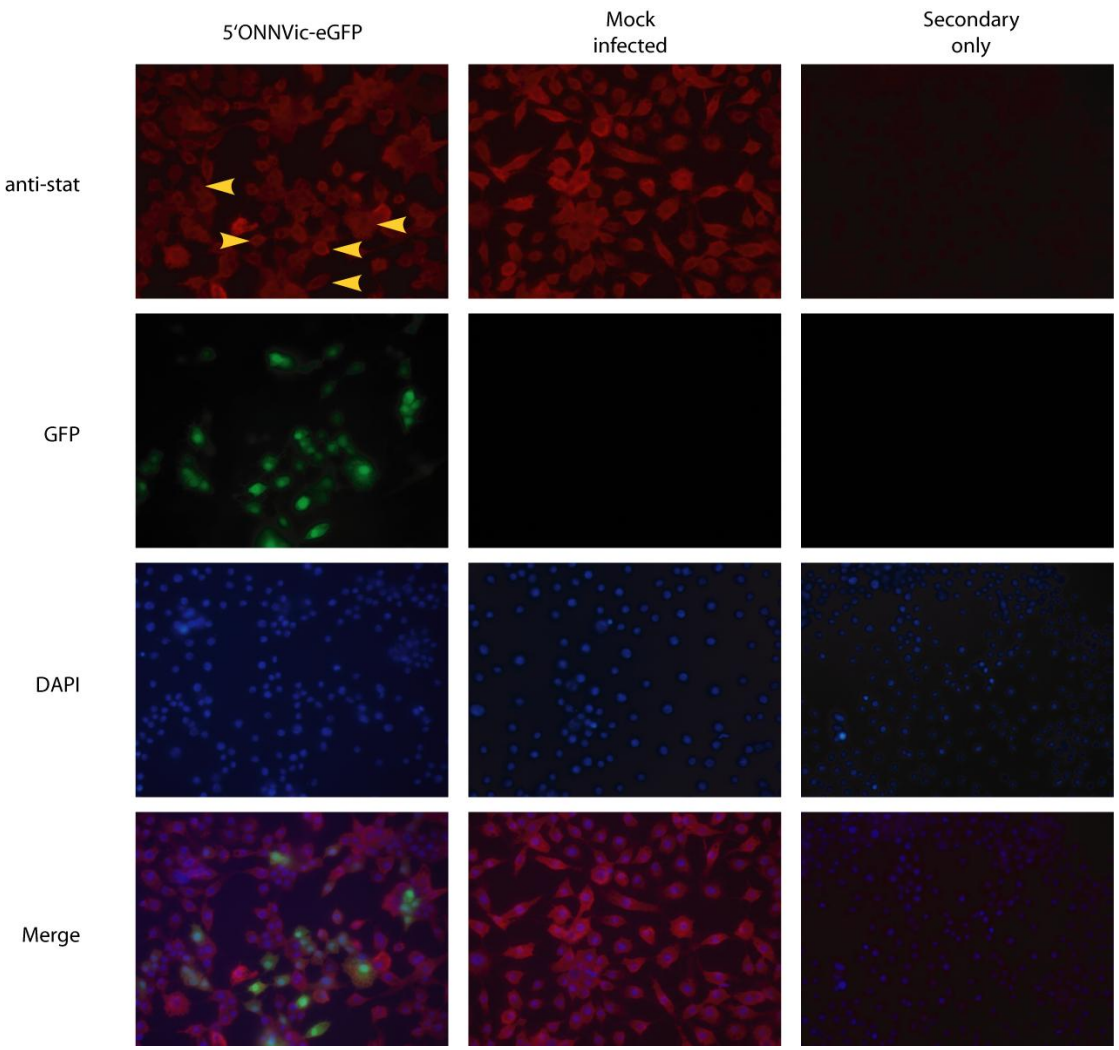
Figure 30. Co-infection of *A. gambiae* with *P. berghei* and 5'ONNVic-eGFP. G3 mosquitoes were inoculated with ~1640 PFU 5'ONNVic-eGFP (+) or mock inoculated (-). 48 hours later mosquitoes were fed on a mouse infected with *P. berghei*. **A)** 7 days post blood feeding midguts were dissected, live oocysts/melanised ookinetes were counted, and guts were scored positive or negative for 5'ONNVic-eGFP expression in midgut musculature or nerves. Parasite numbers were only included for the + virus category when 5'ONNVic-eGFP expression was observed. **B)** Prevalence of parasite infection at 7 dpi in uninfected (mock) and infected (ONNV) midguts.

An approximately 40% reduction in the number of live oocysts was observed in virally infected mosquitoes showing 5′ONNVic-eGFP infection, although this decrease is not statistically significant using the Man Whitney test. A statistically significant decrease in the numbers of melanised ookinetes was observed (** P=0.001). Additionally the prevalence of melanised ookinetes decreased from 35% in mock infected to 18.5% in virally infected mosquitoes (figure 30B).

5.9 STAT1 is not translocated to the nucleus upon viral infection of L35 cells

The observation that depletion of HOP increased titres of ONNV poses the question if JAK/STAT signalling is triggered by viral infection in L35 cells. Downstream of HOP in the JAK/STAT pathway is the transcription factor STAT. In *A. gambiae* two STATs have been characterised, STAT-1 and STAT-2. A previous worker in this laboratory generated a polyclonal antibody against STAT-1 and described translocation of STAT-1 into the nucleus of fat body cells in *A. gambiae* after exposure to bacterial pathogens[100]. Immunofluorescence assays (IFAs) were carried out using L35 cells to see if STAT-1 is also translocated into the nucleus upon viral infection. L35 cells were grown on coverslips and infected with 5′ONNVic-eGFP (or mock infected) when 100% confluent. At 24, 48 and 72hpi cells were fixed in 4% PFA before incubating with anti-STAT-1 antibody. Stained cells were mounted in vector shield with DAPI and observed under a fluorescent microscope. Secondary antibody only controls were also carried out. STAT-1 staining was observed mostly in the cytoplasm of cells, but occasionally throughout the cytoplasm and the nucleus. Secondary only controls showed very little background (see Figure 31A). No difference in STAT-1 staining was observed between uninfected cells and cells expressing GFP (and thus infected), no translocation of STAT-1 into the nucleus was seen in virally infected cells or other cells in virally infected cultures (figure 31B) compared to mock infected cultures.

A


5'ONNVic-eGFP
Mock infected
Secondary only
anti-stat
GFP
DAPI
Merge

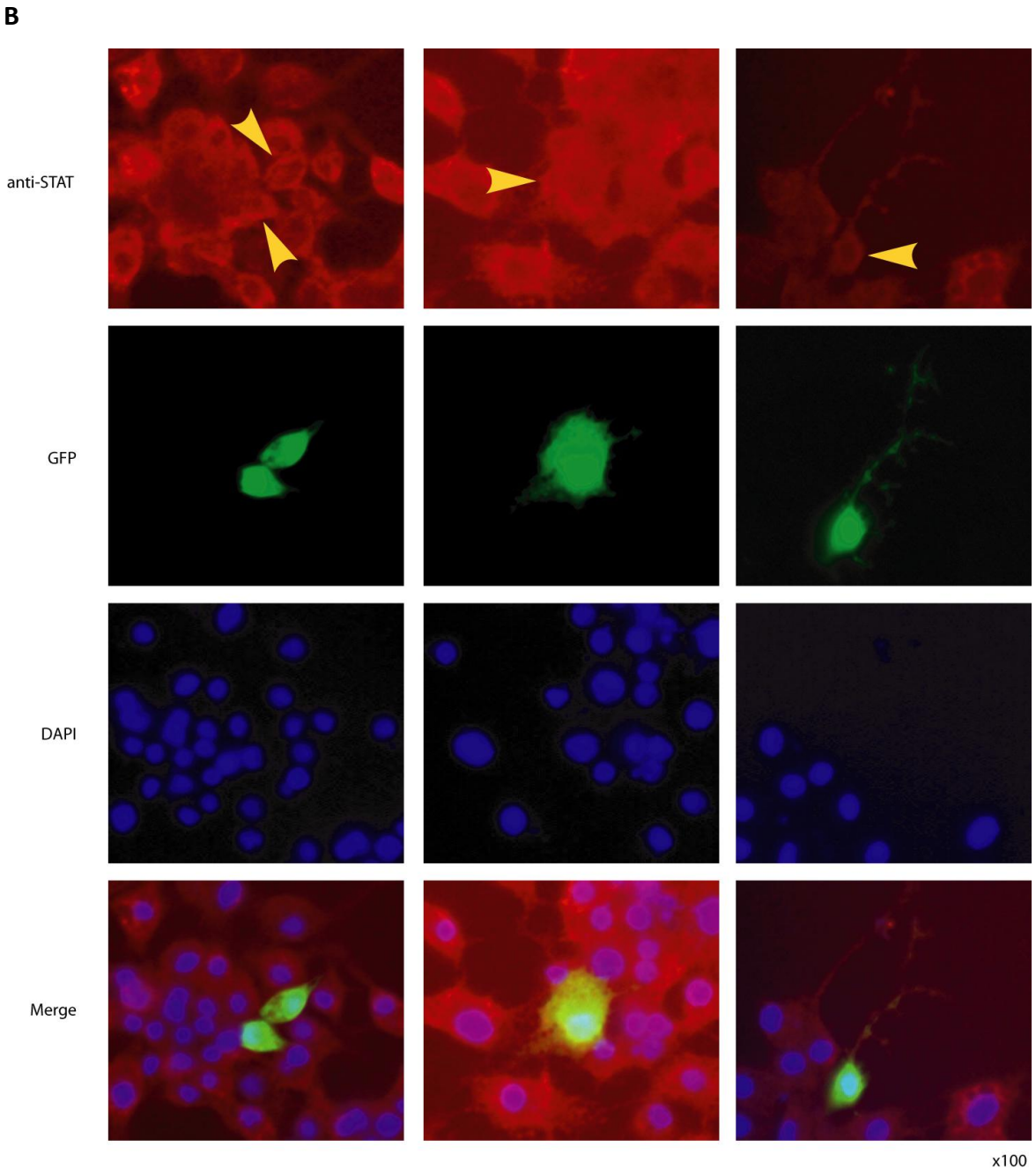


Figure 31. Patterns of STAT-1 staining in L35 cells infected with 5'ONNVic-eGFP at 48h post infection. Confluent L35 cells were mock infected or infected with 1MOI 5'ONNVic-eGFP. Cells were fixed and stained with anti-STAT-1 antibody and DAPI. **A)** 63X magnification of virally infected (5'ONNVic-eGFP) and mock infected (Mock) cells labelled with a-STAT1 and DAPI. A secondary antibody only control (secondary only) was carried out. **B)** 100X magnification of virally infected cultures stained with a-STAT1 and DAPI. Yellow arrows indicate infected cells.

5.10 Summary

This chapter has covered:

- Transcriptional responses to 5′ONNVic-eGFP infection
- Development of a qrt-PCR based screen for viral antagonists in *A. gambiae* mosquitoes
- Identification of four viral antagonists in *A. gambiae* mosquitoes
- Effect of silencing viral anatagonists and agonists on *per os* viral infections
- Co-infection of *A. gambiae* with 5′ONNVic-eGFP and the rodent malaria model *P. berghei*
- The activation of the JAK/STAT pathway in virally infected cells

6 Discussion

6.1 Transcriptional responses to ONNV infection

Due to their small genome size, RNA viruses are dependent on a variety of host cell factors to complete their life cycle. In addition to this, they must co-ordinate and regulate the host machinery to produce new viral proteins and replicate their RNA [136]. Investigating the transcriptional responses to viral infection can help to identify the factors that are used by viruses to complete their unique life cycle, and in turn, shed light on the mechanisms employed by mosquitoes to combat viral infection.

Analysis of the transcriptional responses of *A. gambiae* to 5′ONNVic-eGFP infection revealed a large number of genes that are differentially regulated by viral infection. During a time-course of infection, transcriptional responses at D1, D4 and D9 post infection were studied. D1 represents early infection, when inoculated virus has infected cells and begun to replicate. D4 represent the acute phase of infection, where virus is spreading from cell to cell and from tissue to tissue in the mosquito. D9 represent the persistent phase of infection, where levels of virus drop substantially and a low, but lifelong level infection is formed. The majority of

differential gene regulated occurs at D4, with 214 genes showing 2-fold or greater changes in expression. This might be expected as D4 represents the acute phase of infection, which is systemic. The huge decrease in differential gene expression seen by D9 follows the dramatic decrease in viral titre observed at D9. Differentially expressed genes can be grouped based on their predicted function, and temporal pattern of expression. Several patterns emerge when the genes are grouped in this way, with groups of genes representing fatty acid biosynthesis, cell division, translation and transcription, RNA degradation and interestingly immunity. These genes will be discussed in detail below:

6.1.1 Fatty acid synthesis up-regulation and viral infection

Several genes associated with fatty acid biosynthesis are upregulated: in the early onset genes a fatty acid sythase is upregulated; in the early to mid-onset genes another fatty acid synthase and a gene with putative fatty acid elongation function are upregulated; in the mid onset genes a further fatty acid sythase is up-regulated, along with two further genes with putative functions in fatty acid elongation. The differential regulation of these six genes indicates that fatty acid biosynthesis is up-regulated during early infection, and maintained during the acute phase of infection. Fatty acid sythases are up-regulated by a variety of viruses (Epstein Bar [137] and HCV [138-139] included). A study carried out by Cherry et al [136] looked at genes identified as being essential for viral replication in a genome wide RNAi screen of DCV infection in *D. melanogaster* cells. Several genes involved in fatty acid biosynthesis were identified in the RNAi screen, including CG3523, a fatty acid sythase that is the first rate limiting enzyme in the fatty acid biosynthesis pathway[140]. Depletion of this gene using RNAi resulted in a 10-fold decrease in virus in cell culture. Additionally animal mutants for the fatty acid biosynthesis pathway showed decreased levels of DCV infection. Cherry et al [136] observed that DCV infection induces a vesicular compartment where viral RNA replication is localised. They also observed that within 10 hours of DCV infection, membrane surface area within infected cells doubled, and that depletion of CG3523 blocked the

formation of virus induced vesicular compartments. The fatty acid sythase regulated at D1 and D4 post 5′ONNVic-eGFP infection (AGAP009176) is the ortholog of CG3523. Cherry et al [136] commented that the genes identified during their RNAi screen were biased towards highly conserved genes, indicating that the virus targeted highly conserved cellular functions. They speculate that this may be helpful for infection of disparate hosts, as is common in the life cycle of many insect viruses e.g. mosquito and man for the arboviruses. The up-regulation of the orthologous fatty acid sythase in a mosquito-virus system supports this idea.

6.1.2 Cell division/growth slowing

Several genes associated with the regulation of the cell cycle are down-regulated during early-mid infection (2) and mid infection (8). A cell cycle checkpoint kinase, and a gene putatively involved in the regulation of cell proliferation are down regulated during early to mid infection. Another cell cycle checkpoint protein, several genes with putative roles in regulating cellular proliferation, an ortholog of a DNA replication licencing factor MDM7 (part of a complex required for initiation and elongation of replicative forks during S-phase) and a putative regulator of chromosome condensation, are all down-regulated at D4 post infection. This suggests that normal cell division is disrupted by viral infection. The majority of these genes have roles in the promotion of cell division and/or proliferation, suggesting that cell growth slows during viral infection. It has been reported that a peptide produced by *Ae. aegypti* cells in response to SINV infection, when used to treat uninfected cells leads to a temporary arrest in cell growth, followed by recovery to normal rates of growth. These treated cells are subsequently refractory to viral infection [114]. Whether a similar peptide is produced in infected *A. gambiae* cells that causes a similar reduction in cell growth, remains to be seen.

An ortholog of CDC42 (cell division cycle protein 42) although initially classified as a cell division related protein, actually has functions in actin polymerisation in insects where roles in the formation of polarised actin filaments, elongation of cell shape

and cell signalling have been reported [141]. In *A. gambiae* CDC42 has been shown to be required for the formation of actin zones around invading *P. berghei* ookinetes, and, through an unknown mechanism mediates melanisation. However it has no effect on parasite killing [142]. Interestingly CDC42 has been shown in endothelial cells to be required for the formation of filipodia that facilitate the internalisation of DENV-2 particles [143]. Inhibition of actin re-organisation in endothelial cells led to dramatic decreases in DENV-2 infection. CDC42 was shown to be active in infected cells, and to be required for the formation of filipodia triggered by binding of DENV-2 to the surface of cells, leading to the uptake of virus [143]. The ortholog of CDC42 is down-regulated in 5′ONNVic-eGFP infected *A. gambiae*, suggesting the inhibition of actin re-organisation. It would be interesting to see if depletion of the CDC42 gene would affect ONNV infection, as seen in human endothelial cells infected with DENV2.

6.1.3 Inhibition of RNA degradation including RNAi

RNAi is a potent and well characterised anti-viral mechanism in mosquitoes. One of the components of the RNAi pathway, Tudor-SN (TSN), is down-regulated by 5′ONNVic-eGFP infected *A. gambiae* at D4 post infection. TSN is a known viral antagonist in *Ae. aegypti* mosquitoes; depletion of TSN in *Ae. aegypti* mosquitoes results in significant increases in prevalence of SINV infection, and a modest increases in viral titre [82]. TSN is a component of the RISC complex in *D. melanogaster* [144] although its specific function remains elusive. The down-regulation of TSN may be advantageous to the virus, by limiting the destruction of viral RNA targets. Whether TSN is regulated by the virus itself, or perhaps by the host through a negative feedback loop, is unclear.

Several other genes with putative functions in non-classical RNAi are also down-regulated. A member of the Argonaute protein family (termed AgAgo5 by Campbell et al [77]) that is closely related to the *D.melanogaster* Piwi and Aubergine proteins is down-regulated at D4 post infection. In *D. melanogaster*, Piwi, Aubergine and

Ago3 form a subclass of the Argonaute protein family. They contain the characteristic Piwi and Paz domains; Paz is a small RNA binding domain and Piwi is an RNAse-H type domain that relies on divalent cation binding to facilitate dsRNA-guided cleavage of ssRNA [77]. In *D. melanogaster*, small RNAs bound to Piwi and Aubergine and Ago3 were found to be derived from hotspots of the genome, the majority of which were genomic repeats or transposons, and so there were called repeat-associated small RNAs (rasiRNAs) [145] or more recently, piwi-associated RNAs (piRNAs). Generation of piRNA does not require a dicer enzyme, as they are believed to be generated from ssRNA [146]. They have been associated with a variety of functions in germ-line tissue; ensuring genomic stability by silencing endodgenous selfish genetic elements such as transposons, maintenance of telomeres and mRNA silencing in germ-line cells [77, 146]. In a comparative genomic study of components of the small RNA regulatory pathways in mosquitoes by Campbell et al [77], AgAgo5 was shown to form a clade with several Piwi proteins from Aedes and Culex but distinct from the Drosophila piwi proteins. Campbell et al suggest that the piwi subfamily comprises several gene families, some of which have arisen from recent gene duplication. The function of AgAgo5 has not been investigated to date. There is limited evidence that Piwi mediated RNAi also contributes to anti-viral defence in insects along with the classical RNAi pathway; *D. melanogaster* Piwi mutant flies are more susceptible to WNV infection [61]. Additionally depletion of AgAgo3 results in increases susceptibility of adult *A. gambiae* to ONNV infection [80]. How this non-classical RNAi pathway may be targeting viral genomes remains unclear.

In addition to AgAgo5, Dicer-1 is also down-regulated at D4 post infection. In *D. melanogaster*, Dicer-1 (DCR-1) generates miRNA from endogenous RNA and mediates gene expression regulation through RNAi [78]. DCR-1 is required for the production of miRNA, however it may also contribute towards the production of siRNA in the classical RNAi pathway [78]. Although DCR-1 and DCR-2 have distinct (although somewhat overlapping) functions in the fruitfly, it remains to be seen if this division in functionality is maintained in *A. gambiae*. In *Ae. aegypti* Dicer-2 has been shown to be important in regulating DENV2 infection [83]; whether or not

Dicer-1 has a similar role has yet to be tested. The down-regulation of several components of the RNAi pathway (albeit it from different arms of the RNAi pathway in *Drosophila*) suggests that RNAi is inhibited in *A. gambiae* infected with 5′ONNVic-eGFP. Suppression of RNAi would be advantageous to the virus, and could be a result of interference with host gene expression. Viruses have developed a range of mechanisms to suppress RNAi – strategies range from direct interference with components of the RNAi pathway, such as the FHV B2 [85] protein and the tombusviral P19 protein [147]; recruitment of endogenous inhibitors of RNAi; out-competition of the RISC complex by unproductive viral-derived small RNAs; alteration of host gene expression, illustrated by Geminiviral transactivator proteins that either promote the transcription of genes that interfere with RNAi, or suppressing expression of mediators of RNAi [148]. This wide variety of evasion mechanisms illustrates the strong pressure exerted on viruses by the RNAi pathway in plants and invertebrates. The changes in gene expression of several RNAi components seen in this study may be mediated by ONNV to allow efficient replication within infected cells. ONNV is not thought to encode a protein that suppresses RNAi directly [11], however viral products are known to translocate into the nucleus of cells infected with alphaviruses, for example 90% of SFV nsP2 protein localises to the nucleus of infected cells [149]. Viral products that translocate to the nucleus may interfere with the host cell transcriptome.

6.1.4 The DEAD-box helicases

Whereas invertebrates utilise the siRNA pathway to recognise exogenous dsRNA, vertebrates employ other pathways that trigger interferon (IFN) signalling [150]. The mammalian RIG-1 like receptors (RLRs) recognise exogenous dsRNA and contain a DExD/H box domain, similar to the dicer enzymes in invertebrates. It is suggested that the RLRs may have evolved from a dicer-like ancestor [150]. The DEAD-box helicases are a large family of conserved proteins found in almost every organism, with diverse functions in all aspects of RNA metabolism [150].

A DEAD-box helicase called Belle was identified in *D. melanogaster* as a component of the RNAi pathway by Ulvila et al [151]. In addition, a Belle paralog in *C.elegans* was identified as an RNAi component [152]. Interestingly the human ortholog of belle, DDX3X, was recently shown to be anti-viral; in response to viral pathogens, DDX3X promotes IFN production [153]. Additionally multiple viruses have been shown to interact with DDX3X, and modulate its function [153]. These studies indicate intriguing roles for the DEAD-box helicases in anti-viral immunity and warrant further investigation into this function of the protein family [150].

In this study a group of five DEAD-box helicases, and a further possible helicase, were down-regulated at D4 post infection. This raises the possibility that they may have functions in sensing viral infection/contributing to RNAi responses, although any function in anti-viral immunity has yet to be investigated.

6.1.5 Regulated Immunity genes

Almost 1/3 of regulated genes at D1 and D9 post infection and 1/5 of genes regulated at D4, are immune related. The vast majority of these genes are up-regulated, with only five being down-regulated in response to ONNV infection, indicating that *A. gambiae* is capable of recognising viral infection, resulting in the activation of immune signalling. The mosquito immune response comprises several arms of response to pathogens. The cellular response involves phagocytosis and melanotic encapsulation of invading bacteria and parasites. The humoral respose is made up of hemolymph proteins that recognise invading pathogens and trigger melanisation, lysis or opsonisation. The immune signalling pathways respond to specific pathogens and regulate the expression of a variety of immune genes. Of the numerous putative immunity genes regulated by ONNV, a large number are associated with the humoral response and are hemolymph proteins including many complement-associated genes (several TEPs and LRIM1) and a smaller number of melanisation-associated genes (CLIPA9 and CTLMA2) and opsonisation-associated genes (SRCB8/9). Considering that virus was inoculated directly into the hemolymph

of mosquitoes during these experiments, it is not surprising that a large number of hemolymph genes respond to infection.

The early responsive genes are mostly complement associated and recognition associated genes. LRIM1 functions during Plasmodium infection to lyse parasites through the targeting of TEP1 to parasite surfaces. During 5'ONNVic-eGFP infection, LRIM1, but not other members of the parasite targeting complex APL1C and TEP1, is up-regulated. This suggests that LRIM1 may also function independently of the described complex, possibly regulating other immune mechanisms that may target viral infection. During early and mid infection 3 further LRIMs and 5 TEPs are up-regulated. These may comprise a different branch of the complement pathway that targets viral infection. Two MD2-like receptors are up-regulated in response to 5'ONNVic-eGFP infection. The MLs will be discussed in detail in section 6.2.2.

During mid infection, in addition to complement and recognition-associated genes, genes with known functions in the inhibition of melanisation (CTLMA2 and CLIPA2) are up-regulated. This suggests that viral infection triggers a shift in hemolymph proteins towards those with roles in lysis and opsonisation and shuts down the melanisation pathway. Additionally the IMD, JAK/STAT and RNAi pathway appear to be inhibited through the down-regulation of components of these pathways (IKK1, HOP and TSN respectively).

D. melanogaster IKK1, a member of a complex that activates NF-kB in the IMD pathway, was identified as being required for the activation of Relish [154]. As part of the IKK complex, it targets relish for cleavage and subsequent activation, leading to the up-regulation of several AMPs [155]. The down-regulation of IKK suggests that the IMD pathway is being inhibited. An anti-microbial peptide, ortholog of *D. melanogaster* Diptericin (DPT) is up-regulated during early to mid infection (2 and 3.6 fold regulated at D1 and D4 respectively). DPT is though to be down-stream of the IMD pathway in *D. Melanogaster* [156], indicating that the IMD pathway is switched on by ONNV infection, although it has yet to be confirmed in *A. gambiae* that DPT is regulated by the IMD pathway. The IMD pathway has recently been

implicated in anti-viral immunity in *D. melanogaster*; two studies demonstrated increases in SINV [157] and CrPV [158] infection in Drosophila mutants for components of the IMD pathway. To date no evidence has suggested that the IMD pathway is involved in anti-viral defence in mosquitoes, by contrast, the IMD pathway was shown to have no effect on DENVs infection via silencing of Caspar (the negative regulator of the pathway) in *Ae. aegypti* mosquitoes [72].

As outlined in the introduction to this study, the JAK/STAT pathway is a well characterised anti-viral signalling pathway in mammalian systems. Recent evidence has also implicated the JAK/STAT pathway in anti-viral immunity in mosquitoes [75]. HOP, the Janus kinase of the JAK/STAT pathway, is down-regulated during mid-infection. The role of JAK/STAT signalling in *A. gambiae* will be discussed in detail in relation to gene KD experiments in *A. gambiae* cells.

The down-regulation of an inhibitor of apoptosis (IAP1), and the up-regulation of a caspase (CASPS6) indicate that apoptosis may be triggered by ONNV infection. Although apoptosis has not been shown to be an important regulator of SIN in *Ae. aegypti* derived cells [38], and is generally not shown to be an important invertebrate response to viral infection, apoptosis is certainly a well characterised response to viral infection in mammalian cells. It would be interesting to investigate the effect of activation or inhibition of apoptosis in *A. gambiae* mosquitoes on viral infection.

6.1.6 Genes with unknown function

Despite the differential regulation of a number of putative immunity genes, a large number of uncharacterised genes, many with no predicted function and/or gene architecture are differentially regulated by 5'ONNVic-eGFP infection. Some of these genes may represent uncharacterised anti-viral mechanisms that are specific to *A. gambiae*. Genes with now well characterised mosquito-specific immune functions have been discovered previously through transcriptional profiling of pathogen

challenged *A. gambiae*. An example of this being LRIM1, a potent malaria parasite antagonist that led to the discovery of a mosquito specific family of LRR proteins with functions in mosquito immunity [126]. Sixty six genes with unknown functions were differentially regulated by ONNV infection. A number of the most highly regulated genes are found in a cluster on chromosome 2R; AGAP003773/5/6/7/8. All are predicted to encode mosquito specific conserved small proteins (<200KDa) with orthologs in *Ae. aegypti* and *C. quinquefaciatus* but no orthologs in *D. melanogaster*. AGAP003773/5/7/8 are up-regulated during early infection (4.3/2.7/4/4.1 fold respectively), and AGAP003776 is up-regulated during mid-late infection (8.3 and 2.8 fold at D4 and D9 respectively). As the most highly differentially regulated gene, the anti-viral activity of AGAP003776 was investigated in *A. gambiae* cells, however did not appear to have a substantial effect on viral replication.

6.1.7 Limitations of the microarray study

The main limitation of the microarray analysis of viral infection in *A. gambiae* relates to the rejection of data through the strict filtering criteria applied during analysis to limit the number of false positive data. Data can be rejected for several reasons. The first is low intensity and poor feature quality; features with intensity ratios that do not exceed background levels are excluded from analysis through a flagging system. These features are flagged as 'absent' and are excluded from analysis in Genespring. Further to this features with diameters that exceed two times the SD of the mean are excluded, along with manually flagged features (such as those with scratches and/or dirt on the surface of the array) and features with saturated intensity (i.e. the intensity of one or more dyes is higher than the upper threshold of measurement). These data are excluded prior to normalisation (which counteracts dye specific effects) and thus cannot be 'retrieved' from the raw data and compared to normalised data. The second major reason for the rejection of data is poor consistency across biological replicates; T-test P values of >0.05 are excluded from analysis. The limitation that arises from this rejection of data, is the loss of power to

cluster data temporally with accuracy. Where fold change ratios have passed all the filters for only one of three time points, it is not necessarily accurate to say that this gene is only differentially regulated at this one time point. However, including data with poor consistency, or data that fails other filtering criteria increases the number of false positive data and decreases accuracy. As such, there is a balance between how strict or relaxed the filtering criteria are, and often this will depend on the reason for conducting the microarray analysis in the first place. In this study, the broad responses to viral infection were investigated as a starting point for identifying anti-viral immune mechanisms in *A. gambiae* mosquitoes. Considering the large amount of variation that is seen in viral infection between individually infected *A. gambiae* mosquitoes, it was decided not to exclude data where fold change ratios for all three time points were not present. Although perhaps increasing the number of false positive data, this prevents the exclusion of a number of potentially immune related genes. Although the data cannot be accurately clustered, the overall trend of expression of each gene was plotted over the time course including data with high P values. These plots gave two valuable insights into the data that have passed all the filtering criteria: firstly the trend of expression over the whole time course was seen; secondly, using the multiple probes generated for immunity genes, the trend of multiple probes corresponding to the same gene was visualised. The plots, although containing data that is statistically in-accurate, gave an increase or decrease in confidence of the data that passed the filtering criteria.

6.2 Viral antagonist discovery *in vivo*

An assay utilising RNAi and quantitative real time PCR was developed in order to identify genes that play a role in anti-viral immunity in *A. gambiae*. Target genes were silenced using standard RNAi techniques (as described by [117]) in adult mosquitoes that were subsequently infected with 5'ONNVic-eGFP. Qrt-PCR was then used to assay the titre of a viral gene, nsP3, in pools of 10 or 30 mosquitoes. Using this technique, high levels of variation in infection levels were observed between biological replicates of experiments, despite normalising data to the internal LacZ

control to remove overall changes in infection level from experiment to experiment. The protocol was modified, and two controls were added; dsRNA against the viral gene, nsP3, and against the known viral antagonist Ago2 were used during experiments. High levels of variation between pools of mosquitoes were still observed, however the controls behaved as expected with higher titres of virus in Ago2 KD mosquitoes, and very low viral titres in nsP3 KD mosquitoes (as seen by Keene et al [80]). The high levels of variation reflect the inherent high levels of variation seen between individuals infected with ONNV. Using the fast and efficient qrt-PCR method, 19 genes were screen for roles in anti-viral immunity, and five candidates were taken forward for further study. Carrying out standard plaque assays determined the viral load from individual mosquitoes, allowing the spread of the data (originally lost when pooling groups of mosquitoes) to be observed, and taken into account for statistical analysis of the results. This gives a more powerful dataset for analysis and gives more accurate and robust statistical results. Four of the five genes tested showed consistent and significant changes in viral titre when depleted by RNAi, and are discussed below.

6.2.1 The Galectins – GALE8

The defining feature of the Galectin family of proteins in the ability to bind β-galactosidase sugars [159]. Galectins are found in a wide variety of organisms from parasites to sponges to mammals [160]. The complex galectin repertoire found in many organisms, multiple isoforms and observed plasticity in sugar binding suggests substantial diversity in the glycan recognition properties of the Galectins [160]. Most Galectins are soluble, do not contain transmembrane (TM) or signal peptides and are secreted by the non-classical secretion pathway [159]. Three forms of Galectins are found in mammalian systems; the prototype galectins, containing a single carbohydrate recognition domain (CRD), often forming homodimers of non-covalently linked subunits; the chimera galectins, and the tandem repeat galectins that contain two CRDs. In *D. melanogaster,* galectins are of the tandem repeat type, however in *A. gambiae*, the galectins are of the prototype galectins [159].

In mammalian systems galectins have diverse functions in cell adhesion, proliferation, migration, apoptosis, inflammation, immunity and immunoregulation [159]. In *D. melanogaster* expression patterns of Dmgal during embryogenesis also suggest important roles in mesoderm and neural layer tissue formation during development. *A. gambiae* IGALE20 was found to be up-regulated in response to bacterial and malaria infection, and was postulated to be involved in pathogen recognition by binding saccharide ligands on microbial surfaces [161].

There are 10 putative galectins in *A. gambiae* (CEGG ImmunoDB), three of which are up-regulated by ONNV infection – GALE 6/7/8. All three galectins are mosquito specific (part of a mosquito species specific expansion including AgGALE4-8), and have no orthologs in *D. melanogaster* (see figure 32). Both AgGALE6 and AgGALE8 have orthologs in *Culex quinquefasciatus*, and show duplications in *Ae. aegypti* (AeGALE6A/6B and AeGALE8A/8B). GALE7 is specific to *A. gambiae* and does not have othologs in either mosquito (Cegg immuno dB http://cegg.unige.ch/Insecta/immunodb). This mosquito species specific expansion maybe due to the haematophagous lifestyle of the mosquitoes, and subsequent exposure to a disparate group of pathogens compared to *D. melanogaster*, including viral pathogens. The up-regulation of several galectins in response to ONNV infection suggests that this group of mosquito specific galectins have anti-viral roles.

In mammals, galectins have known functions during viral infection. Nipah virus, a paramyxovirus that causes severe encephalitis, infects cells through attachment of its envelope glycoproteins to host endothelial cell ephrinB2/B3 receptors, additionally triggering cell-cell fusion forming syncytia [160]. Galectin 1 crosslinks the N-glycans displayed in the envelope proteins of Nipah virus, and Hendra virus (another paramyxovirus), blocking cell infection and cell-cell fusion [160]. Additionally Galectin expression is regulated by herpesvirus 1, Newcastle disease [162], Epstein Barr Virus [163], Hepatitus C virus [164] and Human papiloma virus (HPV)[165]. Additionally Galectin 3 secretion and carbohydrate binding increase upon herpesvirus 1 infection [166], suggesting that Galectins function at several

levels of anti-viral defence, from initial recognition and blocking of envelope and fusion glycoproteins to the activation and amplification of the innate and adaptive immune reponses [160].

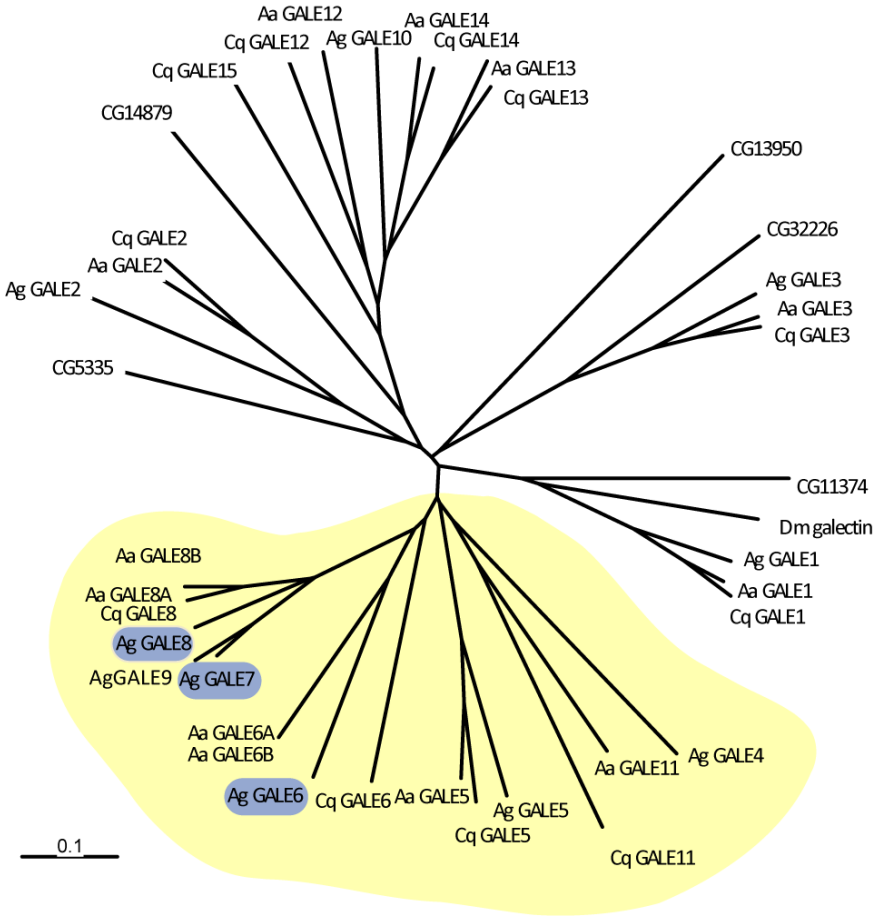


Figure 32. Phylogenetic tree of Galectins in three mosquitoes and *Drosophila*. Phylogenetic tree of Galectins in *D. melanogaster* (Dm or CG), *Ae. aegypti* (Aa), *C. quinquefaciatus* (Cq) and *A. gambiae* (Ag) adapted from Waterhouse et al 2007 [167]. The mosquito specific expansion of GALE4/5/6/7/8/9/11 is indicated by the pale yellow balloon. GALEs up-regulated by ONNV infection are indicated by blue boxes.

GALE8, shown here to be an antagonist of ONNV infection, may function in several ways to limit viral infection. Firstly a direct function can be envisaged – GALE8 may bind to and cross link the sugars on ONNV envelope proteins and thus directly prevent infection of cells. Alternatively galectin binding may target viral particles for opsonisation by hemocytes. A second function of the galectins may be to recruit or activate hemocytes circulating in the hemolymph, leading to the activation of other anti-viral immune mechanisms, and so may have an indirect effect on viral infection.

6.2.2 The MD2-like receptors – ML1

Lipopolysaccharide (LPS), a component of bacterial cell walls, is one of the best characterised pathogen associated molecular patterns (PAMPs). Many studies have shown that in mammals LPS binds to the toll-like receptor (TLR)-4 (reviewed by [168]). In order to trigger TLR-4 signalling, other ancillary proteins are required, one of which is MD-2 [169]. In fact MD-2 is absolutely required for TLR-4 signalling induced by LPS as shown by the fact that TLR-4 and MD-2 form a heterodimer complex capable of sensing LPS [168]. A family of proteins with high sequence homology to MD-2 has been reported by Inohara and Nunez [170], called the MD2-like receptors (MLs). It consists of genes with four conserved cysteine residues, an MD2-like lipid binding domain, and an N-terminal signal peptide, indicating the protein family is secreted [170]. At that time 155 MLs were identified in animals and plants. MLs, like MD-2, have been implicated in lipid recognition and metabolism. It is postulated that MLs may act as a cofactor for the recognition of LPS in flies, where LPS triggers an immune response, however a TLR that recognises LPS has not been identified [170]. In a recent comparison of the transcriptional responses of *A. gambiae* to *P. falciparum* and *P. berghei* infection, AgM1 was shown to be up-regulated during *P. falciparum* ookinete invasion of the midgut [134]. Subsequent RNAi of the AgML1 gene resulted in significant increases in *P. falciparum* oocysts after infectious blood meals. In *A. gambiae* there are 15 MLs (CEGG ImmunoDB), an expansion of the 10 predicted MLs in *D. melanogaster*, which may be a result of adaptation to a haematophagous lifestyle and thus, changes in pathogen exposure [134].

ONNV infection differentially regulated two MLs; ML1 and ML9. AgML1 has single copy orthologs in *Ae. aegypti* and *C. quinquefasciatus*. AgML9 has an ortholog in *C. quinquefasciatus* but has been duplicated in *Ae. aegypti* (AeML9A/AeML9B) (Cegg ImmunoDB). ML1 KD significantly increases viral titre in *A. gambiae* mosquitoes (~5-fold increase in virus). There is evidence in mammalian systems that the envelope glycoprotein of the Ebola virus triggers expression of pro-inflammatory cytokines in a TLR-4/MD-2 receptor dependent manner [171]. Several other studies also link viral infection (Vesicular stomatitus virus (VSV), respiratory syncytial virus, mouse mammary tumor virus [171] and Kaposi sarcoma herpesvirus [172]) with TLR-4 signalling, however the role of MD-2 in these interactions is not clear. The mechanisms through which ML1 may modulate viral infection are unknown. A possible mode of action (based on MD2 signalling) could be that ML1 acts as a recognition protein that recognises a viral PAMP and subsequently is able to bind to a signalling receptor on the surface of cells, resulting in dimerisation and activation of signalling that induces expression of anti-viral genes. There are 10 predicted Toll-like receptors in *A. gambiae* with unknown functions (ImmunoDB, http://cegg.unige.ch/Insecta/immunodb). TOLL9 is of particular interest due observed patterns of midgut expression, and its location within the 2La inversion that is accociated with melanotic encapsulation of Plasmodium parasites [173]. It would be interesting to see if silencing any of the TOLLs would give the same phenotype as ML1 KD, indicating that they may be functioning in the same genetic pathway. An alternative possibility is that ML1 bound to a PAMP can activate other immune genes in a direct signalling cascade, activating anti-viral mechanisms.

6.2.3 The Lysozymes – LYSC4/6

Lysozymes have been implicated in anti-bacterial immunity for over 40 years [174]. They are described as proteins that hydrolyse the 1-4 glycosidic linkage between the alternating N-acetylglucosamine and N-acetylmuramic acid residues of the peptidoglycan layer of bacterial cell walls [174]. Lysozymes are found in many organisms, including a wide range of insects including the Lepidoptera, Orthoptera

and the Diptera. There are two types of lysozyme identified in insects – the c (or chicken) type and the i (or invertebrate) type. There are eight predicted lysozyme Cs in *A. gambiae*; all are predicted to be secreted, 7 having a single lysozyme domain and are found in 2 clusters on chromosome 2L – LYSC1,2,3,5 &8, and LYSC4 & 7 [174]. LYSC6 is unusual in that is has 5 lysozyme domains. Similarly long lysozymes have also been identified in other insects, including *D. melanogaster*. A pattern of 8 cysteine residues is required for carbohydrate binding of the lysozymes, and 2 other amino acids have been shown to be essential for muramidase activity [174]. LYSC2 has all the conserved amino acids, however LYSC4 is missing the one of the residues important for muramidase activity. Several of the lysozyme domains in LYSC6 are also missing this amino acid, suggesting that LYSC4 and 6 have reduced or abolished enzymatic activity [174]. Whether this putative change in enzymatic activity alters their ability to kill bacteria is unclear; evidence has shown that reducing or abolishing enzymatic activity in several lysozymes does not alter their bactericidal activity [175-177]. Both proteins still retain the putative ability to bind N-acetylglucosamine or other oligoschharides [174]. Three killing mechanisms have been put forward for the lysozymes. Each begins with an interaction between the lysozyme and polyanionic molecules in the bacterial cell wall followed by a) stimulation of autolysin activity, b) membrane perturbation or c) peptidoglycan hydrolysis [174].

Although evidence of immune function for LYSC1 and 2 has been seen in *A. gambiae*, the roles of the other putative lysozymes is unclear. It is suggested that LYSC3 and 8 may be involved in the digestion of bacteria as a food source, fitting with the expression profiles of the two genes being restricted to the larval stage of the mosquito, where the larvae live in a moist and bacteria rich environment [174]. Expression of LYSC4 is ubiquitous throughout the developmental stages of the mosquito, but it restricted to expression in the fat body, ovaries and malphigian tubules [174]. LYSC6 expression is restricted to adult mosquitoes where expression was detected in the head and abdomen of adult mosquitoes, however transcripts could not be detected in individual tissues tested (fat body, ovaries, midgut, malphigian tubules and salivary glands)[174].

LYSC2, 4 and 6 were all seen to be up-regulated by ONNV infection. Screening the three genes for anti-viral activity showed that LYSC4 and 6 are antagonists of ONNV infection. Human lysozyme demonstrates anti-viral immunity: Human urinary Lysozyme C as well as lysozyme from chicken egg whites, human milk lysozyme and human neutrophil lysozyme all have an anti-Human Immunodeficiency Virus (HIV) activity [178]. A lysozyme from a marine organism was shown to inhibit Pseudo Rabies Virus (PRV) growth in cell culture [179]. The mechanism of lysozyme anti-viral activity is not clear.

6.3 JAK/STAT pathway and ONNV infection in *A. gambiae* derived cells

The JAK/STAT pathway has been shown to be involved in anti-viral immunity in a variety of organisms, from humans to flies and mosquitoes. The data in this study indicate that the JAK/STAT pathway regulates viral infection in L35 cells, however infection of cells does not trigger translocation of STAT-1, the transcription factor of the JAK/STAT pathway, into the nucleus. Therefore JAK/STAT signalling switches on anti-viral genes, but the pathway is not activated by ONNV infection. There are two STAT genes in *A. gambiae*, and it is possible that STAT-1 is not the transcription factor that responds to JAK/STAT activation in response to viral infection. It would be interesting to investigate whether HOP is phosphorylated in ONNV infected cells, thus indicating that the pathway is switched on in response to viral infection. These results suggest that there is some constitutive activation of the JAK/STAT pathway in L35 cells that results in a moderate decrease in virus, however the pathway is not switched on in response to viral infection. It is not clear if silencing PIAS, the putative negative regulator of the JAK/STAT pathway, is sufficient to activate the pathway, since markers for JAK/STAT activation are not known in *A. gambiae* mosquitoes.

6.4 Co-infection of *A. gambiae* with ONNV and *P. berghei*

During the outbreak of ONNV in the 1960's, viral transmission occurred across many regions where malaria is endemic. *A. gambiae* is one of the most important vectors of human malaria in the world, and is the only known Anopheline vector of an arbovirus. A paper published in 1962 observed that during the epidemic of ONNV, there was a decrease in the reported malaria cases in a study area in central Uganda [180]. The authors studied parasite numbers in captured mosquitoes, and monitored the number of local children infected with malaria over 12 months. They found during May 1960, when the ONNV epidemic was just starting to spread throughout central Uganda, that the number of mosquitoes harbouring parasites was unusually low (in *A. gambiae*, 0.6% compared to an average of 3.2% for the whole 12 months of the study), considering the ideal conditions for high transmission and abundance of *A. gambiae* and *A.funestus*, the major vectors of malaria in the area. They also noted a sharp drop in the number of school children positive for parasites in the blood during May-November (a drop from 53% in December 1959 to 11% in May and 21% in November 1960). The authors speculate that the development of the parasite may have been inhibited by the virus, either in the mosquito or human host [180]. It is possible that mosquito infection by ONNV, which is lifelong, could trigger immune responses that indirectly attack invading malaria parasites, preventing infection and reducing transmission rates of malaria. In order to investigate this, co-infections of *A. gambiae* with 5'ONNVic-eGFP and the rodent malaria parasite *P. berghei* were carried out. During normal parasite development, the invasive ookinete form of the parasite crosses the midgut of the mosquito, and forms a cessile cyst (oocyst) on the basal laminal. The oocyst matures over 10 days, forming thousands of motile sporozoites which eventually disrupt the oocyst and are released into the hemolymph of the mosquito [181]. Sporozoites invade the salivary glands, and are injected into a new host when the mosquito takes its next blood meal. Even in susceptible mosquitoes, only a small number of ookinetes will successfully form an oocyst. Parasite losses can be attributed to the mosquito innate immune system [62, 126]. Some parasites are lysed and cleared, but a small proportion are coated in melanin as they emerge from the midgut in a

process called melanisation. It is still not clear if the melanisation process directly kills the invading parasite, or simply coats already dead parasites. In some refractory strains of *A. gambiae* such as the L35 strain, all invading parasites are melanised during midgut invasion.

There are a number of hemolymph genes that regulate the responses of the immune system in response to pathogens. Three major immune mechanisms employed by mosquitoes (melanisation, lysis or opsonisation of pathogens) are regulated by these proteins. The co-infection experiments were designed such that parasite invasion of the midgut and exposure to the hemolymph occurred at approximately 4 dpi with 5′ONNVic-eGFP. Therefore it is probable that any differences in immune response to invading parasites are mediated by the genes that are differentially regulated by 5′ONNVic-eGFP at 4dpi. When mosquitoes infected with 5′ONNVic-eGFP are exposed to *P. berghei*, there is a significant decrease in the number of melanised ookinetes and a small decrease in the number of live oocysts in infected midguts. This indicates that in virally infected mosquitoes, there is an inhibition of melanisation and promotion of lytic immune mechanisms. Indeed 2 genes up-regulated by 5′ONNVic-eGFP infection are known inhibitors of melanisation in *A. gambiae*: silencing of both CTLMA2 and CLIPA2 results in massive melanisation of invading parasites [62, 67]. This finding is corroborated by the observation that melanisation is inhibited in mosquitoes when these two genes are up-regulated. There are 4 further CLIPs and 1 CTL that are up-regulated by 5′ONNVic-eGFP infection. These proteins can be hypothesised to have roles in the inhibition of melanisation or the promotion of lysis, and it would be interesting to test these genes for interactions with both viral infection and Plasmodium infection.

A large number of TEPs and LRIMs are differentially regulated at 4dpi. Several TEPs and LRIM1 are complement proteins and some have known functions in the lysis of parasites. LRIM1 silencing in *A. gambiae* mosquitoes results in large increases in live oocyst numbers, demonstrating that LRIM1 functions in the killing of parasites independently of melanisation [126]. TEP1 has a similar phenotype in *A. gambiae* and functions as part of a complex with LRIM1 that targets parasites for lysis [126].

Two further LRIMS (LRIM7 and LRIM10) and 5 TEPs (TEP4/9/10/12/14) are up-regulated at 4dpi and may have similar functions in complement mediated lysis or opsonisation of pathgens. Strikingly there are some key differences in the immune response triggered by viral infection compared to that triggered by Plasmodium infection. For example, LRIM1 is up-regulated but APL1C, which forms a complex with LRIM1 and is often co-regulated with LRIM1, is not up-regulated. This suggests that viral specific pathways may be activated, utilising different components of the immune system to target viral infection than those used to target parasite infection. This may explain why there is only a small reduction in the number of live oocysts in virally infected mosquitoes. The viral responsive genes inhibit melanisation (reducing the number of melanised parasites) and divert signalling to another arm of the immune system that, although slightly decreases oocyst numbers, does not lyse all invading parasites.

Immune responses to the model rodent malaria parasite and the human malaria parasite *P. falciparum* have been shown to be different; as such it is important that these experiments be carried out *using P. falciparum* parasites. In addition to differences in susceptibility to immune responses, *P. falciparum* infections tend to have lower infection loads than *P. berghei* infections carried out in the laboratory, with oocyst numbers per midgut in the field being more in the range of 1-10 depending on field isolates and mosquito strains, rather than the average of ~30 seen in these experiments. This is very important when considering the impact of co-infection on *transmission* of malaria, as only a single oocyst can produce thousands of infective sporozoites. It would also be interesting to investigate any changes in the number of sporozoites in co-infected mosquitoes, as they ultimately determine the outcome of infection in the mosquito.

6.5 Final comments

The transcriptional response of *A. gambiae* mosquitoes to ONNV infection reveals that the mosquito does indeed recognise and respond to viral infection by differentially regulating a number of genes, including many involved in the innate immune system. Some of the differentially regulated genes are in-keeping with the literature on invertebrate anti-viral immunity. Indeed, there are genes from the RNAi pathway and from the JAK/STAT pathway are differentially regulated. However genes from immune signalling pathways shown not to be involved in anti-viral immunity in *Ae. aegypti* are differentially regulated, primarily the IMD pathway. This indicates that *A. gambiae* mosquitoes may utilise different immune mechanisms to target viral infection to other *spp* of mosquitoes, possibly contributing to their poor vectorial capacity. Indeed screening the effect of gene KD for the major immune signalling pathways in invertebrates demonstrated that the Toll pathway, known to be important for anti-viral immunity in *Ae. aegypti* and *D. melanogaster*, is not involved in targeting virus in *A. gambiae*. Additionally the JAK/STAT pathway has little effect on viral infection in *A. gambiae* mosquitoes. Experiments using *A. gambiae* cells revealed that the JAK/STAT pathway is capable of moderately reducing viral infection, although viral infection does not trigger activation of the pathway through translocation of the STAT1 transcription factor. The role of the JAK/STAT pathway in *A. gambiae* anti-viral immunity warrants further investigation. Four anti-viral immunity genes were identified through gene KD in *A. gambiae* mosquitoes. All 4 of these genes have novel roles in mosquito anti-viral immunity and require further investigation to begin to elucidate the mechanism of viral antagonism. It appears that *A. gambiae* utilises some core anti-viral mechanisms, such as RNAi, in combination with possibly species specific anti-viral mechanisms to target ONNV infection.

7 Gene silencing and RNA PAMPs in *A. gambiae* derived cells

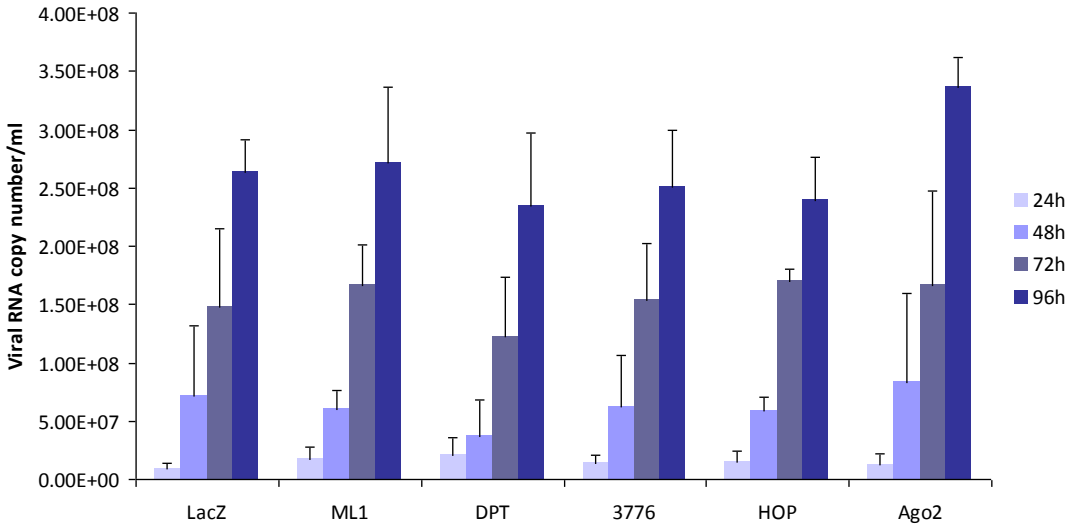
7.1.1 Introduction

The forward genetics approach of transcriptionally profiling virally infected mosquitoes identified many genes with roles in the regulation of ONNV infection. However genes that may have an effect on viral infection are not always differentially regulated upon infection. An example is AGO2, a component of the RNAi pathway (which has a central role in regulating viral infection), is not transcriptionally induced upon viral infection. It is assumed that components of the RNAi pathway are already present at sufficient levels for efficient RNAi to take place. An alternative method for identifying genes such as AGO2 is a reverse genetics approach. Although a small scale reverse genetic RNAi screen was carried out in adult mosquitoes, cell based assays are essential for developing a high throughput RNAi screen. This would allow rapid identification of anti-viral genes, as well as genes associated with the replication of the virus. These assays would require an easily measured output, for example GFP expression from the 5'ONNVic-eGFP infectious clone, and high throughput methods for conducting genome wide RNAi. In this chapter, the testing of two independent methods of gene silencing are described. Observations made during these experiments led to the investigation of foreign RNA as a PAMP in *A.gambaie*.

7.2 Developing cell based RNAi assays to detect viral agonists and antagonists

With a view to developing a genome wide RNAi screen, a protocol was designed for identifying viral antagonists using the L35 cell line and 5'ONNVic-eGFP. Two independent methods for knocking down mosquito genes in *A. gambiae* cell lines were tested; transfection of dsRNA into cells and bathing cells in dsRNA. In initial experiments 5 genes were selected for gene knock down via transfection of dsRNA in L35 cells; AGO2 – the known viral antagonist was selected as a control, 3 genes up-regulated by viral infection identified by transcriptional profiling (ML1, DPT and AGAP003776 – referred to as 3776), and HOP, a known viral antagonist in other mosquito species. Cells were transfected with dsRNA and infected 48 hours later

A



B

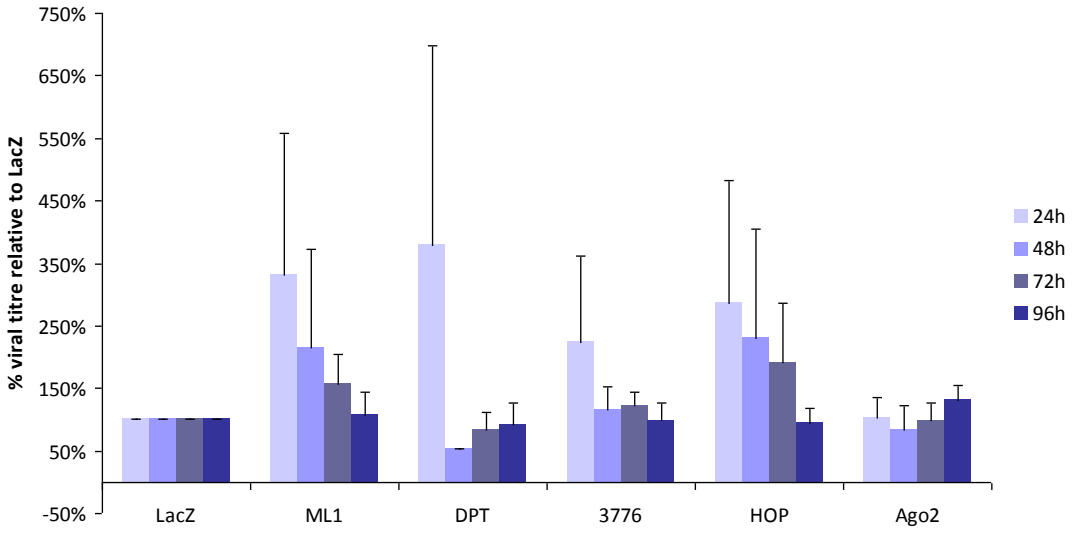


Figure 33. Viral RNA copy number/ml in conditioned media of gene KD L35 cells. Cells were transfected with dsRNA complementary to 5 genes of interest. 48 hours post tansfection cells were infected with 1MOI of P1/P2V 5′ONNVic-eGFP. Conditioned media was collected at 24, 48, 72 and 96 hpi. Viral RNA genome copy number was ascertained using qrt-PCR. **A)** Viral genome copy number/ml. **B)** Viral RNA genome copy number represented as a percentage of the infection level observed in the LacZ control. Error bars represent SD of 3 biological replicates

with 1MOI of 5'ONNVic-eGFP. Conditioned media containing virus was collected at 24, 48, 72 and 96h post infection, and viral RNA genome copy number/ml was calculated using qrt-PCR. Three biological replicates were carried out. Figure 33 A shows the viral RNA copy number/ml for each of the 5 gene knock downs. In order to view the differences in viral load within replicates, the viral RNA copy number/ml was normalised to the LacZ (non-specific RNA control) removing the variation in overall level of infection from replicate to replicate (figure 33B). The results show great levels of variation (as experienced with infection *in vivo*), however ML1 and HOP both show increased viral titres at multiple time points. Curiously no difference in viral genome copy number was observed in AGO2 depleted mosquitoes. The effectene transfection protocol was specifically optimised for use with L35 cells, and further transfections were carried out with dsRNA complimentary to the ML1 and HOP and GALE8 genes. Higher, although not significantly so, levels of 5'ONNVic-eGFP infection were observed in both HOP and ML1 KD cells (see figure 34). GALE8 KD had no effect on viral infection.

7.2.1 Treatment of *A. gambiae* cells with non-specific exogenous RNA dramatically impairs the ability of 5'ONNVic-eGFP to infect and/or replicate in *A. gambiae* cells

With a view to developing a genome wide RNAi screen in *A. gambiae* cells, a second method of gene silencing that would increase throughput of analysis was tested. In this method, cells are bathed in dsRNA for 48h prior to infection. Initially 2 genes were silenced prior to infection; AGO2, a known ONNV antagonist was used as a positive control; HOP, the kinase in the JAK/STAT pathway. Treatment of cells with exogenous dsRNA had a dramatic inhibitory effect on viral replication in all dsRNA treated conditions (see figure 35).

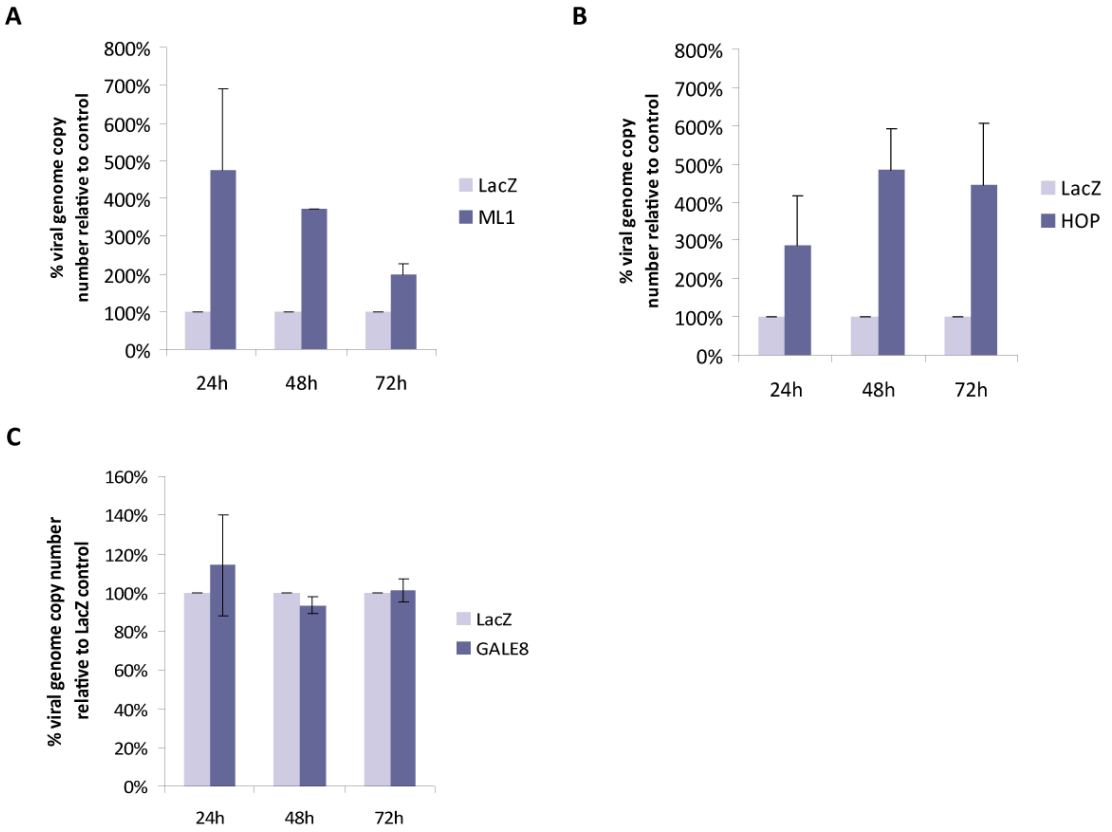


Figure 34. Viral genome copy number in A) ML KD and B) HOP KD and C) GALE8 KD L35 cells. Viral genome copy numbers/ml was calculated using qrt-PCR, and is represented as a % of the viral titre observed in the LacZ (non-specific dsRNA) control. Four biological replicates were carried out for HOP KD, 2 biological replicates were carried out for ML1 and 3 biological replicates were carried out for GALE8. Error bars represent standard error.

Bathing cells in dsRNA resulted in a dramatic reduction in viral titre, importantly this phenotype was not reversed by silencing of the RNAi pathway. To further investigate this cells were treated for 48 hours with non-specific LacZ dsRNA prior to infection with 1MOI of 5'ONNVic-eGFP. Viral genome copy number/ml was calculated from conditioned media from treated and untreated cells, collected at 24, 48, 72 and 96 hpi. Fluorescent microscopy clearly showed a reduction in the number of GFP expressing cells in dsRNA treated compared to untreated cells (Figure 36A). Viral genome copy number was significantly lower in conditioned media from cells treated with dsRNA at 48 and 72hpi (T Test P values of * $P<0.05$ and ** $P<0.001$) (Figure 36B).

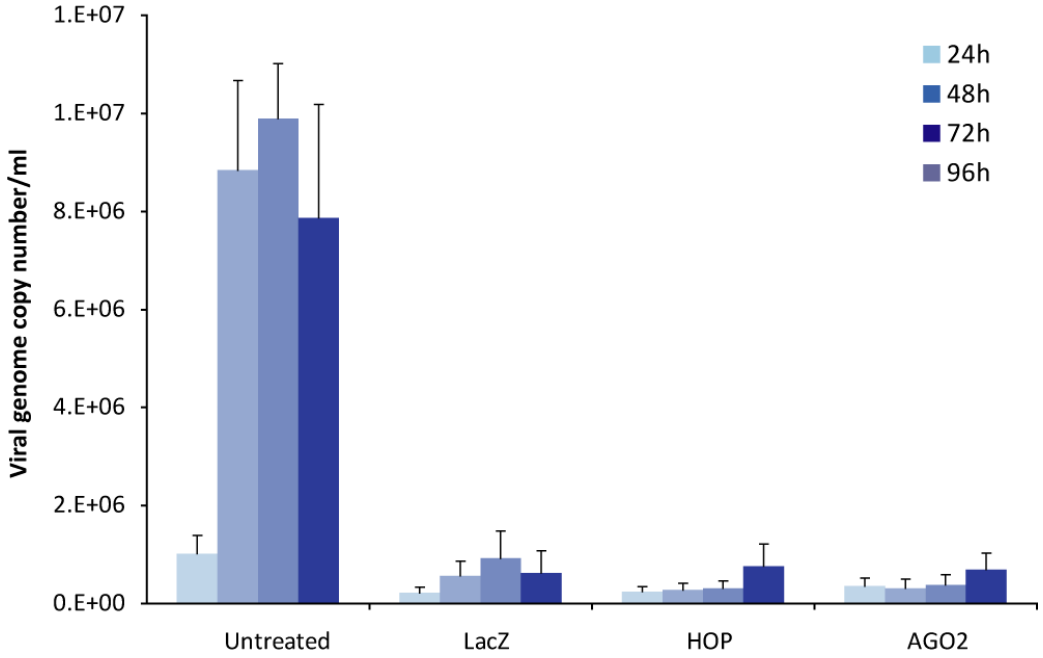


Figure 35. Effect of 'bathing' method of gene silencing on 5'ONNVic-eGFP infection in L35 cells. Cells were untreated or 'bathed' in dsRNA corresponding to LacZ, HOP or AGO2 for 48h prior to infection with 1MOI P1/P2V 5'ONNVic-eGFP. Viral genome copy number/ml in conditioned media was ascertained using qrt-PCR at 24, 48, 72 and 96 hpi. Error bars represent standard error of 3 biological replicates.

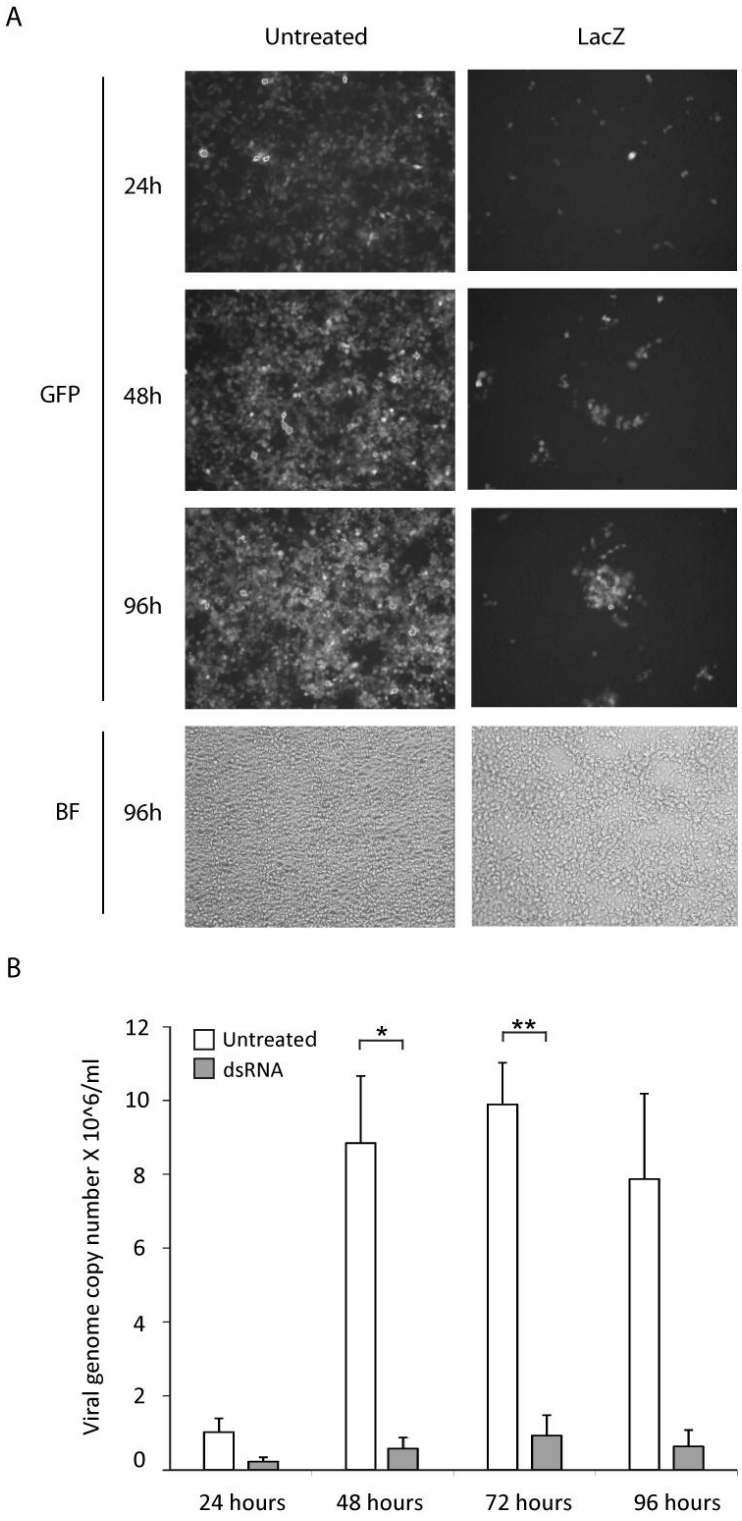


Figure 36. Non-specific dsRNA inhibits viral infection in *A. gambiae* cells. Cells were treated with LacZ dsRNA in conditioned media for 48h prior to infection with 1MOI 5′ONNVic-eGFP **A**) GFP (white) in cells treated with LacZ dsRNA compared to untreated cells. **B**) Viral genome copy number/ml from conditioned media of cells treated with dsRNA compared to untreated cells at 24, 48, 72 and 96 hpi. Error bars represent standard error of 3 biological replicates. *P<0.05, **P<0.001 using students T-test

7.2.2 Treatment of cells with ssRNA

To see if the response to RNA was specific to dsRNA, the experiment was repeated using ssRNA. ssRNA specific to the LacZ was generated from a linearised plasmid containing the full LacZ sequence. Cells were bathed in ssRNA for 48h prior to infection with 1MOI 5′ONNVic-eGFP. Viral RNA genome copy number/ml was ascertained using qrt-PCR of the viral E2 gene at 24, 48 and 72hpi (see figure 37). A similar phenotype to dsRNA treatment was observed, with ssRNA treatment significantly reducing the viral titre of conditioned media (T test P values of * P<0.05 and ** P<0.001).

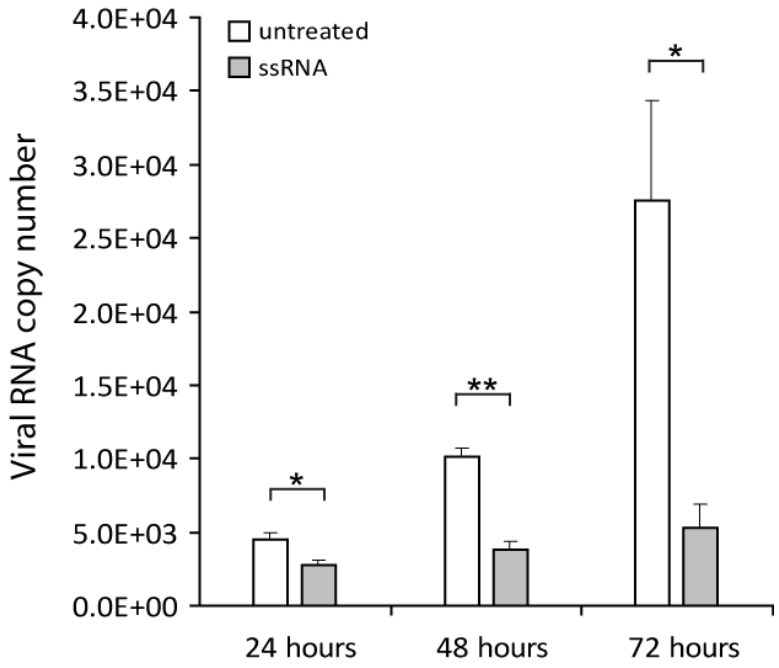


Figure 37. Non-specific ssRNA inhibits viral infection in *A. gambiae* cells. Cells were treated with LacZ ssRNA in conditioned media for 48h prior to infection with 1MOI 5′ONNVic-eGFP. Viral genome copy number/ml was calculated from conditioned media of cells treated with ssRNA compared to untreated cells at 24, 48, 72 and 96 hpi. Error bars represent standard error of 3 biological replicates. * P<0.05, **P<0.001 using students T-test.

7.2.3 Transcriptional responses of *A. gambiae* cells to treatment with non-specific dsRNA, and non-specific ssRNA

To investigate the cause of the dramatic impairment of virus infection and/or replication in RNA treated cells, the transcriptional responses of *A. gambiae* cells to ds/ssRNA treatment were characterised. Cells were treated for 48h with non-specific LacZ dsRNA, ssRNA or water as a control. Total RNA from treated cells was collected in TRIzol and extracted as described in material and methods. Total RNA was used to generate labelled probes for hybridisation to the aglient 4 X 44K array (see section 3.15). Hybridised slides were washed to remove unbound labelled RNA and were scanned using a GenePix semiconfocal microarray scanner (AXON Instruments, Foster city, CA). Gene Pix Pro 6.1 was used to record feature signal intensity, to eliminate local backgrounds, for grid alignment and manual inspection of feature quality. Average feature diameter was calculated and features lying outside two standard deviations of the mean were excluded from analysis. The ratio of feature intensity verses local and global backgrounds were calculated and features not exceeding background intensities were excluded from analysis. Features were normalised using Genespring 6.1 (AXON Instruments, Foster city, CA) by locally weighted linear regression methods (Lowess). Feature intensities over the three biological replicates were averaged. T-test p-values were calculated, and normalised data was filtered to exclude data with p-values greater than 0.05. Data was further filtered to include only genes showing 2-fold and greater regulation.

A. gambiae cells showed significant transcriptional responses to both ds and ssRNA; 227 genes were differentially regulated by treatment with exogenous dsRNA and 141 genes were differentially regulated by ssRNA. Of these genes, 97 were regulated by treatment with both ds and ssRNA (figure 38). These genes were clustered into three groups – those regulated by ssRNA alone, those regulated by dsRNA alone, and those regulated by both ds and ssRNA.

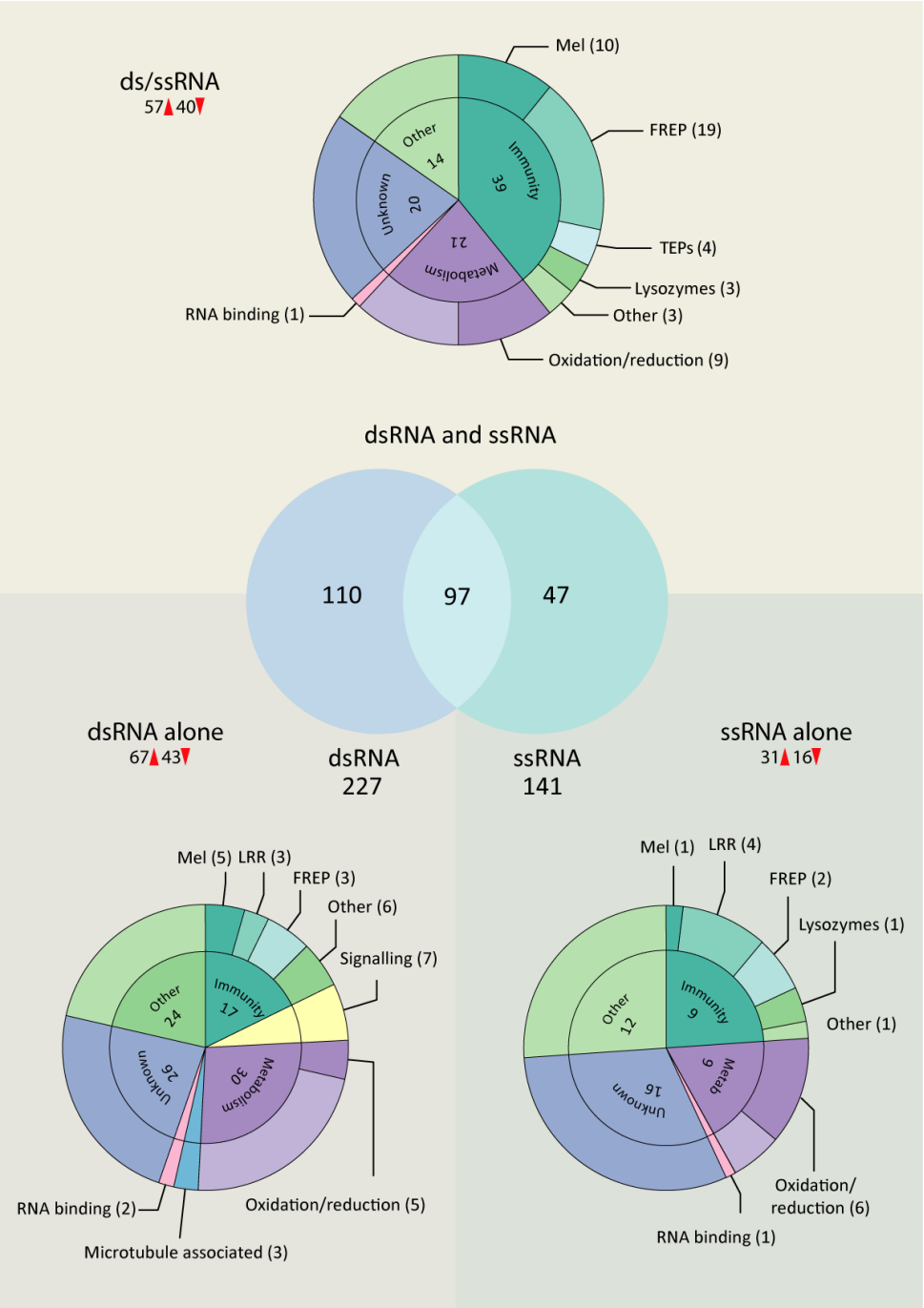


Figure 38. Transcriptional responses to treatment of *A. gambiae* cells with ds or ssRNA compared to water. The transcriptional responses of *A. gambiae* cells to ss and dsRNA were profiled using 4X44K Agilent RNA microarrays. Gene lists include only features that pass strict criteria outlined in figure 20. Genes included the analysis are 2-fold or greater regulated, with T-test P values of <0.05. Genes were categorised based on gene ontology and orthologs in other insects. Red arrows indicate direction of regulation. Mel, Melanisation; LRR, Leucine Rich Repeat.

7.2.4 Transcripts regulated by ssRNA only

Fourty seven genes are differentially regulated specifically by ssRNA alone. Nine genes associated with metabolism are regulated, 6 of which are associated with oxidation/reduction (5 of which are up-regulated). Sixteen genes have no predicted gene architecture and/or putative function. A putative RNA binding protein is up-regulated; it contains a putative dsRNA binding domain, however its function is not known. Nine immunity related genes are differentially regulated. Of these only 4 are up-regulated by ssRNA treatment; 2 LRR proteins, and 2 Clip domain serine proteases. Down-regulated immunity genes include 2 Fibrinogen-like proteins (FREPs), a further 2 LRR proteins and a lysozyme. Other differentially regulated genes have diverse putative functions including protein-protein interaction, sugar transport, DNA repair, and a cuticle constituent. Appendix 2 Table 3 gives a full list of genes regulated by ssRNA treatment.

7.2.5 Transcripts regulated by dsRNA only

A much larger number of genes are transcriptionally regulated by dsRNA alone; 110 genes, 67 of which are up-regulated, and 43 are down-regulated. Thirty regulated genes have diverse putative functions in metabolism, 5 are involved in oxidation/reduction. Two RNA binding genes are up-regulated. One is a putative inhibitor of translation. The other is a Vasa-like RNA helicase, which may function in unwinding of dsRNA. Interestingly 7 genes with roles in signalling are differentially regulated, including a tyrosine phosphorylation-regulated kinase (up-regulated), a tyrosine phosphatase (down-regulated), a putative transcription factor (up-regulated), a G-protein (up-regulated) and a protein with multiple predicted domains including the integrin beta subunit, and EGF-like region and a Von Williebrand factor domain (up-regulated). Seventeen immunity regulated genes are differentially regulated, 10 of which are down-regulated. The most highly up-regulated immunity related gene is an LRR protein that contains a predicted transmembrane domain (TM) and encodes 2 possible Immuno Tyrosine Activating Motifs (ITAMs) in the

cytoplasmic tail of the protein, indicating that it may be an immune receptor that can signal upon activation. A further LRR gene is up-regulated, along with an MD-2 like protein ML5, 2 serine protease inhibitors SRPN5/16, a peptidoglycan recognising protein (PGRP-LB1) and a sugar binding protein (CTLMA4). *D. melanogaster* PGRP-LB is a Gram negative antibacterial protein that is regulated by the IMD pathway, and itself down-regulates the IMD pathway in a negative feedback to regulate immune responses[182]. CTLMA4 is a C-type lectin that has also been shown to play a role in defence against Gram negative bacteria in *A. gambiae[183]*, indicating that dsRNA triggers the IMD pathway. However, DEF1, ortholog of DmDefensin-1 (which is regulated by both the Toll and the IMD pathways in *Drosophila*) is down-regulated by dsRNA treatment. Two prophenoloxidases and three FREPs are down-regulated. Additionally a lysozyme, a scavenger receptor, LRIM1 and SRNP6 are down-regulated. Appendix 2 table 2 gives a full list of genes differentially regulated by dsRNA.

7.2.6 Transcripts regulated by ds and ssRNA

Ninety seven genes are differentially regulated by ds and ssRNA. Fifty seven of these are up-regulated and 40 are down-regulated. Of the up-regulated genes, 20 have no putative function and/or predicted genes architecture. Nineteen genes have putative functions in metabolism, nine with oxido-reductase activity including a glutathione-S-tranferases and two cytochrome P450s. Eight putative immunity genes are up-regulated, 7 of which have roles in the melanisation cascade and regulation of melanisation; 5 clip domain serine proteases (CLIPs) are regulated (A4/A8/A14/B5/B13), a super-oxide dismutase (SOD1) (known to reduce melanisation of parasites in mosquitoes[184]), and CTLMA1. The final immunity gene is GALE8. The remaining up-regulated genes have diverse functions, including a nucleic acid binding protein with a Tudor domain (found in Tudor-SN, a component of the RNAi pathway in *D. melanogaster*).

Of the 40 down-regulated genes, the majority are putative immunity genes (31). 19 of these are Fibrinogen-like proteins (FREPs). Four TEPs are down-regulated, along with 2 anti-microbial peptides (CEC1/3) and three lysozymes (LYSC1/2/5). Three pro-phenoloxidases (PPOs; enzymes in a cascade that produces melanin during melanisation) are also down-regulated. Appendix 2 Table 1 gives a full list of genes differentially regulated by ds and ssRNA.

7.2.7 Overlap in gene expression between viral responsive genes and ds/ssRNA regulated genes

The genes differentially regulated by ds and ssRNA were compared to virally responsive genes (Table 8). Five genes are differentially regulated in the same direction by all 3 conditions, and 1 is differentially regulated by ssRNA and 5'ONNVic-eGFP infection. Two of these genes are putative immunity genes; SRPN5 and GALE8. GALE8 was found to be an antagonist of ONNV infection in G3 mosquitoes. The other genes include a transport protein that putative functions in transporting ligands between membranes, a dehydrogenase-like hydrolase that has a putative function in metabolism, and a sugar transporter that belongs to the Major Facilitator Superfamily (MFS) (a family of transporter capable of moving small solutes in response to chemiosmotic ion gradient.

Eleven genes are differentially regulated in the opposite direction by ds/ssRNA treatment and 5'ONNVic-eGFP (Table 8), including 7 immunity genes, indicating that viral infection triggers distinct signalling to RNA recognition.

Table 8. Transcripts differentially regulated by ds/ssRNA treatment and 5′ONNVic-eGFP infection.

AGAP	Common name/function	dsRNA	ssRNA	Fold change ratio ONNV Day 1	ONNV Day 4	ONNV Day 9
Genes regulated in the same direction						
AGAP003733	Cellular retinaldehyde binding/alpha-tocopherol transport	2.217	n/a		2.11	
AGAP004954	Haloacid dehalogenase-like hydrolase	2.622	2.215		2.91	
AGAP009145	unknown	2.156	n/a		2.06	
AGAP009221	SRPN5	2.127	n/a		2.15	
AGAP012529	GALE8	4.57	3.866111		2.08	
AGAP007753	sugar transporter integral to membrane		2.652		3.02	
Genes regulated in the opposite direction						
AGAP000694	CEC3	0.34633	0.398		2.04	
AGAP002198	Glycine N-methyltransferase	0.421	n/a		2.7	
AGAP006278	Insect pheromone/odorant binding protein PhBP	0.489	n/a	2.4	2.1	
AGAP007343	LYSC2	0.3873	0.424		2.08	
AGAP008578	Vasa-like ATP dependent RNA helicase	2.122	n/a		0.4	
AGAP006348	LRIM1	0.495	n/a	2.2835		
AGAP010814	TEP17	0.406	0.493	2.601		
AGAP010774	FREP27	0.415	n/a		2.13	
AGAP010819	TEP10	0.401	0.423		2.10	
AGAP010830	TEP9	0.345	0.41	2.3		
AGAP011333	AMP dependent synthase and ligase		2.048		0.50	

7.2.8 Confirmation of microarray analysis using qrt-PCR

In order to confirm the results seen in the microarray analysis, an independent method was used to ascertain the differential regulation of several genes found to be 2-fold or greater regulated by ds and ssRNA treatment. Qrt-PCR was used to confirm the differential regulation of CLIPA8, CEC1, CEC3 and DEF1 in *A. gambiae* cells. The two independent methods gave similar fold-changes in transcripts for all genes tested for dsRNA treatment (figure 39A) and ssRNA treatment (figure 39B). Additionally the transcriptional profile of three genes, (CEC1, CEC3 and DEF1) in response to dsRNA treatment, were confirmed using a different *A. gambiae* cell line (figure 39C).

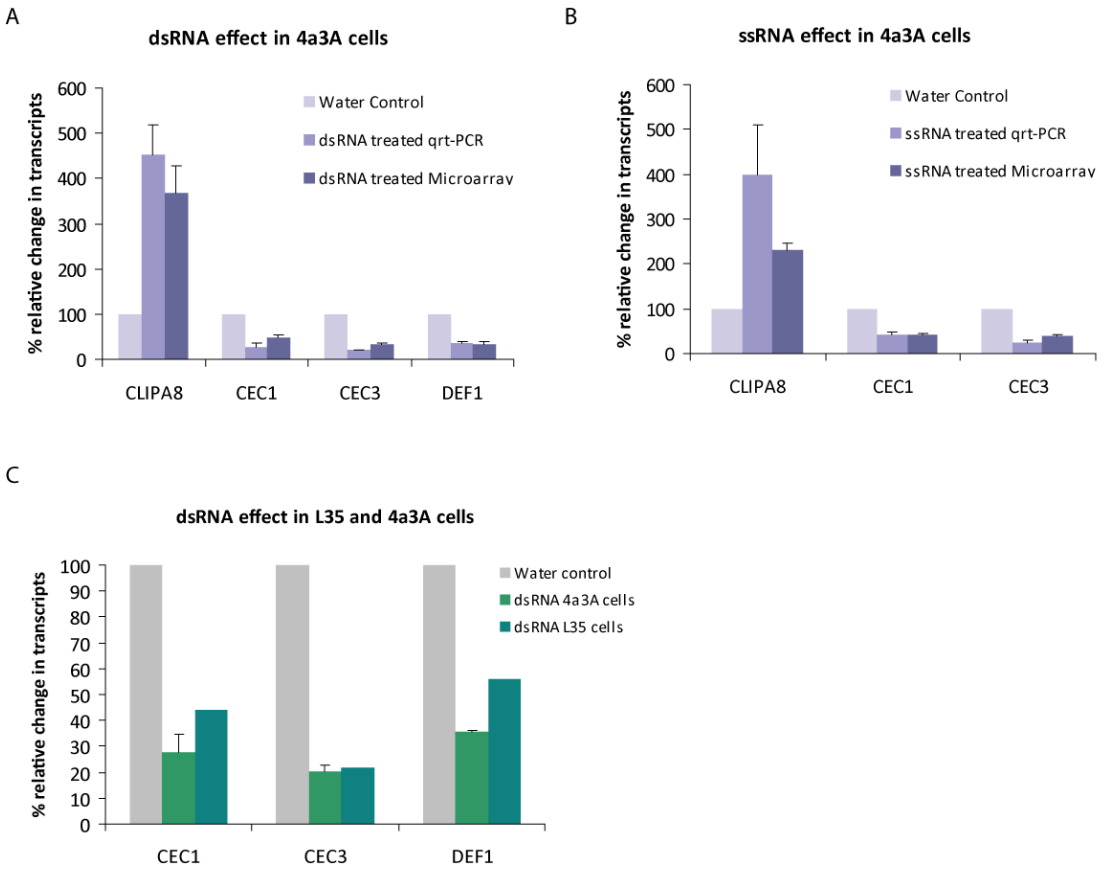


Figure 39. Confirmation of transcriptional responses to RNA using qrt-PCR. Relative transcripts levels (compared to controls) were calculated in cells **A)** treated with dsRNA **B)** treated with ssRNA. **C)** Conformation of transcriptional responses to dsRNA in 2 different *A. gambiae* derived cell lines.

7.2.9 Effect of non-specific dsRNA treatment in *A. gambiae* mosquitoes

Considering the dramatic effect of non-specific dsRNA on viral infection in *A. gambiae* cell lines, the effect of non-specific dsRNA treatment in adult G3 mosquitoes was investigated. Mosquitoes were inoculated with dsRNA or sterile PBS 4 days prior to infection with 5'ONNVic-eGFP. Viral RNA genome copy number/mosquito was calculated using qrt-PCR against the viral E2 gene (figure 40). Injection of LacZ dsRNA did not dramatically reduce viral infection as observed in the L35 cell line. Viral titre was slightly decreased at 1 dpi, however infection in LacZ dsRNA treated mosquitoes at 3 and 6 dpi was higher than in PBS treated mosquitoes, although this was highly variable from between replicates and not statistically significant.

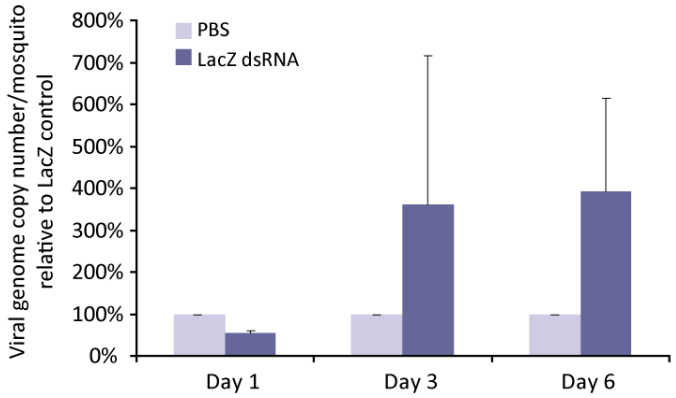


Figure 40. Effect of non-specific dsRNA treatment on 5'ONNVic-eGFP infection in adult mosquitoes. Mosquitoes were injected with either PBS or dsRNA specific to the LacZ gene 4 days prior to infection with 5'ONNVic-eGFP. Viral genome copy number per mosquito was calculated using qrt-PCR at 1, 3 and 6 dpi. Error bars represent standard deviation of 3 biological replicates.

7.2.10 Summary

This chapter has covered:

- Silencing of ML1 and HOP but not GALE8 increases viral replication in *A. gambiae* cells
- Treatment of *A. gambiae* cells with exogenous non specific RNA impairs the ability of 5'ONNVic-eGFP to infect and replicate
- *A. gambiae* cells transcriptionally respond to both ds and ssRNA treatment
- Qrt-PCR confirmation of microarray analysis; transcriptional responses were confirmed using an independent method and additionally using a different *A. gambiae* derived cell line
- Non-specific dsRNA does not inhibit viral replication in adult mosquitoes

7.3 Discussion

The development of a high throughput genome-wide RNAi screen would greatly facilitate the identification of genes with roles in regulating viral infection in *A. gambiae*. Any method of silecing genes using the RNAi pathway must include the introduction of dsRNA against genes of interest into cells to initiate gene silencing. In mammalian systems dsRNA and ssRNA are potent danger signals. They are recognised or sensed by multiple receptors and proteins (RIG-1, MDA5, LGP-2, TLR-3), resulting in the activation of anti-viral mechanisms [185]. Non-self ds and ssRNA are believed to be viral specific PAMPs, and are recognised both intra- and extra-cellularly. The mammalian response to RNA PAMPs is highly discriminative; the cytoplasm of cells are full of host RNA, however signalling only occurs in infected cells [185]. Viral PAMPs recognised by RIG-1 include viral genomic RNA, Viral transcripts and replication intermediates, non-capped RNA and Pol III transcribed RNA [185]. Although RNAi is an established anti-viral mechanism in invertebrates, which recognises and destroys foreign dsRNA, other mechanisms for sensing ds/ssRNA, like those seen in mammalian systems, have not been described. In this

section the development of RNAi screens in *A. gambiae* cells and the response of *A. gambiae* cells to foreign ds and ssRNA are discussed.

7.3.1 Testing genes in *A. gambiae* derived cell lines; a view to developing a genome wide RNAi screen

A secondary objective of this project was to develop protocols for carrying out a genome wide RNAi screen to identify viral antagonists/agonists utilising the GFP marker inserted into 5′ONNVic-eGFP, and *A. gambiae* cell lines. During this study a method to assay the effect of gene knock down (via transfection of dsRNA) on viral titres in L35 cells was developed, however, experiments using a plate reader to quantify GFP expression as a marker for infection were unsuccessful in differentiating between high and low levels of infection (data not shown). This is due to the nature of the marker – cells infected with 5′ONNVic-eGFP express GFP shortly after infection, however the GFP expression is a qualitative rather than a quantitative marker, and does not reflect the viral genome copy number generated during infection. Thus, GFP can only be used to identify differences in very early infection i.e. if silencing a gene results in 100% of cells becoming infected compared to 20% of cells in a control, or blocks the entry of virus into cells, preventing infection. However, silencing a gene that limits, but does not prevent, infection would show no/very little difference in GFP expression. As such, in this study, the viral genome copy number in conditioned media was used as a measure of viral infection of L35 cells, and was calculated using qrt-PCR.

Two genes were found to be viral antagonists in L35 cells, ML1 – an MD-2 like receptor discussed previously, and HOP – the Janus kinase in the JAK/STAT pathway. HOP KD in *A. gambiae* mosquitoes also showed some interesting, however inconsistent, results *in vivo*, and as such was not taken forward for detailed analysis by plaque assay. In view of the effect on viral replication in cells, it would be interesting to carry out plaque assays of HOP KD mosquitoes to see if the function is maintained *in vivo*. Due to the high variation between replicates observed in viral

infections, and the low number of replicates carried out, neither HOP nor ML1 KD resulted in statistically significantly increased viral titres using a students T-test, however carrying out T-tests with such a small number of data is not appropriate. Curiously silencing of AGO2 had no effect on viral replication in L35 cells. This was unexpected seeing as AGO2 is a known viral antagonist in *A. gambiae* mosquitoes and *A. gambiae* L35 cells do have a functioning RNAi pathway. It is possible that AGO2 is redundant and silencing this gene does not prevent RNAi in L35 cells. To test if RNAi is capable of limiting viral infection in L35 cells, further members of the RNAi pathway would need to be silenced.

In order to increase the throughput of gene KD assays in cells, a second method of gene silencing was investigated. Cells were bathed in dsRNA for 48h prior to infection. Treating cells with dsRNA, both mosquito specific and non-specific, led to a dramatic decrease in viral infection. This reduction of viral infection was not reversed by silencing AGO2, indicating that the virus is suppressed by a mechanism independent of RNAi (however as previously mentioned, the ability of RNAi to limit viral infection in L35 cells has not been demonstrated to date). The observation of this dramatic phenotype led to the question of whether non-specific dsRNA, and additionally ssRNA, can act as a danger signal in mosquito cells.

In mosquitoes there are a wide variety of immune responses that vary in their specificity. There are broad spectrum responses, for example LRIM1 and several TEPS transcriptionally respond to diverse pathogens (bacterial, fungal and parasitic) (Leanna Upton, unpublished data), and there are more specific responses to individual pathogens, such as the Toll pathway responding to gram positive bacteria and fungal infection, and the IMD pathway responding to gram negative bacteria. RNAi demonstrates a still more specific response, which targets individual and specific dsRNA sequences from a single virus.

In mammalian cells, dsRNA and ssRNA are potent triggers of innate immune responses, with multiple mechanisms for sensing ds/ssRNA and downstream signalling. Unlike RNAi, the induction of innate immune signalling by ds/ssRNA in

mammalian cells is a non-specific response to foreign RNA, where foreign RNA is acting as a danger signal. Intracellular dsRNA can be recognised by Protein kinase R (PKR), binding of which induces the phosphorylation of the eukaryotic translation initiation factor eIF2 and the inhibition of translation [186]. dsRNA is also recognised by 2'5'Oligoadenlyate synthetase, leading to the activation of RNAse L, which cleaves dsRNA [187]. Intracellular ds/ssRNA is recognised by RIG-1 and the RIG-1 like helicases (RLHs), MDA5 and Lgp2, all of which lead to signalling via CARD-CARD interactions and trigger NF-kB signalling, regulating the expression of a plethora of cytokines, interferon and potent anti-viral innate immune responses [185]. Lastly dsRNA is also recognised within endosomal vesicles of cells by Toll-like receptor 3 (TLR-3). TLR-3 signals through intracellular TIR domains, recruiting TRIF, and culminating in the activation of NF-kB, IRF3 and IRF7 signalling [188]. The ligands, or PAMPs, that are recognised by these proteins can be derived from viral transcripts and replication intermediates, viral genomic RNA, uncapped transcripts and Pol III transcribed RNA [185].

It has been suggested that in *D. melanogaster*, Dicer-2, despite lacking CARD domains, may still have a signalling function downstream of RNAi. Vago is a gene expressed in a Dicer-2 dependent manner in *D. melanogaster*, and has an anti-viral function in the fat body against Drosophila C virus (DCV) infection [189]. The helicase domain of Dicer-2 is required for Vago expression [189], and may have a secondary function in Dicer-2 downstream signalling, mimicking the function of the RLH's to which it's helicases domain is closely related [190]. The signalling events leading to Vago expression remain uncharacterised. In mosquitoes, the RNAi response is known to respond to non-self dsRNA, however, alternative mechanisms for sensing foreign RNA are little known. Whether mechanisms exist that can sense RNA PAMPs through pathways other than RNAi in invertebrates, is not clear.

The transcriptional responses of an *A. gambiae* cell line, 4a3A, to dsRNA and ssRNA were profiled using microarrays analysis. Both ds and ssRNA elicited a significant transcriptional response in *A. gambiae* cells, including a large number of putative

immunity genes. Ninety seven of these genes were regulated by both ds and ssRNA, however distinct groups of genes are regulated by dsRNA alone, and ssRNA alone.

7.3.2 Genes regulated by ds and ssRNA

The treatment of L35 cells with both dsRNA and ssRNA dramatically impairs the ability of ONNV to infect and/or replicate within cells. This suggests that the genes involved in limiting viral infection are responsive to both ds and ssRNA. These genes are discussed below.

7.3.2.1 Melanisation

Of the immunity genes up-regulated by both dsRNA and ssRNA, nearly all have roles in the melanisation cascade. Five CLIPs are up-regulated. The CLIPs are serine proteases, so called because of a 'paper-clip'-like fold formed by disulphide bridges present in the protein family [191]. In *A. gambiae* there are 54 CLIPs, divided into several sub-families. The CLIPAs are serine protease homologs (SPHs) where the catalytic triad of the enzyme has been mutated and is no longer functional [124]. The CLIPBs are functional serine proteases. The CLIPs have functions in hemolymph coagulation, anti-microbial peptide synthesis and melanin synthesis in invertebrates [192]. In *D. melanogaster* CLIPs are known to activate the Toll pathway; Eater, Snake and Persephone (3 CLIPs) are involved in the cleavage of Spatzle, the cleaved form of which is the ligand for the Toll pathway [191]. Additionally CLIPs have a role in regulating the melananotic cascade. Melanisiation is an important immune reaction in invertebrates. It involves a cascade of serine proteases that culminates in the cleavage of pro-phenoloxidase (PPO) into phenolxidase (PO). PO is the first enzyme in a cascade that produces melanin: a compound that is deposited on the surface of pathogens & foreign objects, or around damaged tissue during wound healing [193]. During the production of melanin, highly reactive and toxic intermediates are produced [193]; as such tight regulation of the melanotic cascade is required. In *A. gambiae*, CLIPA8 (up-regulated by both ds and ssRNA) is a known positive regulator of melanisation of *P. berghei* ookinetes [67] and is also required for PPO activation

after bacterial infection [194]. The other up-regulated CLIPs do not have described roles in *A. gambiae*, however it seems that ds/ssRNA alters the regulation of melanisation. CTLMA1 is also up-regulated; CTLMA2 inhibits the melanisation of *P. berghei* parasites [62], however if CTLMA1 has a similar function it is unknown. In addition to the differential regulation of several CLIPs, three PPOs are down-regulated by ds/ssRNA treatment, indicating that the melanisation cascade is inhibited by ds/ssRNA treatment. Considering the poor characterisation of the majority of the CLIPs and other components of the melanisation cascade, it is not clear how ds/ssRNA treatment is altering the regulation of melanisation.

7.3.2.2 Redox status

Nine genes up-regulated by ds/ssRNA have oxido-reductase activity, including a glutathione-S transferase, two cytochrome P450s and a flavin containing mono-oxygenase (FMO). Two aldo-keto reductase-like enzymes are up-regulated; these enzymes are NADPH dependent, catalyse the reduction of a variety of carbonyl compounds and have diverse cellular functions [195]. Another NAD(P) binding enzyme with oxido-reductase activity is also up-regulated.

Gluathione-S transferases are a family of enzymes that have a wide variety of functions. Much of the work on mosquito glutathione-S transferases has focused on insecticide resistance, demonstrating glutathione-s transferases to have an import role in maintaining the redox status of cells, and protecting against harmful reactive oxygen species [196]. FMOs are flavo-enzymes that require FAD, NADPH and oxygen for catalysis [197]. They function within a catalytic redox cycle much like the cytochrome P450s. FMOs catalyse the oxidation of nitrogen, sulphur, phosphorous or selenium in normal metabolism, and in the metabolism of xenobiotics [197]. Cytochrome P450s similarly catalyse the oxidation of a variety of compounds, including xenobiotics such as sulphur and nitrogen containing insecticides [197] and are known to be involved in resistance to insecticides in a variety of insects [198]. The up-regulation of several enzymes capable of detoxifying xenobiotics and reactive

oxygen species indicates that ds/ssRNA treatment may be inducing oxidative stress in cells. A further detoxifying exzyme, superoxide dismutase (SOD) is also up-regulated by both ds/ssRNA, further supporting this idea.

The respiratory burst is a well characterised immune response in mammalian systems that is triggered in activated macrophages; during phagocytosis a large number of ROS are released by macrophages that are toxic to invading microorganisms. In *D. melanogaster* the oxidative burst has been shown to be indispensable for protection against microbes in the gut [199]. The redox status of mosquitoes has a known effect on the immune reactions to parasite infection; a strain of *A. gambiae* that is refractory to *P. berghei* infection (L35) has been shown to be in a state of heightened oxidative stress compared to susceptible strains [184]. This strain of *A. gambiae* melanises all ookinetes as they cross the midgut. Microarray analysis identified several genes involved in redox metabolism as being up-regulated in the refractory strain. Additionally the refractory stain was a) seen to have higher levels of reactive oxygen species (ROS) after blood feeding than susceptible mosquitoes, b) that treatment with anti-oxidants reversed the melanisation of ookinetes and c) that SOD expression was higher than in susceptible mosquitoes [184]. It is not clear whether the heightened oxidative stress directly targets the invading parasites, which are then melanised, or whether it speeds up the immune responses of the mosquito.

There is limited evidence that viruses stimulate the respiratory burst; incubation of human neutrophils with the influenza virus stimulates the production of ROS and increased oxygen metabolism [200]; DENV infection in mouse spleen induces the production of hydrogen peroxide and superoxide anion [201]; in plants, infection of Tobacco with tobacco mosaic virus (TMV) stimulates the release of ROS [202], and additionally previous infection of Tobacco with TMV leads to higher levels of SOD during a second infection resulting in resistance to infection[203]. ds/ssRNA may be recognised as a viral PAMP, triggering oxidative stress, which in turn may target viral infection in cells. It would be interesting to test susceptibility to viral infection in the

strain of *A. gambiae* that has been shown to be in a state of heightened oxidative stress.

7.3.2.3 Fibrinogen-like proteins (FREPs)

Of the genes down-regulated by ds/ssRNA treatment, almost half (19) are FREPs. The FREPs are a family of PRRs common to vertebrate and invertebrates. In *A. gambiae* there has been significant species specific expansion of the family, with 59 FREPs, compared to 14 in *D. Melanogaster* [204]. In invertebrates the FREPs have been implicated in innate immunity [134, 205-207]. A comprehensive study of the protein family in *A. gambiae* revealed strong correlation between expansions and chromosome location with immune responsiveness of the FREPs [204]. Different FREPs respond to bacterial, fungal and parasitic infection, demonstrating a wide repertoire of recognition capability.

There are five different clusters of FREPs on several chromosomes [204] – the FREPs down-regulated by ds/ssRNA are from all 5 clusters, and represent genes differentially regulated by *E.coli*, *S.aureus*, *P. berghei*, *B.bassiana* and *P. falciparum*, in both directions i.e. there does not seem to be any immune responsive specificity in the FREPs that are down-regulated by ds/ssRNA. Why such a broad range of FREPs are down-regulated by ds/ssRNA treatment is not clear.

7.3.3 Genes regulated by dsRNA only

There are 110 dsRNA specific responsive genes. Just less than half the genes have no known putative function, or have very diverse functions. Those with putative functions can be divided into several categories.

7.3.3.1 Signalling

Six genes with functions in cellular signalling are differentially regulated – five of these are up-regulated. Additionally an interesting LRR protein that appears to have signalling capability is the most highly regulated gene that is dsRNA responsive. This LRR is a transmembrane protein with two possible immuno tyrosine-based activation motifs (ITAMs) in the intracellular portion of the protein (a 'classical' ITAM is YxxI/$Lx_{(6-12)}$YxxI/L [208], the two possible ITAMs are Yxx$Lx_{(14)}$YxxI and Yxxx$Lx_{(8)}$YxxL). These two possible ITAM motifs are highly conserved between *A. gambiae*, *Ae. aegypti* and *D. melanogaster*. The extracellular portion contains 18 leucine rich repeats (LRRs) that are also well conserved across the three species. ITAMs are used by the classical immunoreceptors; T cell receptors (TCRs), B cell receptors (BCRs), FcRs, activating NK cell receptors, TREM1 and 2 all signal through ITAMs, activating the Scr family of kinases [208]. ITAMs are also utilised by a variety of non-immune receptors for signalling via the Scr family of kinases [208]. LRRs are structural motifs of 20-29 residues containing a conserved pattern of leucine residues. Typically each LRR forms a β strand and an α helix connected by loops, with each LRR structural unit being arranged on a parallel axis and so forming a horseshoe structure [209]. From surveys of LRR containing proteins, it is believed that the main function of LRRs is to provide a structural framework for protein-protein recognition[209]. Importantly, although the leucines are highly conserved, the residues between the patterns of leucines can be highly variable [210]. LRRs are found in many immune related genes; the TLRs have tandem LRRs, which deviate in sequence in the LRR domains, conferring the ability to bind a wide variety of PAMPs[210]; the LRIM family of proteins in *A. gambiae* all contain LRR domains, and have roles in anti-plasmodium immunity [126]; the variable lymphocyte receptors (VLRs) in the jawless vertebrates contain highly diverse LRRs generated from a VLR locus with banks of LRR cassettes that can be inserted into incomplete germline VLR genes, thus individual lymphocytes express unique arrangements of LRRs for recognition of a variety of pathogens [211]. Examples of PAMPs recognised by LRR containing proteins are bacterial DNA (TLR7/8) [210], dsRNA (TLR3) [212], LPS (TLR4) [213], flagellin (TLR5) [210] and fungal effectors [214].

In addition to this receptor, a dual specificity tyrosine phosphorylation regulated kinase is up-regulated, a tyrosine phophatase is down-regulated, and a possible transcription factor is up-regulated. The tyrosine phosphatase is an ortholog of *D. melanogaster* TRAM1; TLR4 induces signalling via two sets of adaptor proteins, one of which is the TRAM/TRIF pair of adaptor proteins[168]. It is temping to speculate that they may belong to a pathway that recognises exogenous dsRNA. However, due to the wide variety of PAMPs recognised by LRR proteins, and the huge diversity in LRR sequences, it is impossible to theorise what is the ligand of this receptor.

7.3.3.2 Immunity genes

The immune genes regulated by dsRNA treatment have a variety of predicted functions. Several genes associated with defence against Gram negative bacteria are up-regulated; PGRP-LB1 is the ortholog of a Gram negative anti-bacterial protein regulated by the IMD pathway in *D.melanogaster* [182]; CTLMA4 has a role in defence against Gram negative bacteria in *A.gambiae* [183]; ML5 (an MD2-like receptor) may bind to the same substrate as the MD2 receptor (LPS found in the wall of Gram negative bacteria). Several immune genes are down-regulated; DEF1 is an antimicrobial peptide that is regulated by both the Toll and the IMD pathway in *A. Gambiae* [135] and is predominantly active against Gram positive bacteria [49]; LRIM1, SCRC1 and three FREPs are all putative recognition proteins that respond to a wide variety of pathogens, but it is noteworthy that all three FREPs have been shown to be down-regulated in response to *E.coli* (a Gram negative bacteria) in *A. Gambiae* [204].

The differential regulation of the above immune genes by dsRNA indicates that pathways triggered by Gram negative bacteria are activated and/or modulated. However, the gene expression induced by the Toll and IMD pathways in *A. gambiae* are not as clear as they are in *D. melanogaster*, with several AMPs responding to signalling by both pathways [135]. The remaining immunity genes have roles in the

regulation of melanisation; SRNP6 and PPO2/4 are down-regulated (SRPN6 inhibits melanisation in *A. Gambiae* [215] and PPOs are the precursor to the enzyme at the start of the cascade that produces melanin), but SRPN5/16 are up-regulated. Due to poorly defined roles of most of these genes, it is not possible to speculate how melanisation is being altered by dsRNA treatment.

7.3.4 Genes regulated by ssRNA

A smaller number of genes are regulated by ssRNA alone, however (as will be discussed in limitations of the microarray study) the majority of fold change ratios for the corresponding genes in the dsRNA microarray are missing, and as such it is difficult to say that these genes are specifically regulated by ssRNA alone. Two groups of genes are ssRNA responsive; **Redox status** - A group of 5 genes with functions in maintaining the redox status of cells, including a glutathione-S transferase, are up-regulated by ssRNA. As discussed previously this may be part of a respiratory burst response to a danger signal triggered by non-self RNA. **Immune genes –** a smaller number of immunity genes are ssRNA responsive (9). The majority of differentially regulated immune genes have uncharacterised functions and belong to gene families with diverse pathogen recognition capability (four LRRs and two FREPs) or diverse functions in modulating immune signalling (two CLIPs). Their differential regulation suggests that ssRNA is also recognised as a PAMP by *A. gambiae* cells.

7.3.5 Overlap of RNA responsive genes with other gene lists

The genes differentially regulated by ss/dsRNA were compared with virally responsive genes. Only a small number of genes are differentially regulated by ss/dsRNA and viral infection; 6 genes are up-regulated by both ss/dsRNA and viral infection, and 11 genes are regulated in opposite directions by ds/ssRNA and viral infection. This very small overlap indicates that 5'ONNVic-eGFP infection triggers

different signalling pathways to ss/dsRNA, suggesting that the virus is not recognised through foreign ss/dsRNA. It is perhaps not surprising that the two data sets do not overlap considering that ss/dsRNA treatment was extracellular, and is taken up into cells, as seen in *Drosophila* cells [151, 216]. Infection of mosquito cells with 5′ONNVic-eGFP does not appear to trigger apoptosis, and so we would not expect extracellular foreign ss/dsRNA to be present. The transcriptional responses to exogenous ss/dsRNA may be triggered by receptors that recognise extracellular RNA, whereas viral infection may be recognised through a variety of PAMPs, such as the coat proteins of the virus, viral genomic RNA, dsRNA replication intermediates or secondary structure of RNA and viral transcripts. ss/dsRNA represents only a single PAMP compared to the many present during a natural viral infection. However the observation that recognition of foreign RNA leads to gene regulation that can target viral infection supports the idea that foreign RNA is recognised as a viral PAMP. Considering that viruses utilise host cells to produce and post translationally modify their proteins and use host membranes to form new virions, foreign RNA is possibly one of the most accessible and recognisable PAMPs.

Interestingly none of the genes differentially regulated by RNA are putative members of the RNAi pathway. It is possible that components of the pathway are already sufficiently expressed in the cell line, and so differential expression of these genes is not required for efficient RNAi. The dataset of RNA responsive genes was also compared to those identified in a dsRNA uptake pathway in *D.melanogaster* [151, 216]; again none of the RNA responsive genes are involved in the dsRNA uptake pathway in *D.melanogaster*, however whether this pathway is conserved between *D. melanogaster* and *A.gambiae* is not clear.

7.3.6 Limitations of the microarray study

The first limitation of this microarray study involves an issue already discussed in section 6.1.8. When attempting to compare differentially regulated genes from two separate microarray experiments, there is often data 'missing' from one of the

arrays for a number of reasons – data can be missing because it is flagged as 'absent' before analysis in Genespring software. This includes features that have diameters outside two standard deviations of the mean of all features, features that do not exceed local and/or global backgrounds (i.e. are low intensity) and features removed due to scratches, dust, dirt or patches that have dried out during hybridisation on the array. Data can also be missing because of high variability between biological replicates: T-test p values greater than 0.05 are excluded from analysis. In order to compare between gene lists, the 'missing' data that can be retrieved (in this case those with inconsistent ratios and thus high P values) have been added into tables of regulated genes for comparison (see Appendix 2, text in red), however data removed before analysis in Genespring software cannot be retrieved, due to the fact that any data retrieved would be in a raw format prior to normalisation. Considering that Lowess normalisation takes into account each feature from the array, un-normalised data cannot be directly compared with normalised data. During this microarray study, the hybridisation for ssRNA was considerably better than for dsRNA, and as such, much of the data found to be differentially regulated by ssRNA does not have an equivalent value in the dsRNA analysis. This means it is difficult to say accurately that a gene regulated by ssRNA is exclusively regulated by ssRNA alone. The same problem does not occur in the opposite direction, where data for both genes regulated by dsRNA and ssRNA are present. As a technical problem with microarray analysis this is difficult to overcome.

The second limitation in this study occurs when we try to compare the experiment to a natural situation that occurs during viral infection. During the experiments short (~500bp) fragments of ss/dsRNA were used to treat cells. In a natural system, viral RNA and viral RNA transcripts are much longer than ~500bp, and may be recognised and processed using different mechanisms. Also, cells were treated exogenously with ss/dsRNA; the majority of virally derived RNA is expected to be intracellular considering that their replication cycle takes place within the cytoplasm of host cells, and in the case of ONNV infection of *A. gambiae*, does not cause apoptosis, thus limiting the exposure of cells to exogenous foreign RNA. This does not detract from the finding that recognition of foreign RNA and subsequent signalling results in the

dramatic impairment of ONNV infection and/or replication, however it is likely that natural viral infection is recognised through different mechanisms, as highlighted by the very small overlap between virally responsive and RNA responsive genes.

7.3.7 Final comments

This study has revealed that ss and dsRNA trigger immune mechanisms in *A. gambiae* cells that can inhibit the infection and/or replication of ONNV. Both ds and ssRNA were shown to alter the transcriptional profile of a large number of genes, including immune related genes. Both ds and ssRNA appear to regulate genes mainly associated with regulation of the melanisation cascade, and a number of genes involved in detoxification of ROS indicating that ds and ssRNA induce an oxidative burst in *A. gambiae* cells. In addition to responding to both ds and ssRNA, groups of genes are also dsRNA specific and ssRNA specific. dsRNA induces the regulation of a number of interesting signalling genes that may be involved in dsRNA recognition and subsequent signalling. ssRNA triggers the differential regulation of a smaller number of genes, the immune category of which primarily have uncharacterised or diverse functions. It is not clear which of the immune genes differentially regulated are responsible for the dramatic decrease in virus in ds/ssRNA treated cells, however the oxidative burst that appears to be induced has been shown in other systems to be active against viral infection. The oxidative burst may be another immune mechanism employed by *A. gambiae* to tackle viral infection. These observations lead further work in the direction of identifying the molecular basis of ds/ssRNA recognition, and identifying the responses to foreign RNA that are anti-viral.

8 Final discussion and future perspectives

8.1 Final discussion

The invertebrate immune system was long considered to be archaic and simple compared to its mammalian counterpart, however over the past 20 years, the complexity of invertebrate innate immunity has been revealed. Several distinct immune signalling pathways respond to a variety of pathogens, resulting in the differential regulation of potent anti-microbial peptides (considered to be the hallmark of invertebrate immunity), the activation of complement-like signalling cascades that recognise and kill pathogens, the activation of white blood cell-like hemocytes that can phagocytose or melanotically incapsulate pathogens, and activation of the melanisation cascade. Although it lacks the adaptive capability of the mammalian immune response, the invertebrate immune system is far from simple. Much research has been carried out into the responses of *A. gambiae* to malaria parasites and bacterial infections, however it is only in recent years that anti-viral immunity has been in the spotlight. Mosquito-borne viruses are responsible for huge numbers of human infections every year, including DENV, YFV, WNV and CHICKV to name but a few. Just as mosquitoes have been shown to respond to bacterial, fungal and parasitic infections, they also respond to viral infection, with recent evidence implicating three signalling pathways as key in invertebrate anti-viral immunity – the JAK/STAT pathway, the Toll pathway, and most importantly the RNAi pathway. The majority of studies carried out to date on anti-viral immunity in mosquitoes have focused on the *Aedes* genus (one of the most important group of vectors of human viruses). Mosquito-borne arboviruses are almost exclusively transmitted by *Aedes* or *Culex* mosquitoes, with one exception – ONNV, which is transmitted by Anopheline mosquitoes. The ONNV-*A. gambiae* combination offers a unique opportunity to study the interactions occurring in a natural infection of *A. gambiae* mosquitoes with an important human pathogen, possibly giving insights into the poor viral vectorial capacity of Anopheline mosquitoes. One of the advantages of studying *A. gambiae* mosquitoes is the variety of tools available in the laboratory, including DNA microarrays with full genome coverage of *A. gambiae*, and a selection of cell lines derived from several strains of *A. gambiae*.

The main objective of this study was to identify genes involved in the anti-viral immune response of *A. gambiae*, using 2 approaches; the first based on the literature currently available to select candidates that may have a role in anti-viral immunity; and a second using forward genetics to identify viral responsive genes through transcriptional profiling. A secondary objective of the project was to develop a reverse genetic approach to rapidly identify genes for functions in ONNV infection using RNAi, a GFP-encoding ONNV infectious clone and *A. gambiae* cell lines, with a view to conducting a genome wide RNAi screen.

The study was successful in indentifying immune responses to ONNV infection, and in demonstrating that some of these responses are involved in limiting viral infection in *A. gambiae* mosquitoes. Four genes were identified with novel functions in anti-viral immunity in invertebrates. The study also clearly demonstrated that development of a cell-based genome wide RNAi screen approach to identify viral antagonists may not be feasible with the current tools available in *A. gambiae*. It was observed that even non-specific RNA triggers a response in cells that dramatically reduced viral infection. This result although discouraging with regards to the development of a high-throughput reverse genetic approach, triggered another very interesting question: whether non-specific foreign RNA acts as a PAMP, as seen in mammalian systems. Transcriptional profiling of responses to ds and ssRNA were conducted revealing that non-specific foreign RNA induces immune gene expression and triggers oxidative stress in *A. gambiae* cells and opening new avenues for the study of mosquito anti-viral defence. An additional inherent problem with the development of a cell-based assay for viral infections was the use of GFP expression as a readout, which did not allow accurate quantification of the viral infection or monitoring the different phases of the infections. It was concluded that different quantitative readouts tightly linked to the presence of the virus must be developed.

The main findings of this study can be divided into four sections discussed below, each of which adds to our understanding of the interactions between *A. gambiae* and ONNV, and opens up new areas of research to build upon these findings.

8.1.1 Susceptibility of *Anopheles spp* to ONNV

The results obtained in this study support the idea that *A. gambiae* may not be the main vector of ONNV, with low rates of infection, slow replication rates and poor dissemination compared to many other vector-virus combinations. Testing ONNV infection in several Anopheline mosquitoes revealed large differences in susceptibility to viral infection. *A. quadriannulatus* is more susceptible to ONNV infection than *A. gambiae* or *A. arabiensis*. Interestingly the reverse is true for infection with *Plasmodium* parasites. *A. quadriannulatus* is refractory to *Plasmodium* infection, whilst *A. gambiae* and *A. arabiensis* are susceptible [63]. It has been shown that KD of immune genes in *A. quadriannulatus* reverses the refractory phenotype of the mosquito rendering it a capable vector of malaria [63]. This demonstrates that differences in immune status determine the vectorial capacity of a mosquito *spp* for *Plasmodium* infection. *A. gambiae* can be infected with ONNV, but frequently limits viral infection, and restricts tissue tropism of infection. Silencing immune genes, such as AGO2, in *A. gambiae* leads to widespread infection of mosquitoes demonstrating that most tissues in *A. gambiae* can be infected with ONNV [80] i.e. restricted tissue tropism is not caused by the lack of a receptor necessary for infection, but that most tissues are not infected, possibly due to immune responses to infection. The differences in vectorial capacity of *A. gambiae* and *A. quadriannulatus* for malaria transmission are hypothesised to be caused by evolutionary co-adaptation of mosquito species frequently exposed to malaria parasites, thus *A. gambiae*, which is frequently exposed to malaria parasites, has an attenuated immune response to parasite, whereas *A. quadriannulatus*, which has limited exposure to malaria parasites, retains the ability to kill invading parasites [63].

Just as *A. quadriannulatus* receives limited exposure to *Plasmodium*, epidemics of ONNV have occurred in areas where *A. gambiae* and *A. arabiensis* are abundant, but *A. quadriannulatus* is scarce. It is possible that the continual exposure of *A. gambiae* and *A. arabiensis* to ONNV has resulted in the evolution of anti-viral mechanisms in *A. gambiae*, however this has not occurred in *A. quadriannulatus* due to

geographical isolation. Indeed if ONNV infection of *A. gambiae/arabiensis* caused mortality as observed in *A. quadriannulatus*, there would be significant pressure to evolve mechanisms to cope with viral infection.

After invasion of the midgut wall, *Plasmodium* parasites are exposed to the hemolymph of the mosquito, which contains a multitude of immune proteins that can recognise pathogens and regulate an appropriate immune response. Co-infections of ONNV and *P. berghei* revealed that viral infection shifts the humoral immune response away from melanisation and towards lysis or opsonization of pathogens. Parasites entering the hemocoel of mosquitoes infected with ONNV showed reduced melanisation and increased parasite losses through lysis. Conversely *A.quadrinnaulatus* melanises large numbers of invading *P. berghei* parasites compared to *A. gambiae*, suggesting that its immune system is skewed towards melanisation of invading pathogens. This shift in immune responses may account for the increased susceptibility to viral infection, where genes that can target viral infection through lysis or opsonisation are inhibited and genes that promote melanisation are activated. This implies that the same immune responses that lyse parasites may be targeting viral infection, such as the mosquito complement system. Indeed ML1, a gene found to be antagonistic to viral infection, is also a *P. falciarum* antagonist, KD of which increases live oocyst number, indicating a role in the lysis of parasites [134].

8.1.2 *A. gambiae* can recognise and respond to viral infection through a combination of conserved anti-viral mechanisms and other *spp* specific immune responses

Transcription profiling of *A. gambiae* responses to ONNV infection demonstrated that *A. gambiae* differentially regulates many immunity related genes, presumably to cope with viral infection. Indeed several of these genes were shown to antagonise viral infection. Interestingly some key differences between *A. gambiae*-ONNV interactions compared to other vector-virus combinations were observed.

Gene silencing experiments confirmed that RNAi has a central role in limiting viral infection, as previously demonstrated in *A.gambiae* and a variety of other invertebrate systems including *Aedes* mosquitoes and *D.melanogaster*. However it was demonstrated that other immune signalling pathways known to be involved in anti-viral immunity in insects do not appear to function in modulating ONNV infection in *A.gambiae*. The Toll pathway, an established anti-viral pathway in *D. melanogaster* and known to regulate DENV2 infection in *Ae.aegypti* [72], does not modulate ONNV infection in *A.gambiae*. Activating or inhibiting the pathway had no effect on viral titres in adult mosquitoes. Also, genes known to be regulated by the Toll pathway in *A.gambiae,* such as CEC1 (personal communication, G K Christophides), were not identified in transcriptional profiling of infected mosquitoes. This suggests that ONNV infection does not trigger Toll signalling, and the Toll pathway does not up-regulate genes involved in regulating ONNV infection.

The JAK/STAT pathway is also an established anti-viral pathway in *D.melanogaster* and *Ae.aegypti*. Although inhibition of the JAK/STAT pathway by silencing of HOP led to a moderate increase in ONNV infection in L35 cells, *A.gambiae* mosquitoes show a small and inconsistent increase in viral titre when the JAK/STAT pathway is inhibited (through silencing of HOP, STAT1 and STAT2). Activation of the pathway through KD of PIAS, an inhibitor of the JAK/STAT pathway, does not decrease viral infection (as demonstrated in *Ae.aegypti* infected with DENV2 [75]). Additionally infection of L35 cells with ONNV does not induce translocation of STAT1 into the nucleus, implying that there is no activation of the JAK/STAT pathway by ONNV infection. These results suggest that the JAK/STAT pathway may have a limited role in regulating ONNV infection but is not triggered by viral infection. *A.gambiae* has two described STAT genes [100]. It is possible that STAT2, rather than STAT1, is translocated to the nucleus following pathway activation by viral infection. Also PIAS has not been shown to inhibit the JAK/STAT pathway in *A.gambiae*: its function has been inferred by the role of orthologous genes in *D.melanogaster* and *Ae.aegypti*. As such, silencing PIAS may not induce JAK/STAT signalling. To date, no markers of JAK/STAT activation in *A.gambiae* have been identified, and so determining if the pathway has

indeed been activated is challenging. Although the data presented here suggest that the JAK/STAT pathway does not respond to viral infection, to confirm this, direct evidence that PIAS silencing activates the JAK/STAT pathway, and whether STAT2 translocates to the nucleus upon ONNV infection would be required. The IMD pathway also appears to have no effect on ONNV when inhibited through REL2 KD, despite the indications from transcriptional profiling that ONNV infection triggers IMD signalling.

In addition to the major immune signalling pathways, mosquitoes have a humoral branch of the immune system, consisting of complement-like proteins, serine proteases and enzymes that can recognise invading pathogens and target them for melanisation, lysis or opsonisation. The proteins that regulate these processes circulate in the hemolymph, and through the recognition of pathogens, lead to the activation of signalling cascades that can tip the balance of an immune response towards melanisation, lysis or opsonisation. Recently a complex of complement-like proteins that recognise invading *Plasmodium* parasites and target them for lysis was described in *A.gambiae* [126]. Transcriptional profiling of ONNV infected mosquitoes showed that several complement-like proteins are up-regulated following infection, including a member of this complex, LRIM1. Although viral particles are difficult to visualise during an infection, *Plasmodium* parasites are easily visible and can be used to visualise the responses of the mosquito immune system: melanisation is clearly visible as invading ookinetes are coated in black melanin and lysis of parasites can be inferred from changes in the number of live oocysts observed during infection compared to controls. Co-infections, designed such that ookinetes invading the midgut come into contact with hemolymph proteins at 4 dpi with ONNV, reveal that melanisation is inhibited by ONNV infection, diverting immune responses towards lysis and/or opsonisation (as demonstrated by a decrease in the number of melanised parasites and an overall decrease in the number of live parasites in ONNV infected mosquitoes). This is consistent with the up-regulation of two genes that have roles in melanisation inhibition in *A. gambiae*: CTLMA2 and CLIPA2 [62, 183]. Interestingly the L35 cell line (which is highly susceptible to ONNV infection) is derived from a strain of *A. gambiae* that melanised

all invading *Plasmodium* parasites, and *A. quadriannulatus* (also highly susceptible to ONNV infection) melanises large numbers of invading *Plasmodium* parasites. This suggests that diversion of immune responses towards melanisation and away from lysis or opsonisation increases susceptability to ONNV infection.

The overall number of live parasites was decreased in ONNV infected mosquitoes, indicating that some parasites are lysed after invasion of the midgut. Four LRIMs and five TEPs are up-regulated by ONNV infection, and may form the basis of recognition and targeted lysis of viral particles in a similar mechanism to the LRIM1/APL1C/TEP1 complex. Interestingly, although LRIM1 is up-regulated, APL1C and TEP1 are not, suggesting that LRIM1 may interact with other binding partners in response to different PAMPS. Although the number of live parasites decreased in ONNV infected mosquitoes, the reduction in oocyst number was not dramatic, suggesting that other mechanisms including opsonisation may be triggered by viral infection. Two scavenger receptors with putative roles in opsonisation are up-regulated, along with three Galectins that may act as opsonins.

The finding that two pathways with known anti-viral roles in other invertebrate-virus systems do not extensively modulate ONNV infection indicates that *A. gambiae* may use other immune mechanisms to recognise and fight viral infections. Transcriptional profiling of infected mosquitoes and co-infections with *Plasmodium* parasites suggests these mechanisms involve the complement-like proteins of the humoral immune response functioning in lysis and opsonisation of pathogens. The anti-viral immune response in *A. gambiae* is thus composed of some key conserved mechanisms to target viral infection including RNAi, but also other diverse and possibly species-specific mechanisms. Four genes found to be anti-viral in *A. gambiae* have novel roles in invertebrate anti-viral immunity, pointing to *A. gambiae*-specific recognition of viral PAMPs and mechanisms for targeting viral infection.

8.1.3 Four genes were discovered to have novel anti-viral functions in *A. gambiae*

The four novel viral antagonists identified in this study are all predicted to be secreted proteins and are not known to function in the described immune signalling pathways in *A. gambiae*. ML1, an MD2-like receptor was found to be an antagonist of ONNV infection in both mosquitoes and L35 cells. Interestingly when looking at the pattern of GFP expression in *per os* infected ML1 KD mosquitoes, no difference in infection prevalence is observed but a skew towards higher numbers of cells being infected is seen. This indicates that ML1 does not function in preventing initial infection, but does limit the spread of virus from cell to cell. ML1 may act as a recognition protein that binds to a receptor on cells inducing signalling that regulates anti-viral effectors, as seen in LPS recognition in mammalian systems where LPS binds to heterodimers of MD2 and TLR4 [168]. ML1 may act as part of an extracellular surveillance system to detect viral infection. In contrast AGO2 KD has no effect on the number of cells that become infected, but does increase the prevalence of infection, indicating that RNAi is important for preventing the establishment of infection. This is consistent with the intracellular role of the RNAi pathway, where RNAi can prevent the virus from replicating within an initially infected cell, and thus occasionally clearing infection before it is established. When RNAi is inhibited initial infection is more often successful, thus increasing the prevalence of infection. RNAi is an established intracellular surveillance mechanism for detecting viral infection, confirmed in this study as being key in regulating viral infection in *A. gambiae*.

GALE8 was shown to be a viral antagonist in *A. gambiae* mosquitoes, and is up-regulated by both ONNV infection, and RNA treatment of cells. GALE8 is a prototype galectin; prototype galectins contain a single CRD and often form homodimers. It is one of three GALEs that transcriptionally respond to ONNV infection, all of which are part of a mosquito specific expansion of the GALE family. One possible mechanism of action is suggested by evidence from galectin-virus interactions observed in mammalian systems; GALE8 may bind and cross-link the envelope glycoproteins of virus particles, blocking the ability of the virus to infect cells. The differential

regulation of GALE8 during mid-infection suggests that GALE8 functions downstream of initial recognition of viral infection. Galectins also have reported roles in immune signalling, induction of cytokines, and driving proliferation of immune cell populations [164]. As such, GALE8 may have diverse roles in anti-viral immunity, including activation of hemocytes or inducing proliferation of hemocytes. Two lysozymes were also demonstrated to have anti-viral roles in this study; LYSC4/6. As classical anti-bacterial proteins, the mechanism of anti-viral immunity is unclear.

The four genes discovered to have novel anti-viral functions in *A. gambiae* suggest new mechanisms of viral recognition and subsequent inhibition of viral infection through the recognition protein families ML and GALEs and the effector protein family of the lysozymes, with multiple members of each family responding transcriptionally to ONNV infection. Although speculative modes of action are outlined in figure 41, much work must be carried out to begin identifying the actual mechanism of action of each of these interesting proteins.

8.1.4 Foreign RNA acts as a PAMP in *A. gambiae* cells

This study has demonstrated that RNA acts as a PAMP in *A. gambiae* cells and triggers transcriptional responses that can limit viral infection. In invertebrate immunity there are numerous PAMPs recognised as danger signals. Some of these PAMPs are very specific, and will trigger equally specific immune responses. For example, ML1 is regulated in response to *P. falciparum* infection and is an antagonist of parasite development, however it has no function in limiting *P. berghei* parasite development. Immune responses can be species specific (as just described), structure specific, such as the Toll pathway responding to Gram positive bacteria, and the IMD pathway responding to Gram negative bacteria, and pathogen specific, such as the Toll pathway responding to fungal infection. The PAMPs that trigger these responses can equally be species-specific or broad spectrum. Experiments in this study sought to answer if foreign RNA can act as a broad spectrum PAMP in

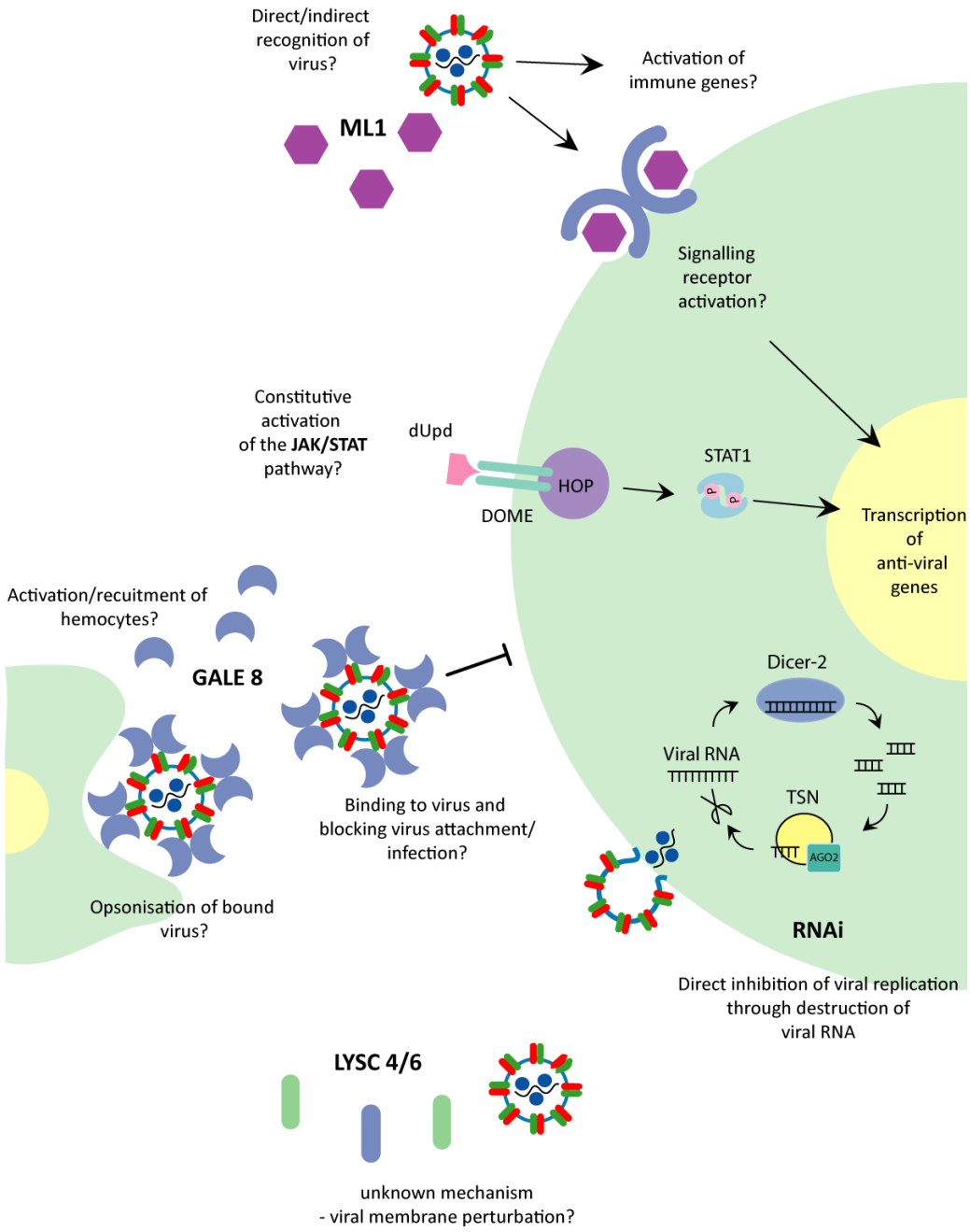


Figure 41. Possible mechanisms of anti-viral immunity in *A. gambiae* mosquitoes. Putative functions of immune genes that regulate ONNV infection in *A. gambiae*. Modes of action have been inferred from the function of orthologous genes in other organisms. ML1 may directly or indirectly recognise extracellular virus, and bind to a TLR-like receptor inducing dimerisation and signalling. The JAK/STAT pathway, although not activated by viral infection, regulates the transcription of unknown anti-viral genes. GALE8 may bind glycoproteins in the coat of circulating virus directly preventing infection of cells or acting as an opsonin. Alternatively GALE8 may induce proliferation of, or activate, hemocytes circulating in the hemolymph. The RNAi pathway limits viral infection through the destruction of viral specific transcripts within infected cells. LYSC4/6 may perturb the membrane of circulating viruses.

A. gambiae, as seen in mammalian immune systems. Intra- and extra-cellular ss and dsRNA are recognised by a number of mechanisms in mammalian cells, and trigger a variety of immune responses, in particular those targeting viral infection. Foreign RNA indeed does illicit immune responses in *A. gambiae* derived cells, and the transcriptional responses induced by foreign RNA dramatically limit viral infection. This is, to date, the first demonstration of non-specific RNA as a PAMP in mosquitoes, and opens an interesting new avenue to study the mechanisms of RNA recognition, RNA uptake and subsequent triggering of signalling, and to try and identify the mechanism of viral suppression downstream of RNA recognition.

8.1.5 Differences in responses between adult *A. gambiae* and *A. gambiae* derived cells to ONNV infection

Some interesting differences were observed between adult mosquitoes and the L35 *A. gambiae* cell line. Identified antagonists through gene silencing experiments in L35 cells were ML1, HOP and non-specific ds/ssRNA. AGO2 and GALE8 KD showed no effect on viral titre. In contrast in adult mosquitoes, ML1, AGO2 and GALE8 were identified as antagonists, but HOP KD and non-specific dsRNA had no effect on viral titre. Additionally the L35 cell line is highly susceptible to ONNV infection, in contrast to adult mosquitoes where infection patterns are restricted. The observation that AGO2 silencing has no effect on viral titre in L35 cells, coupled with the high susceptibility to infection suggests that L35 cells do not have a functioning RNAi pathway. However, silencing of genes using dsRNA has been achieved in L35 cells, demonstrating that they are capable of RNAi (data not shown). In fact, the efficiency of gene silencing of a single gene was tested in all 7 *A. gambiae* cell lines, and no correlation between the efficiency of gene silencing and susceptibility to 5'ONNVic-eGFP infection was observed (data not shown). Although it has been demonstrated that AGO2 is required for RNAi in the Sua1B cell line [217], it is possible that AGO2 is redundant in L35 cells, and is not required for RNAi.

The *A. gambiae* cell lines were derived from minced neonates; newly hatched larvae were used for primary culture from three *A.gambiae* strains (Suakoko 2La, 4a r/r and L-35). These primary cell lines have not been immortalised or cloned and contain a mixed population of cells that is assumed to be continually changing. Each cell line has varying properties, for example the 4a3A cell line does not express PPOs (enzymes involved in the melanisation response) in comparison to the 4a3B cell line which expresses 6 PPOs [99] despite both cell lines being derived from the same initial pool of newly hatched larvae. The cell lines currently used in the laboratory were selected for immuno-responsiveness (for example to bacterial challenge) and are described as hemocyte-like from immune gene expression patterns, making them appropriate for use in studies of immune responses to pathogens [99, 218]. Little characterisation of the L-35 derived cell line used in this study has been carried out. Microarray analysis of cell line responses to pathogen challenge compared to adult mosquito responses to pathogen challenge showed similar patterns of gene expression (Dr G christophides, personal communication) and as such the cells have been considered a useful model for experiments. However the due to the heterogeneous population of cells within these cultures, properties including gene expression may change over time decreasing the accuracy of modelling mosquito-arbovirus interactions.

Although *A.gambiae* derived cells will not represent the complex responses of multiple tissues to viral infection, and we can expect that through changes in cell populations during passage different cells lines will respond to viral infection in different ways, the use of cell lines for high throughput analysis is still invaluable. The disparity between cell culture and whole mosquitoes is a problem common to all *in vitro* systems and underlies the importance of carrying out research *in vivo* where possible. However, some genes have similar functions *in vitro* and *in vivo*, for example ML1, and cell lines offer a practical solution for the rapid identification of genes with interesting functions. Continuing to develop a high throughput method for screening genes for functions in ONNV infection warrants further investment and will greatly facilitate the identification of anti-viral mechanisms in *A. gambiae* mosquitoes.

This series of experiments is the start of the process of understanding the interactions between *A. gambiae* and a natural viral pathogen of this mosquito. It is clear that Anopheline mosquitoes are poor vectors of viral disease, although they are efficient vectors of malaria parasites. The genetic differences that determine the vectorial capacity of mosquitoes for different groups of pathogens are just beginning to be dissected, including immune responses to infection. This study has revealed that not all of the immune signalling pathways with known anti-viral roles in invertebrates are utilised by *A. gambiae* mosquitoes to fight ONNV infection. The novel anti-viral genes identified in this study indicate that other uncharacterised immune responses can limit viral infection and may be important in defining the ability of mosquito *spp* to harbour and spread arboviral disease.

8.2 Future perspectives

The results of this study point to several avenues for further research in anti-viral immunity in *A. gambiae*. Much of the work conducted in this series of experiments represents the first steps, in a relatively young field of study. There is much work that can be carried out, building upon the results found here, which will broaden our understanding of the interactions between mosquitoes and their viral pathogens.

During this study it was noted that there is a large level in variation between individual mosquitoes to support ONNV infection. Some individuals were highly susceptible to viral infection, showing high viral titres when infected intrathoracically or showing broad patterns of midgut infection after *per os* infection. In contrast some individuals appear to be refractory to infection, showing very low titres after intrathoracic inoculation, or no infection in the gut after *per os* infection. The question arises if there are genetic differences between the individuals within these populations that confer susceptibility or refractoriness to infection? One way to investigate this is to identify any single nucleotide polymorphisms (SNPs) that exist between groups of highly susceptible verses refractory individuals using a SNP chip. Genes that confer susceptibility or refractoriness will have a higher density of SNPs

compared to other genes, and so these genes can be identified through finding areas of dense SNPs typically within a single gene. SNP chip analysis has been used to identify genes involved in insecticide resistance and in disease resistance, for example, SNPs have been identified in immune genes that are associated with *P. falciparum* infection [219]. During this study, large numbers of N'gousso strain *A. gambiae* (a recently colonised strain of *A. gambiae* that still retains a relatively large amount of genetic diversity within the colony) were *per os* infected with ONNV, and individuals demonstrating refractoriness to infection, or demonstrating high susceptibility to infection were collected and genomic DNA was extracted. In a collaboration with the Broad Institute, the DNA samples are currently being used for SNP analysis using an array containing 400,000 SNPs in *A. gambiae*. The results of this project are pending, but will hopefully give some insight into the genes that may confer the ability of *A. gambiae* mosquitoes to act as vectors of ONNV.

In addition to the high levels in variation observed between individuals of the same species, large differences in the ability to host ONNV infection were observed between different species of Anopheline mosquitoes. It would be interesting to investigate the transcriptional responses of *A. quadriannulatus* to ONNV infection and to contrast this with the transcriptional responses of *A. gambiae*. As previously discussed, in a study conducted by a colleague in the laboratory differences in gene copy number between *A. gambiae* and *A. quadriannulatus* were determined (Dr T Habtewold, unpublished data), and revealed large differences in gene copy number for apoptosis related genes. Apoptosis may be responsible for the mortality observed in ONNV infected *A. quadriannulatus* mosquitoes, which could be confimed using a number of fluorescent markers of apotosis in *A. quadriannulatus*. Additionally positive and negative regulators of apotosis could be silenced in *A. quadriannulatus* to see if this has any effect on viral infection or mortality.

During this study a small scale RNAi screen of 19 genes was carried out in *A. gambiae* mosquitoes to identify genes with roles in anti-viral immunity. The transcriptional profiling of infected mosquitoes revealed a large number of genes to be responsive to viral infection, many of which have no putative function. It would be important to

continue screening these genes for functions in anti-viral immunity, as they may identify *A. gambiae* specific mechanisms for targeting viral infection. The development of a high throughput cell based assay to screen genes would greatly assist this process, but has to date been unsuccessful. Methods to develop these assays should continue to be investigated.

This study identified four genes as being novel anti-viral genes in *A. gambiae*, however the mechanism by which each of these genes targets viral infection is not clear. Identifying these mechanisms is the next challenge in this field. Firstly it would be important to show if there is any direct interaction between the genes identified and invading virus. Localisation studies using antibodies against these genes may give insight into their function. For example can GALE8 bind to the glycosylated coat proteins of ONNV and/or does ML1 act as a recognition receptor that can bind virus or perhaps binds to receptors on immune/other cells. It would be interesting to see if these proteins bind not only to virus, but also what other proteins may be involved in viral recognition complexes, or signalling receptors that may be activated by these proteins. Producing tagged proteins and using techniques including immunoprecipitation assays could identify such proteins. Whether ML1 has a role as a signalling receptor could be investigated: determining the transcriptional profile of ML1 KD mosquitoes may identify genes with anti-viral functions. A possible function of GALE8 may be to activate/and or recruit hemocytes - proliferation of hemocyte populations after viral infection could be identified using markers of hemocyte activation (however these remain poorly characterised in *A. gambiae*) and markers of cellular proliferation. Each of the four genes belongs to a protein family with several predicted members in *A. gambiae*. There are 15 predicted MLs (two of which are up-regulated by ONNV infection), 11 predicted GALEs (3 of which are up-regulated by ONNV infection) and 8 predicted LYSCs (three of which are up-regulated by ONNV infection) in *A. gambiae*. It would be interesting to screen the other members of these protein families to see if they also have roles in anti-viral immunity.

Foreign RNA ellicits the transcriptional regulation of hundreds of genes in *A. gambiae* derived cells. Work now must begin to elucidate which of these genes is responsible for the inhibition of viral infection and/or replication of cells exposed to foreign RNA. Initially an RNAi screen could be carried out, to see if silencing any of the differentially regulated genes reverses the phenotype observed. Additionally, the mechanisms for recognition and uptake of exogenous foreign RNA remain to be characterised in *A. gambiae*. Silencing genes with putative roles in the mechanisms identified through the transcriptional analysis carried out in this study, for example the signalling genes that are differentially regulated by dsRNA may reveal functions in recognition and signalling in response to dsRNA recognition.

8.3 Concluding comments

In summary this study has attempted to identify components of the anti-viral immune system of *A. gambiae* mosquitoes. Through transcriptional profiling of virally infected mosquitoes, it was demonstrated that *A. gambiae* responds to viral infection through the differential regulation of a number of putative immunity genes, in addition to many genes with non-immune or uncharacterised function. Further to this, an RNAi based screen identified several virally responsive genes as being antagonistic to viral infection. One of these genes is a component of the RNAi pathway and has previously been described as a viral antagonist in *A. gambiae*. The other 4 genes have novel roles in anti-viral immunity, and do not appear to belong to the classical invertebrate anti-viral immune signalling pathways i.e. the Toll and the JAK/STAT pathways. Additionally the Toll pathway has no effect on ONNV infection and the JAK/STAT does not appear to activated by ONNV infection. Thus it appears that *A. gambiae* mosquitoes utilise a combination of key conserved anti-viral mechanisms to target viral infection, such as RNAi, with other mosquito specific, and possibly even species specific anti-viral immune mechanisms. Transcriptional profiling of infected mosquitoes and co-infections with *P. berghei* suggests these mechanisms involve the complement-like proteins and lysis and/or opsonisation. This study also revealed non-specific RNA to be a broad spectrum PAMP in *A.*

gambiae derived cells, triggering transcriptional responses that antagonise viral infection.

The study of the interactions between the mosquito immune system, and viral pathogens is not only interesting in terms of the fascinating cell biology but is important for the development of novel approaches to prevention of arboviral disease transmission. Viruses have very different lifestyles compared to many other pathogens, being obligate intracellular pathogens that rely on host cells to complete their lifecycles. They represent a unique challenge to the immune system, and require sophisticated mechanisms of recognition and targeting. Completing their replication cycle within host cells enables viruses to 'hide' from conventional recognition from the immune system, and their use of the host cell membrane to form new viral particles reduces the number of PAMPs that the immune system can recognise. The mechanisms immune systems have evolved to detect these stealthy pathogens are not only ingenuous, but provide us with novel targets for the development of disease prevention methods, for example rendering mosquitoes refractory to viral infection through genetic engineering, using transmission blocking vaccines or chemical modification of immune responses. Although many years away from realisation, each of these theoretical disease prevention methods requires a detailed understanding of the interactions between mosquitoes and their viral pathogens. This study is one of the first steps towards elucidating the unique ways in which *A. gambiae* mosquitoes respond to and fight viral infection.

9 References

1. Gillett, J.D., *Mosquitoes*. The World Naturalist, ed. R. Carrington. 1971, London: The Chaucer Press.
2. Mellor, P.S., *Replication of Arboviruses in Insect Vectors.* J.Comp.Path, 2000. **123**: p. 231-247.
3. Beaty, B.J., *Control of arboviruses diseases: is the vector the weak link? .* Arch Virol Suppl, 2005. **19**: p. 73-88.
4. Tolle, M.A., *Mosquito-borne diseases.* Curr Probl Pediatr Adolesc Health Care, 2009. **39**(4): p. 97-140.
5. Mukhopadhyay, S., R.J. Kuhn, and M.G. Rossmann, *A Structural Perspective of the Flavivirus Lifecycle.* Nature Reviews Microbiology, 2005. **3**: p. 13-22.
6. Gubler, D.J., *Dengue/dengue haemorrhagic fever: history and current status.* Novartis Found Symp, 2006. **277**: p. 3-16; discussion 16-22, 71-3, 251-3.
7. Initiative, P.D.V. *Pediatric Dengue Vaccine Initiative*. 2009.
8. Takken, W. and G.J. Knols, *Emerging pests and vector-borne diseases in Europe*. Ecology and control of vector-borne diseases. Vol. 1. 2007: Wageningen Academic Publishers. 499.
9. Enserink, M., *Entomology. A mosquito goes global.* Science, 2008. **320**(5878): p. 864-6.
10. Mackenzie, J.S., D.J. Gubler, and L.R. Petersen, *Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses.* Nat Med, 2004. **10**(12 Suppl): p. S98-109.
11. Strauss, J.H. and E.G. Strauss, *The alphaviruses: gene expression, replication, and evolution.* Microbiol review, 1994. **58**(3): p. 491-562.
12. Ross, R.W., *The Newala epidemic. III. The virus: isolation, pathogenic properties and relationship to the epidemic.* J Hyg (Lond), 1956. **54**(2): p. 177-91.
13. Chhabra, M., et al., *Chikungunya fever: a re-emerging viral infection.* Indian J Med Microbiol, 2008. **26**(1): p. 5-12.
14. Chevillon, C., et al., *The Chikungunya threat: an ecological and evolutionary perspective.* Trends Microbiol, 2008. **16**(2): p. 80-8.
15. Pialoux, G., et al., *Chikungunya, an epidemic arbovirosis.* Lancet Infect Dis, 2007. **7**(5): p. 319-27.
16. Rezza, G., et al., *Infection with chikungunya virus in Italy: an outbreak in a temperate region.* Lancet, 2007. **370**(9602): p. 1840-6.
17. Watson, R., *Europe witnesses first local transmission of chikungunya fever in Italy.* BMJ, 2007. **335**(7619): p. 532-3.
18. Knudsen, A.B., R. Romi, and G. Majori, *Occurrence and spread in Italy of Aedes albopictus, with implications for its introduction into other parts of Europe.* J Am Mosq Control Assoc, 1996. **12**(2 Pt 1): p. 177-83.
19. Foy, B.D. and K.E. Olson, *Alphavirus Transducing Systems. Trangenesis and the Management of Vector-Borne Disease* S. Aksoy, Editor. 2006, Landes Bioscience: Austin. TX.

20. Shore, H., *O'nyong-nyong Fever: an epidemic virus disease in East Africa: ome clinical and epidemiological observations in the northern province of Uganda.* Transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1961. **55**: p. 361-373.
21. Williams, M.C., *O'nyong-nyong Fever: an epidemic virus disease in East Africa. VII. Virus isolations from man and seriological studies up to July 1961.* transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1965. **59**: p. 186-97.
22. Williams, M.C. and J.P. Woodall, *O'nyong-nyong Fever: an epidemic virus disease in East Africa.II. Isolation and some properties of the virus.* Transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1961. **55**: p. 135-41.
23. Johnson, B.K., *O'nyong-nyong virus disease*, in *The arboviruses: Epidemiology and Ecology*, T.P. Monath, Editor. 1988, CRC Press: Boca Raton, FL. p. 217-222.
24. Williams, M.C., J.P. Woodall, and J.S. Porterfield, *O'nyong-nyong Fever: an epidemic virus disease in East Africa. V. Human antibody studies by plaque inhibition and other seriological tests.* transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1962. **56**: p. 166-72.
25. Corbet, P.S., M.C. Williams, and J.D. Gillett, *O'nyong-nyong Fever: an epidemic virus disease in East Africa. IV. Vector studies at epidemic sites.* transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1961. **55**: p. 463-479.
26. Lutwama, J.J., et al., *Epidemic O'Nyong-Nyong fever in southcentral Uganda, 1996-1997: entomologic studies in Bbaale village, Rakai District.* Am J Trop Med Hyg, 1999. **61**(1): p. 158-62.
27. Levinson, R.S., E.G. Strauss, and J.H. Strauss, *Complete sequence of the genomic RNA of O'nyong-nyong virus and its use in the construction of alphavirus phylogenic trees.* Virol J, 1990. **175**: p. 110-123.
28. Olaleye, O.D., S.A. Omilabu, and A.H. Fagbami, *Igbo-Ora (an alphavirus isolated in Nigeria): a seriological surveyfor haemaglutination inhibiting antibody in humans and domestic animals.* transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1988. **82**: p. 905-906.
29. Lanciotti, R.S., et al., *Emergence of Epidemic O'nyong-nyong Fever in Uganda after a 35-year absence: Genetic characterisation of the Virus.* Virol J, 1998. **252**: p. 10.
30. Williams, M.C., et al., *O'nyong-nyong Fever: an epidemic virus disease in East Africa. 8. Virus isolations from Anopheles mosquitoes.* transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1965. **59**: p. 300-306.
31. Marshall, C., et al., *the epidemiology of O'nyong-nyong in the Kano Plain, Kenya.* Annals of Tropical Medicine and Parasitology, 1982. **76**(2): p. 153-158.
32. Strauss, E.G., et al., *Nonstructural proteins nsP3 and nsP4 of Ross River and O'nyong-nyong viruses: sequence comparison with this of other alphaviruses.* Virol J, 1988. **164**: p. 265-274.
33. Wang, G., et al., *Infection of cells by Sindbis virus at low temperature.* Virology, 2007. **362**(2): p. 461-7.

34. Sawicki, D.L. and S.G. Sawicki, *Alphavirus positive and negative strand RNA synthesis and the role of polyproteins in formation of viral replication complexes.* arch Virol Suppl, 1994. **9**: p. 393-405.
35. Park, E. and D.E. Griffin, *The nsP3 macro domain is important for Sindbis virus replication in neurons and neurovirulence in mice.* Virology, 2009. **388**(2): p. 305-14.
36. Carleton, M., D.T. Brown, and *Disulfide bridge-mediated folding of Sindbis virus glycoproteins. J Virol. 70(8):5541-7.* j virol, 1996. **70**(8): p. 5541-7.
37. Miller, M.L. and D.T. Brown, *Morphogenesis of Sindbis virus in three subclones of Aedes albopictus (mosquito) cells. .* j virol, 1992. **66**(7): p. 4180-90.
38. Wang, H., et al., *Effects of inducing or inhibiting apoptosis on Sindbis virus replication in mosquito cells.* J Gen Virol, 2008. **89**(Pt 11): p. 2651-61.
39. Stollar, V., K. Harrap, V. Thomas., Sarver, N. (1979). , *Observations related to cytopathic effect in Aedes albopictus cells infected with Sindbis virus, Arctic and tropical arboviruses.* . 1979, Academic Press, Inc.,New York.: In E. Kurstak (ed.).
40. Mudiganti, U., et al., *Sindbis virus infection of two model insect cell systems--a comparative study. .* virus res, 2006. **122**(1-2): p. 28-34.
41. Foy, B.D., et al., *Development of a new Sindbis virus transducing system and its characterization in three Culicine mosquitoes and two Lepidopteran species.* Insect Mol Biol, 2004. **13**(1): p. 89-100.
42. Brault, A.C., et al., *Infection patterns of o'nyong nyong virus in the malaria-transmitting mosquito, Anopheles gambiae.* Insect Mol Biol, 2004. **13**(6): p. 625-35.
43. Higgs, S., et al., *Mosquito sensitivity to a scorpion neurotoxin expressed using an infectious Sindbis virus vector.* Insect Mol Biol, 1995. **4**(2): p. 97-103.
44. de Lara Capurro, M., et al., *Virus-expressed, recombinant single-chain antibody blocks sporozoite infection of salivary glands in Plasmodium gallinaceum-infected Aedes aegypti.* Am J Trop Med Hyg, 2000. **62**(4): p. 427-33.
45. Vanlandingham, D.L., et al., *Determinants of vector specificity of o'nyong nyong and chikungunya viruses in Anopheles and Aedes mosquitoes.* Am J Trop Med Hyg, 2006. **74**(4): p. 663-9.
46. Johnson, B.W., et al., *Inhibition of luciferase expression in transgenic Aedes aegypti mosquitoes by Sindbis virus expression of antisense luciferase RNA.* Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13399-403.
47. Olson, K.E., et al., *Genetically engineered resistance to dengue-2 virus transmission in mosquitoes.* Science, 1996. **272**(5263): p. 884-6.
48. Dimopoulos, G., *Insect immunity and its implication in mosquito-malaria interactions.* Cell Microbiol, 2003. **5**(1): p. 3-14.
49. Yassine, H. and M.A. Osta, *Anopheles gambiae innate immunity.* Cell Microbiol. **12**(1): p. 1-9.
50. Cherry, S. and N. Silverman, *Host-pathogen interactions in drosophila: new tricks from an old friend.* Nat Immunol, 2006. **7**(9): p. 911-7.
51. Fritz, J.H., S.E. Girardin, and D.J. Philpott, *Innate immune defense through RNA interference.* Sci STKE, 2006. **2006**(339): p. pe27.

52. Galiana-Arnoux, D., S. Deddouche, and J.L. Imler, *[Antiviral immunity in drosophila.].* J Soc Biol, 2007. **201**(4): p. 359-65.
53. Ip, Y.T., *Drosophila innate immunity goes viral.* Nat Immunol, 2005. **6**(9): p. 863-4.
54. Royet, J., J. Reichhart, and J. Hoffmann, *Sensing and signaling during infection in Drosophila. .* curr opin immunol, 2005. **17**(1): p. 11-7.
55. Zambon, R.A., et al., *The Toll pathway is important for an antiviral response in Drosophila.* PNAS, 2005. **102**(29): p. 7252-7262.
56. Dostert, C., et al., *The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of drosophila.* Nat Immunol, 2005. **6**(9): p. 946-53.
57. Galiana-Arnoux, D., et al., *Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila.* Nature Immunology, 2006. **7**(6): p. 590-597.
58. Wang, X.-H., et al., *RNA Interference Directs Innate Immunity Against Viruses in Adult Drosophila.* Science, 2006. **312**: p. 452-454.
59. Zambon, R.A., V.N. Vakharia, and L.P. Wu, *RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster.* Cellular Microbiology, 2006. **8**(5): p. 880-889.
60. van Rij, R.P., et al., *The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster.* Genes Dev, 2006. **20**(21): p. 2985-95.
61. Chotkowski, H.L., et al., *West Nile virus infection of Drosophila melanogaster induces a protective RNAi response.* Virology, 2008. **377**(1): p. 197-206.
62. Osta, M.A., G.K. Christophides, and F.C. Kafatos, *Effects of mosquito genes on Plasmodium development.* Science, 2004. **303**(5666): p. 2030-2.
63. Habtewold, T., et al., *Transmission blocking immunity in the malaria non-vector mosquito Anopheles quadriannulatus species A.* PLoS Pathog, 2008. **4**(5): p. e1000070.
64. Riehle, M.M., et al., *Natural malaria infection in Anopheles gambiae is regulated by a single genomic control region.* Science, 2006. **312**(5773): p. 577-9.
65. Volz, J., et al., *The roles of two clip domain serine proteases in innate immune responses of the malaria vector Anopheles gambiae.* j biol chem, 2005. **280**(48): p. 40161-8.
66. Michel, K., et al., *Anopheles gambiae SRPN2 facilitates midgut invasion by the malaria parasite Plasmodium berghei.* EMBO Rep, 2005. **6**(9): p. 891-7.
67. Volz, J., et al., *A genetic module regulates the melanization response of Anopheles to Plasmodium.* Cell Microbiol, 2006. **8**(9): p. 1392-405.
68. Binari, R. and N. Perrimon, *Stripe-specific regulation of pair-rule genes by hopscotch, a putative Jak family tyrosine kinase in Drosophila.* Genes Dev, 1994. **8**(3): p. 300-12.
69. Zeidler, M.P., E.A. Bach, and N. Perrimon, *The roles of the Drosophila JAK/STAT pathway.* Oncogene, 2000. **19**(21): p. 2598-606.
70. Agaisse, H. and N. Perrimon, *The roles of JAK/STAT signaling in Drosophila immune responses.* Immunological Reviews, 2004. **198**(72-82).

71. Lin, C.C., et al., *Characterization of two mosquito STATs, AaSTAT and CtSTAT. Differential regulation of tyrosine phosphorylation and DNA binding activity by lipopolysaccharide treatment and by Japanese encephalitis virus infection.* j biol chem, 2004. **279**(5): p. 3308-17.
72. Xi, Z., J.L. Ramirez, and G. Dimopoulos, *The Aedes aegypti toll pathway controls dengue virus infection.* PLoS Pathog, 2008. **4**(7): p. e1000098.
73. Mukherjee, T., U. Schafer, and M.P. Zeidler, *Identification of Drosophila genes modulating janus kinase/signal transducer and activator of transcription signal transduction.* Genetics, 2006. **172**(3): p. 1683-97.
74. Baeg, G.H., R. Zhou, and N. Perrimon, *Genome-wide RNAi analysis of JAK/STAT signaling components in Drosophila.* Genes Dev, 2005. **19**(16): p. 1861-70.
75. Souza-Neto, J.A., S. Sim, and G. Dimopoulos, *An evolutionary conserved function of the JAK-STAT pathway in anti-dengue defense.* Proc Natl Acad Sci U S A, 2009. **106**(42): p. 17841-6.
76. Fragkoudis, R., et al., *Semliki Forest virus strongly reduces mosquito host defence signaling.* Insect Mol Biol, 2008.
77. Campbell, C.L., et al., *Comparative genomics of small RNA regulatory pathway components in vector mosquitoes.* BMC Genomics, 2008. **9**: p. 425.
78. Meister, G. and T. Tuschl, *Mechanisms of gene silencing by double-stranded RNA.* Nature Reviews Microbiology, 2004. **431**(7006): p. 343-9.
79. Gracheva, E., M. Dus, and S.C. Elgin, *Drosophila RISC Component VIG and Its Homolog Vig2 Impact Heterochromatin Formation.* PLoS One, 2009. **4**(7): p. e6182.
80. Keene, K., et al., *RNA interference acts as a natural antiviral response to O'nyong-nyong virus (Alphavirus; Togaviridae) infection of Anopheles gambiae.* PNAS, 2005. **101**(49): p. 17240-17245.
81. Hammond, S.M., et al., *Argonaute2, a link between genetic and biochemical analyses of RNAi.* Science, 2001. **293**(5532): p. 1146-50.
82. Campbell, C.L., et al., *Aedes aegypti uses RNA interference in defense against Sindbis virus infection.* BMC Microbiol, 2008. **8**: p. 47.
83. Sanchez-Vargas, I., et al., *Dengue virus type 2 infections of Aedes aegypti are modulated by the mosquito's RNA interference pathway.* PLoS Pathog, 2009. **5**(2): p. e1000299.
84. Myles, K.M., et al., *Alphavirus-derived small RNAs modulate pathogenesis in disease vector mosquitoes.* Proc Natl Acad Sci U S A, 2008. **105**(50): p. 19938-43.
85. Cirimotich, C.M., et al., *Suppression of RNA interference increases alphavirus replication and virus-associated mortality in Aedes aegypti mosquitoes.* BMC Microbiol, 2009. **9**: p. 49.
86. Xie, Q. and H.S. Guo, *Systemic antiviral silencing in plants.* Virus Res, 2006. **118**(1-2): p. 1-6.
87. Saleh, M.C., et al., *Antiviral immunity in Drosophila requires systemic RNA interference spread.* Nature, 2009. **458**(7236): p. 346-50.
88. Attarzadeh-Yazdi, G., et al., *Cell-to-cell spread of the RNA interference response suppresses Semliki Forest virus (SFV) infection of mosquito cell cultures and cannot be antagonized by SFV.* J Virol, 2009. **83**(11): p. 5735-48.

89. Pacca, C.C., et al., *RNA interference inhibits yellow fever virus replication in vitro and in vivo.* Virus Genes, 2009.
90. Cullen, B.R., *Is RNA interference involved in intrinsic antiviral immunity in mammals?* Nat Immunol, 2006. **7**(6): p. 563-7.
91. Hashimoto, C., K.L. Hudson, and K.V. Anderson, *The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein.* cell, 1988. **52**(2): p. 269-79.
92. Bilak, H., S. Tauszig-Delamasure, and J.L. Imler, *Toll and Toll-like receptors in Drosophila.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 648-51.
93. Gobert, V., et al., *Dual activation of the Drosophila toll pathway by two pattern recognition receptors.* Science, 2003. **302**(5653): p. 2126-30.
94. Weber, A.N., et al., *Binding of the Drosophila cytokine Spatzle to Toll is direct and establishes signaling.* Nat Immunol, 2003. **4**(8): p. 794-800.
95. Christophides, G.K., D. Vlachou, and F.C. Kafatos, *Comparative and functional genomics of the innate immune system in the malaria vector Anopheles gambiae.* Immunol Rev, 2004. **198**: p. 127-48.
96. Rutschmann, S., et al., *The Rel protein DIF mediates the antifungal but not the antibacterial host defense in Drosophila.* Immunity, 2000. **12**(5): p. 569-80.
97. Sanders, H.R., et al., *Sindbis virus induces transport processes and alters expression of innate immunity pathway genes in the midgut of the disease vector, Aedes aegypti.* Insect Biochem Mol Biol, 2005. **35**(11): p. 1293-307.
98. Sinden, R.E., G.A. Butcher, and A.L. Beetsma, *Maintenance of the Plasmodium berghei life cycle.* . methods mol med, 2002. **72**: p. 25-40.
99. Muller, H.M., et al., *A hemocyte-like cell line established from the malaria vector Anopheles gambiae expresses six prophenoloxidase genes.* j biol chem, 1999. **274**(17): p. 11727-35.
100. Barillas-Mury, C., et al., *Anopheles gambiae Ag-STAT, a new insect member of the STAT family, is activated in response to bacterial infection.* EMBO Journal, 1999. **18**(4): p. 959-967.
101. Gusmao, D.S., et al., *First isolation of microorganisms from the gut diverticulum of Aedes aegypti (Diptera: Culicidae): new perspectives for an insect-bacteria association.* Mem Inst Oswaldo Cruz, 2007. **102**(8): p. 919-24.
102. Domingo, E., *Rapid evolution of viral RNA genomes.* J Nutr, 1997. **127**(5 Suppl): p. 958S-961S.
103. Borucki, M.K., et al., *La Crosse virus: replication in vertebrate and invertebrate hosts.* Microbes Infect, 2002. **4**(3): p. 341-50.
104. Griffin, D.E. and J.M. Hardwick, *Regulators of apoptosis on the road to persistent alphavirus infection.* Annu Rev Microbiol, 1997. **51**: p. 565-92.
105. Bowers, D.F., C.G. Coleman, and D.T. Brown, *Sindbis virus-associated pathology in Aedes albopictus (Diptera: Culicidae).* J Med Entomol, 2003. **40**(5): p. 698-705.
106. Girard, Y.A., et al., *Ultrastructural study of West Nile virus pathogenesis in Culex pipiens quinquefasciatus (Diptera: Culicidae).* J Med Entomol, 2005. **42**(3): p. 429-44.
107. Girard, Y.A., et al., *Salivary gland morphology and virus transmission during long-term cytopathologic West Nile virus infection in Culex mosquitoes.* Am J Trop Med Hyg, 2007. **76**(1): p. 118-28.

108. Weaver, S.C., et al., *Togavirus-associated pathologic changes in the midgut of a natural mosquito vector.* J Virol, 1988. **62**(6): p. 2083-90.
109. Weaver, S.C., L.H. Lorenz, and T.W. Scott, *Pathologic changes in the midgut of Culex tarsalis following infection with Western equine encephalomyelitis virus.* Am J Trop Med Hyg, 1992. **47**(5): p. 691-701.
110. Fragkoudis, R., et al., *Advances in dissecting mosquito innate immune responses to arbovirus infection.* J Gen Virol, 2009. **90**(Pt 9): p. 2061-72.
111. Knight, R.L., et al., *Role of N-linked glycosylation for sindbis virus infection and replication in vertebrate and invertebrate systems.* J Virol, 2009. **83**(11): p. 5640-7.
112. Shabman, R.S., et al., *Differential induction of type I interferon responses in myeloid dendritic cells by mosquito and mammalian-cell-derived alphaviruses.* J Virol, 2007. **81**(1): p. 237-47.
113. Shabman, R.S., K.M. Rogers, and M.T. Heise, *Ross River virus envelope glycans contribute to type I interferon production in myeloid dendritic cells.* J Virol, 2008. **82**(24): p. 12374-83.
114. Luo, T. and D.T. Brown, *Purification and characterization of a Sindbis virus-induced peptide which stimulates its own production and blocks virus RNA synthesis.* Virology, 1993. **194**(1): p. 44-9.
115. Luo, T. and D.T. Brown, *A 55-kDa protein induced in Aedes albopictus (mosquito) cells by antiviral protein.* Virology, 1994. **200**(1): p. 200-6.
116. Kim, K.H., et al., *Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts.* Virology, 2004. **323**(1): p. 153-63.
117. Blandin, S., et al., *Reverse genetics in the mosquito Anopheles gambiae: targeted disruption of the Defensin gene. .* EMBO Journal, 2002. **3**(9): p. 852-6.
118. Bowers, D.F., B.A. Abell, and D.T. Brown, *Replication and tissue tropism of the alphavirus Sindbis in the mosquito Aedes albopictus.* Virology, 1995. **212**(1): p. 1-12.
119. Pierro, D.J., E.L. Powers, and K.E. Olson, *Genetic determinants of Sindbis virus strain TR339 affecting midgut infection in the mosquito Aedes aegypti.* J Gen Virol, 2007. **88**(Pt 5): p. 1545-54.
120. Salazar, M.I., et al., *Dengue virus type 2: replication and tropisms in orally infected Aedes aegypti mosquitoes.* BMC Microbiol, 2007. **7**: p. 9.
121. Weaver, S.C., et al., *Detection of eastern equine encephalomyelitis virus deposition in Culiseta melanura following ingestion of radiolabeled virus in blood meals.* Am J Trop Med Hyg, 1991. **44**(3): p. 250-9.
122. Johnson, B.K., et al., *Recovery of O'nyong-nyong virus from Anopheles funestus in Western Kenya.* transactions of the Royal Scoiety of Tropical Medicine and Hygiene, 1981. **75**(2): p. 239-241.
123. Ayala, F.J. and M. Coluzzi, *Chromosome speciation: humans, Drosophila, and mosquitoes.* Proc Natl Acad Sci U S A, 2005. **102 Suppl 1**: p. 6535-42.
124. Christophides, G.K., et al., *Immunity-related genes and gene families in Anopheles gambiae.* Science, 2002. **298**(5591): p. 159-65.

125. Dimopoulos, G., et al., *Genome Gexpression analysis of Anopheles gambiae: responses to injury, bacterial challenge, and malaria infection. .* Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8814-9.
126. Povelones, M., et al., *Leucine-rich repeat protein complex activates mosquito complement in defense against Plasmodium parasites.* Science, 2009. **324**(5924): p. 258-61.
127. Mitri, C., et al., *Fine pathogen discrimination within the APL1 gene family protects Anopheles gambiae against human and rodent malaria species.* PLoS Pathog, 2009. **5**(9): p. e1000576.
128. Riehle, M.M., et al., *Anopheles gambiae APL1 is a family of variable LRR proteins required for Rel1-mediated protection from the malaria parasite, Plasmodium berghei.* PLoS One, 2008. **3**(11): p. e3672.
129. Sim C, H.Y., et al., *Modulation of Anopheles gambiae gene expression in response to o'nyong-nyong virus infection.* Insect Biochemistry and Molecular Biology, 2005. **14**(5): p. 475-81.
130. Wang, L., et al., *A genetic screen identifies new regulators of steroid-triggered programmed cell death in Drosophila.* Genetics, 2008. **180**(1): p. 269-81.
131. Belinda, L.W., et al., *SARM: a novel Toll-like receptor adaptor, is functionally conserved from arthropod to human.* Mol Immunol, 2008. **45**(6): p. 1732-42.
132. Dalod, M., *Studies of SARM1 uncover similarities between immune and neuronal responses to danger.* Sci STKE, 2007. **2007**(417): p. pe73.
133. Peng, J., et al., *SARM inhibits both TRIF- and MyD88-mediated AP-1 activation.* Eur J Immunol. **40**(6): p. 1738-47.
134. Dong, Y., et al., *Anopheles gambiae immune responses to human and rodent Plasmodium parasite species.* PLoS Pathog, 2006. **2**(6): p. e52.
135. Meister, S., et al., *Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito Anopheles gambiae.* Proc Natl Acad Sci U S A, 2005. **102**(32): p. 11420-5.
136. Cherry, S., et al., *COPI activity coupled with fatty acid biosynthesis is required for viral replication.* PLoS Pathog, 2006. **2**(10): p. e102.
137. Li, Y., et al., *Fatty acid synthase expression is induced by the Epstein-Barr virus immediate-early protein BRLF1 and is required for lytic viral gene expression.* J Virol, 2004. **78**(8): p. 4197-206.
138. Fukasawa, M., et al., *Enhancement of de novo fatty acid biosynthesis in hepatic cell line Huh7 expressing hepatitis C virus core protein.* Biol Pharm Bull, 2006. **29**(9): p. 1958-61.
139. Yang, W., et al., *Fatty acid synthase is up-regulated during hepatitis C virus infection and regulates hepatitis C virus entry and production.* Hepatology, 2008. **48**(5): p. 1396-403.
140. Slabas, A.R., et al., *Pivotal reactions in fatty acid synthesis.* Prog Lipid Res, 1994. **33**(1-2): p. 39-46.
141. Bryan, B.A., et al., *The Rho family of small GTPases: crucial regulators of skeletal myogenesis.* Cell Mol Life Sci, 2005. **62**(14): p. 1547-55.
142. Shiao, S.H., et al., *Fz2 and cdc42 mediate melanization and actin polymerization but are dispensable for Plasmodium killing in the mosquito midgut.* PLoS Pathog, 2006. **2**(12): p. e133.

143. Zamudio-Meza, H., et al., *Cross-talk between Rac1 and Cdc42 GTPases regulates formation of filopodia required for dengue virus type-2 entry into HMEC-1 cells.* J Gen Virol, 2009. **90**(Pt 12): p. 2902-11.
144. Caudy, A.A. and G.J. Hannon, *Induction and biochemical purification of RNA-induced silencing complex from Drosophila S2 cells.* Methods Mol Biol, 2004. **265**: p. 59-72.
145. Hartig, J.V., Y. Tomari, and K. Forstemann, *piRNAs--the ancient hunters of genome invaders.* Genes Dev, 2007. **21**(14): p. 1707-13.
146. Choudhuri, S., *Lesser known relatives of miRNA.* Biochem Biophys Res Commun, 2009. **388**(2): p. 177-80.
147. Dunoyer, P., et al., *Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing.* Plant Cell, 2004. **16**(5): p. 1235-50.
148. Voinnet, O., *Induction and suppression of RNA silencing: insights from viral infections.* Nat Rev Genet, 2005. **6**(3): p. 206-20.
149. Peranen, J., et al., *Nuclear localization of Semliki Forest virus-specific nonstructural protein nsP2.* J Virol, 1990. **64**(5): p. 1888-96.
150. Ulvila, J., D. Hultmark, and M. Ramet, *RNA Silencing in the Antiviral Innate Immune Defence – Role of DEAD-box RNA Helicases.* Scandinavian journal of immunology, 2009. **71**(3): p. 146-158.
151. Ulvila, J., et al., *Double-stranded RNA is internalized by scavenger receptor-mediated endocytosis in Drosophila S2 cells.* j biol chem, 2006. **281**(20): p. 14370-5.
152. Kim, V.N., *Small RNAs: classification, biogenesis, and function.* Mol Cells, 2005. **19**(1): p. 1-15.
153. Schroder, M., *Human DEAD-box protein 3 has multiple functions in gene regulation and cell cycle control and is a prime target for viral manipulation.* Biochem Pharmacol. **79**(3): p. 297-306.
154. Lu, Y., L.P. Wu, and K.V. Anderson, *The antibacterial arm of the drosophila innate immune response requires an IkappaB kinase.* Genes Dev, 2001. **15**(1): p. 104-10.
155. Tanji, T. and Y.T. Ip, *Regulators of the Toll and Imd pathways in the Drosophila innate immune response.* Trends Immunol, 2005. **26**(4): p. 193-8.
156. Georgel, P., et al., *Drosophila immune deficiency (IMD) is a death domain protein that activates antibacterial defense and can promote apoptosis.* Dev Cell, 2001. **1**(4): p. 503-14.
157. Avadhanula, V., et al., *A novel system for the launch of alphavirus RNA synthesis reveals a role for the imd pathway in arthropod antiviral response.* PLoS Pathog, 2009. **5**(9): p. e1000582.
158. Costa, A., et al., *The Imd pathway is involved in antiviral immune responses in Drosophila.* PLoS One, 2009. **4**(10): p. e7436.
159. Pace, K.E. and L.G. Baum, *Insect galectins: roles in immunity and development.* Glycoconj J, 2004. **19**(7-9): p. 607-14.
160. Vasta, G.R., *Roles of galectins in infection.* Nat Rev Microbiol, 2009. **7**(6): p. 424-38.
161. Dimopoulos, G., et al., *Innate immune defense against malaria infection in the mosquito.* curr opin immunol, 2001. **13**(1): p. 79-88.

162. Almkvist, J., et al., *Newcastle disease virus neuraminidase primes neutrophils for stimulation by galectin-3 and formyl-Met-Leu-Phe.* Exp Cell Res, 2004. **298**(1): p. 74-82.
163. Gandhi, M.K., et al., *Galectin-1 mediated suppression of Epstein-Barr virus specific T-cell immunity in classic Hodgkin lymphoma.* Blood, 2007. **110**(4): p. 1326-9.
164. Mengshol, J.A., et al., *A crucial role for Kupffer cell-derived galectin-9 in regulation of T cell immunity in hepatitis C infection.* PLoS One. **5**(3): p. e9504.
165. Garcia-Pineres, A.J., et al., *Gene expression patterns induced by HPV-16 L1 virus-like particles in leukocytes from vaccine recipients.* J Immunol, 2009. **182**(3): p. 1706-29.
166. King, R.D., J.M. Lubinski, and H.M. Friedman, *Herpes simplex virus type 1 infection increases the carbohydrate binding activity and the secretion of cellular galectin-3.* Arch Virol, 2009. **154**(4): p. 609-18.
167. Waterhouse, R.M., et al., *Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes.* Science, 2007. **316**(5832): p. 1738-43.
168. Bryant, C.E., et al., *The molecular basis of the host response to lipopolysaccharide.* Nat Rev Microbiol. **8**(1): p. 8-14.
169. Visintin, A., et al., *MD-2.* Immunobiology, 2006. **211**(6-8): p. 437-47.
170. Inohara, N. and G. Nunez, *ML -- a conserved domain involved in innate immunity and lipid metabolism.* Trends Biochem Sci, 2002. **27**(5): p. 219-21.
171. Okumura, A., et al., *Interaction between Ebola virus glycoprotein and host toll-like receptor 4 leads to induction of proinflammatory cytokines and SOCS1.* J Virol. **84**(1): p. 27-33.
172. Lagos, D., et al., *Toll-like receptor 4 mediates innate immunity to Kaposi sarcoma herpesvirus.* Cell Host Microbe, 2008. **4**(5): p. 470-83.
173. Luna, C., et al., *Characterization of four Toll related genes during development and immune responses in Anopheles gambiae.* Insect Biochem Mol Biol, 2002. **32**(9): p. 1171-9.
174. Li, B., et al., *Characterization of the c-type lysozyme gene family in Anopheles gambiae.* Gene, 2005. **360**(2): p. 131-9.
175. Ibrahim, H.R., T. Matsuzaki, and T. Aoki, *Genetic evidence that antibacterial activity of lysozyme is independent of its catalytic function.* FEBS Lett, 2001. **506**(1): p. 27-32.
176. During, K., et al., *The non-enzymatic microbicidal activity of lysozymes.* FEBS Lett, 1999. **449**(2-3): p. 93-100.
177. Masschalck, B. and C.W. Michiels, *Antimicrobial properties of lysozyme in relation to foodborne vegetative bacteria.* Crit Rev Microbiol, 2003. **29**(3): p. 191-214.
178. Lee-Huang, S., et al., *Lysozyme and RNases as anti-HIV components in beta-core preparations of human chorionic gonadotropin.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2678-81.
179. zhang, X., et al., *Purification and properties of lysozyme from a marine strain* Annals of Microbiology, 2008. **58**(1): p. 89-94.
180. De Zulueta, J., et al., *An observation on the possible effect of O'Nyong-Nyong fever on malaria.* Bull World Health Organ, 1962. **26**: p. 135-9.

181. Sinden, R.E. and P.F. Billingsley, *Plasmodium invasion of mosquito cells: hawk or dove?* Trends Parasitol, 2001. **17**(5): p. 209-12.
182. Zaidman-Remy, A., et al., *The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection.* Immunity, 2006. **24**(4): p. 463-73.
183. Schnitger, A.K., et al., *Two C-type lectins cooperate to defend Anopheles gambiae against Gram-negative bacteria.* j biol chem, 2009. **284**(26): p. 17616-24.
184. Kumar, S., et al., *The role of reactive oxygen species on Plasmodium melanotic encapsulation in Anopheles gambiae.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14139-44.
185. Rehwinkel, J. and C. Reis e Sousa, *RIGorous detection: exposing virus through RNA sensing.* Science. **327**(5963): p. 284-6.
186. Meurs, E., et al., *Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon.* cell, 1990. **62**(2): p. 379-90.
187. Zhou, A., et al., *Interferon action and apoptosis are defective in mice devoid of 2',5'-oligoadenylate-dependent RNase L.* EMBO J, 1997. **16**(21): p. 6355-63.
188. Meylan, E. and J. Tschopp, *Toll-like receptors and RNA helicases: two parallel ways to trigger antiviral responses.* Mol Cell, 2006. **22**(5): p. 561-9.
189. Deddouche, S., et al., *The DExD/H-box helicase Dicer-2 mediates the induction of antiviral activity in drosophila.* Nat Immunol, 2008. **9**(12): p. 1425-32.
190. Takeuchi, O. and S. Akira, *RIG-I-like antiviral protein in flies.* Nat Immunol, 2008. **9**(12): p. 1327-8.
191. Jang, I.H., H.J. Nam, and W.J. Lee, *CLIP-domain serine proteases in Drosophila innate immunity.* BMB Rep, 2008. **41**(2): p. 102-7.
192. Gorman, M.J. and S.M. Paskewitz, *Serine proteases as mediators of mosquito immune responses.* Insect Biochem Mol Biol, 2001. **31**(3): p. 257-62.
193. Cerenius, L. and K. Soderhall, *The prophenoloxidase-activating system in invertebrates.* Immunol Rev, 2004. **198**: p. 116-26.
194. Schnitger, A.K., F.C. Kafatos, and M.A. Osta, *The melanization reaction is not required for survival of Anopheles gambiae mosquitoes after bacterial infections.* j biol chem, 2007. **282**(30): p. 21884-8.
195. Seery, L.T., P.V. Nestor, and G.A. FitzGerald, *Molecular evolution of the aldo-keto reductase gene superfamily.* J Mol Evol, 1998. **46**(2): p. 139-46.
196. Ranson, H. and J. Hemingway, *Mosquito glutathione transferases.* Methods Enzymol, 2005. **401**: p. 226-41.
197. Scharf, M.E., et al., *Catalytic activity and expression of two flavin-containing monooxygenases from Drosophila melanogaster.* Arch Insect Biochem Physiol, 2004. **57**(1): p. 28-39.
198. Berge, J.B., R. Feyereisen, and M. Amichot, *Cytochrome P450 monooxygenases and insecticide resistance in insects.* Philos Trans R Soc Lond B Biol Sci, 1998. **353**(1376): p. 1701-5.
199. Ha, E.M., et al., *A direct role for dual oxidase in Drosophila gut immunity.* Science, 2005. **310**(5749): p. 847-50.
200. Mills, E.L., et al., *Initiation of the respiratory burst of human neutrophils by influenza virus.* Infect Immun, 1981. **32**(3): p. 1200-5.

201. Misra, A., R. Mukerjee, and U.C. Chaturvedi, *Respiratory Burst by Dengue-Virus-Induced Cytotoxic Factor.* Med Princ Pract, 1998. **7**: p. 251-260.
202. Allan, A.C., et al., *An early tobacco mosaic virus-induced oxidative burst in tobacco indicates extracellular perception of the virus coat protein.* Plant Physiol, 2001. **126**(1): p. 97-108.
203. Barna, B., et al., *Role of reactive oxygen species and antioxidants in plant disease resistance.* Pest Manag Sci, 2003. **59**(4): p. 459-64.
204. Dong, Y. and G. Dimopoulos, *Anopheles fibrinogen-related proteins provide expanded pattern recognition capacity against bacteria and malaria parasites.* j biol chem, 2009. **284**(15): p. 9835-44.
205. Gokudan, S., et al., *Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen.* Proc Natl Acad Sci U S A, 1999. **96**(18): p. 10086-91.
206. Wang, X., Q. Zhao, and B.M. Christensen, *Identification and characterization of the fibrinogen-like domain of fibrinogen-related proteins in the mosquito, Anopheles gambiae, and the fruitfly, Drosophila melanogaster, genomes.* BMC Genomics, 2005. **6**: p. 114.
207. Wang, X., et al., *A novel lectin with a fibrinogen-like domain and its potential involvement in the innate immune response of Armigeres subalbatus against bacteria.* Insect Mol Biol, 2004. **13**(3): p. 273-82.
208. Abram, C.L. and C.A. Lowell, *The expanding role for ITAM-based signaling pathways in immune cells.* Sci STKE, 2007. **2007**(377): p. re2.
209. Kobe, B. and A.V. Kajava, *The leucine-rich repeat as a protein recognition motif.* Curr Opin Struct Biol, 2001. **11**(6): p. 725-32.
210. Bell, J.K., et al., *Leucine-rich repeats and pathogen recognition in Toll-like receptors.* Trends Immunol, 2003. **24**(10): p. 528-33.
211. Pancer, Z., et al., *Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey.* Nature, 2004. **430**(6996): p. 174-80.
212. Bell, J.K., et al., *The molecular structure of the Toll-like receptor 3 ligand-binding domain.* Proc Natl Acad Sci U S A, 2005. **102**(31): p. 10976-80.
213. Takeda, K. and S. Akira, *Toll receptors and pathogen resistance.* Cell Microbiol, 2003. **5**(3): p. 143-53.
214. Padmanabhan, M., P. Cournoyer, and S.P. Dinesh-Kumar, *The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities.* Cell Microbiol, 2009. **11**(2): p. 191-8.
215. Abraham, E.G., et al., *An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites.* Proc Natl Acad Sci U S A, 2005. **102**(45): p. 16327-32.
216. Saleh, M.C., et al., *The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing.* Nat Cell Biol, 2006. **8**(8): p. 793-802.
217. Hoa, N.T., et al., *Characterization of RNA interference in an Anopheles gambiae cell line.* Insect Biochem Mol Biol, 2003. **33**(9): p. 949-57.
218. Dimopoulos, G., et al., *Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malaria parasites.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11508-13.

219. Horton, A.A., et al., *Identification of three single nucleotide polymorphisms in Anopheles gambiae immune signaling genes that are associated with natural Plasmodium falciparum infection.* Malar J. **9**(1): p. 160.

"To the reader who asks, 'What can you possibly see to interest you in these miserable creatures?' or 'So you work on mosquitoes, so what?' (according to what side of the Atlantic he is on) I would say I find it sobering to reflect that mosquitoes have at times played a more effective part in determining the course of human history than the generals and other leaders who set themselves up to guide our fortunes. But when the *inevitable* question is asked, I can only ask in return, 'What *use* is a man?'

"I think the present time may well go down in history as the period when man sought to conquer and control nature as if he himself were something apart, when the ignorance of the nineteenth century was replaced by the arrogance of the twentieth. But I question the whole present-day attitude that defines progress and civilisation as conquests over nature. Progress can come only from an understanding of and attempt to co-operate with the forces of nature in the full realisation that we ourselves as well as the mosquitoes that bite us are moulded by these very forces and are as inseparable from them as the clouds from the sky."

J.D. Gillett, 1971

Appendix 1

Table 1. Genes regulated by 5'ONNVic-eGFP infection in *A.gambiae* mosquitoes. Two-fold or greater fold change ratios are shown in black text for one day postinfection (D1), 4 days post infection (D4) and nine days post infection (D9). Fold change ratios less than 2-fold regulated that have passed all filters outlines in figure 20 excluding filtering on fold change ratio, are shown in grey text. Putative functions/functional domains were derived from Gene ontology terms, Interpro domains and functions of orthologous genes (www.vectorbase.org).

AGAP	D1	D4	D9	Common name	Functional group	Putative function
Early onset						
AGAP001899	2.03				Fatty acid synthesis	Fatty acid synthase
AGAP000260	2.08				Housekeeping	ATPase, F1 complex, epsilon subunit, mitochondrial
AGAP010814	2.60			TEP5	Immunity	Thioester containing protein
AGAP006348	2.28			LRIM1	immunity	Leucine rich repeat protein
AGAP004248	2.27			GPXH3	Immunity	peroxidase, GPX sub familiy
AGAP002848	2.01			ML9	Immunity	MD2-like receptor
AGAP005901	0.48				immunity/apoptosis?	Sterile alpha and TIR motif containing protein (d.mel ortho), also had Armadillo repeats
AGAP001116	2.02				metabolism	FAD dependent oxidoreductase
AGAP006009	2.49			CPR30	Misc	insect cuticle protein,structural
AGAP006958	0.48				Misc	Heat shock protein
AGAP010895	0.46	0.51			Misc	D.mel ortho beta spectrin, cytoskeletal protein with diverse function
AGAP005913	0.45				Signalling	G-protein with WD-40 repeats
AGAP005912	0.45				Signalling	G-protein with WD-40 repeats
AGAP005911	0.39				Signalling	G-protein with WD-40 repeats
AGAP002171	0.48				Translation	Nucleolar protein nop5 (d.mel ortho), rRNA processing factor
AGAP003773	4.37				unknown	unknown
AGAP003778	4.12				unknown	unknown
AGAP003777	4.06				unknown	unknown
AGAP003939	2.83		1.85		unknown	protein binding, unknown
AGAP003775	2.65				unknown	unknown, PA fragment
AGAP001078	2.28				unknown	unknown
AGAP009974	2.12				unknown	unknown
AGAP008447-RA						
AGAP008444-RA	2.11				unknown	unknown
	23.00					

Early to mid onset						
AGAP008118-RA	0.48	0.35			Cell division	Cell cycle checkpoint kinase, d.mel ortho grp
AGAP004556-RA	0.46	0.32			Cell division	d mel ortho peter pan (ppan), required for larval growth, roles in mitosis and cell growth (mutants show reduced growth)
AGAP009176-RA	2.21	2.19	1.27		Fatty acid synthesis	Fatty acid synthase
AGAP003196-RA	2.17	2.01			Fatty acid synthesis	GNS1/SUR4 membrane protein - unknown function/poss fatty acid elongation
AGAP005620	2.02	3.60	1.66	DPT	Immunity	Dipteracin, anti-microbial peptide
AGAP004845	2.01	3.42		SCRB8	Immunity	Scavenger receptor, cell adhesion,
AGAP012352	2.31	2.86	1.61	ML1	Immunity	MD2-like lipid recognition
AGAP009556	2.10	2.82		FREP50	Immunity	Fibrinogen-like
no AGAP	2.47	2.77	1.44	CLIPA9	Immunity	Clip domain serine protease
AGAP010812-RA	2.88	2.73	1.61	TEP4	Immunity	Thioester containing protein
AGAP008654	2.15	2.58	1.68	TEP12	immunity	Thioester containing protein
AGAP007457	2.10	2.38			Immunity	LRR protein
AGAP004455	2.10	2.55		GNBPB1	Immunity	Gram negative binding protein subgroup B
AGAP011790	2.09	2.36		CLIPA2	Immunity	Clip domain serine protease
AGAP003247	2.14	2.17		CLIPB19	Immunity	Clip domain serine protease
AGAP010819	2.53	2.15		TEP10	Immunity	Thioester containing protein
AGAP010830	2.23	2.06		TEP9	Immunity	Thioester containing protein
AGAP007315	2.04	2.69			metabolism	nitrogen compound metabolism
AGAP005372	3.15	3.01		COEBE3D	Misc	Carboxylesterase, type B
AGAP012320-RA	2.32	2.48		OBP25	Misc	oderant binding protein
AGAP006278	2.05	2.18	1.85		Misc	insect pheromone/odorant binding protein
AGAP008311-RA	2.46	2.12			Misc	acylphophatase, unknown function
AGAP007160	2.12	2.06			Misc	HSP20 domain
AGAP000930-RA	0.47	0.36			Misc	WD40 repeats - protein-protein interactions, D.mel ortho pod1 (actin/microulue binding)
AGAP006904	0.39	0.48	0.61		Misc	Matrix metalloproteinase (MMP1 d.mel ortho) functions in digestion of ECM, release of apoptotic ligands, diverse functions in cell proliferation, angiogensis, apoptosis and immune defence
AGAP010548-RA	2.21	2.85	1.61		signalling	Laminin B (cell adhesion, EM), EGF-like domain (epidermal-growth factor-like)
AGAP001589-RA	0.49	0.32			Translation	initiation factor 2b-related (translation)
AGAP008306-RA	2.76	3.58			unknown	hypothetical protein, unknown function
AGAP008011-RA	2.10	3.48			unknown	Unknown

AGAP006504	2.51	3.24			unknown	Unknown
AGAP006507	2.67	3.00			unknown	Unknown
AGAP005611	2.15	2.89			unknown	Unknown
AGAP011317-RA	2.05	2.79			unknown	General oderant binding precursor familiy
AGAP010066-RA	2.34	2.70			unknown	Peroxisomal membrane protein
AGAP008307-RA	2.44	2.59			unknown	hypothetical protein, unknown function
AGAP007711	2.03	2.44			unknown	Unknown
AGAP003239-RA	2.42	2.31	1.26		unknown	Unknown
AGAP004208	0.46	0.43			unknown	unknown
	38.00					
Mid onset						
AGAP009792-RA	0.68	0.49			Cell division	d.mel ortho pendulin, needed for cell proliferation, comains armadillo repeats, and an importing domain
AGAP003742-RA		0.47			Cell division	regulator of chromosome condensation
AGAP005800		0.47			Cell division	DNA replication licensing factor MDM7, part of a complex required for initialtion and elongation of replicative forks during S-phase
AGAP006165		0.45			Cell division	ortho d.mel domino; hemocyte proliferation, mRNA splicing, DNA helicase activity
AGAP007874-RA	0.58	0.45			Cell division	Initiation factor eIF-4 gamma, MA3, or cell cycle control
AGAP007477		0.40			Cell division	ATP dependent DNA helicase, DNA replication, mitosis, DNA repair (from d.mel ortho Replication Factor C40, RfC40)
AGAP002440-RA		0.34		CDC42_ANOGA	Cell division	CDC42 - cell division control 42; Rho GTPase, diverse functions including regualtion of the cell cycle, also Miro-like - ribosomal homeostasis and apoptosis functions
AGAP007112	0.50	0.27			Cell division	Cell cycle checkpoint (BRCT dmain) and cell proliferation (pescadillo domain)
AGAP010150-RA	1.94	2.45	1.79		Fatty acid synthesis	Cytochrome b5 related, fattcy acid desaturase domains
AGAP010695-RA		2.31			fatty acid synthesis	GNS1/SUR4 membrane protein - unknown function/poss fatty acid elongation
AGAP008468-RA	1.88	2.06	1.57		fatty acid synthesis	Fatty acid synthase
AGAP010461-RA	0.81	0.47			Histone	Histone H1 ortho
no AGAP	1.87	2.86	1.55	CLIPC9	Imminuty	Clip domain serine protease
no AGAP	1.65	2.45		CLIPE4	Immunity	Clip domain serine protease

AGAP003058	1.64	2.44	1.99	CLIPB9	Immunity	Thioester containing protein
AGAP008368		2.44		TEP14	Immunity	Clip domain serine protease
AGAP010774		2.30		FREP27	Immunity	Fibrinogen-like
AGAP004806		2.28		GALE6	Immunity	galectin, sugar binding
AGAP009221	1.82	2.27		SRPN5	Immunity	Serpin
AGAP005334	1.57	2.25		CTLMA2	Immunity	C-type lectin, sugar binding
AGAP012529		2.21		GALE8	Immunity	galectin, sugar binding
AGAP004807	1.54	2.20		GALE7	Immunity	galectin, sugar binding
AGAP010360-RA	1.47	2.19			Immunity	peritrophin-A chitin binding
AGAP007343		2.19		LYSC2	Immunity	Lysozymes
AGAP005717	1.94	2.18		LYSC6	Immunity	lysozyme
no_AGAP		2.15		FREP10	Immunity	Fibrinogen-like
AGAP004920		2.14		CASPS6	Immunity	Caspase, apoptosis
AGAP003246	1.50	2.12	1.90	CLIPB2	Immunity	Clip domain serine protease
AGAP000694-RA	1.60	2.11	1.76	CEC3	Immunity	AMP
AGAP012945-RA	1.58	2.07			Immunity	Caspase (d.mel Decay)
AGAP000443	1.73	2.07		CTL5	Immunity	C-type lectin, sugar binding
AGAP009166	0.50	0.48		IKK1	Immunity	part of IKK complex required for IMD signalling
AGAP004036		0.48		HXP7	Immunity	Peroxidase
AGAP007294	0.68	0.46		IAP1	Immunity	inhibitor of apoptosis
AGAP004038	0.80	0.45		HXP8	Immunity	Peroxidise
AGAP008354		0.35		HOP	Immunity	Janus kinase, Hopskotch
AGAP010363-RA		4.23			Immunity?	Chitin binding,peritrophin-A
AGAP010364-RA		4.08			Immunity?	peritrophin-A chitin binding
AGAP006707		3.23			metabolism	Chymotrypsin
AGAP005752		2.89			metabolism	Metabolic, transferase activity, transferring hexosyl groups
AGAP010530-RA	1.77	2.63			metabolism	chymotrypsin/Hap
AGAP012034-RA						
AGAP012035-RA		2.32			metabolism	chymotrypsin/Hap domain, unknown gene family
AGAP007142		2.29			metabolism	Peptidase S1A, chymotrypsin, trypsin activity, digestion, metabolism
AGAP006926		2.22			metabolism	alcohol dehydrogenase
AGAP012843-RA	1.84	2.19	1.54		metabolism	Chymotrypsin/Hap
AGAP005124	1.51	2.18			metabolism	Aldehyde dehydrogenase
AGAP002721-RA		2.16			metabolism	tryptophan oxygenase - tryptophan metabolism
AGAP010243-RA		2.04			metabolism	chymotrypsin (aedes ortho of trypsin), complement activation?
AGAP007505		0.47			metabolism	serine carboxypeptidase activity
AGAP011948-RA		0.44			metabolism	Threonine dehydrogenase

AGAP002208-RA	0.57	0.38			Metabolism	cytochrome P450 family, monooxygenase activity, iron binding
AGAP008296-RA	1.32	0.34		TRY1_ANOGA	Metabolism	Peptidase S1A, chymotrypsin, trypsin activity, digestion, metabolism
AGAP007753-RA		3.13			Misc	Major facilitator family. Sugar transporter
AGAP007601		3.08			Misc	Major facililtator family, general substrate transporter
AGAP008052-RA		2.85		Q6H8Z3_ANOGA	Misc	Putative sensory appendage protein SAP-2 precursor
AGAP006076	1.44	2.80		OBP50	misc	oderant binding protein
AGAP002198-RA		2.73	1.96		Misc	Methyltransferase, gene/protein regulation?
AGAP010409-RA		2.43		OBP22	Misc	oderant binding protein
AGAP005837		2.37			Misc	Juvenile hormone esterase, d mel ortho
AGAP007918-RA		2.37		Q8I9N3_ANOGA	Misc	aldehyde oxidase/xanthine dehydrogenase
AGAP011426-RA		2.32			Misc	Sodium dependent phosphate transported (D.m and Aedes ortho's), part of MIP familoy of transporters
AGAP004433-RA	1.67	2.31		OBP19	Misc	oderantbinding protein19
AGAP002826-RA						
AGAP002827-RA		2.27			Misc	Major facilitator family, transporter
AGAP008404-RA		2.20	1.57		Misc	N-6 Adenine-specific DNA methylase
AGAP003733-RA	1.70	2.18	1.45		Misc	Cellular retinaldehyde binding/alpha-tocopherol transport
AGAP008182-RA	1.38	2.13			Misc	oderant binding protein
AGAP001966-RA		2.10			Misc	Sodium solute symporter
AGAP009464-RA	1.36	2.08	1.53		Misc	ABC transporter (import/exportof a wide variety of substrates)
AGAP005918	1.70	2.06			Misc	cation diffusion
AGAP006249		0.48			misc	transport of aa's
AGAP001514-RA		0.48			misc	Kv3.4 voltage-gated K+ channel
AGAP002622-RA	0.61	0.48			misc	Sodium/substrate sympoter - imports substrates against the concentration gradient using sodium gradients
AGAP012089-RA	0.72	0.48			Misc	negative regulation of protein amino acid dephosphorylation
AGAP004274-RA	0.56	0.47			misc	Zinc finger, CCHC type

AGAP000889-RA	0.68	0.46		misc	Actin-binding, cofilin/tropomyosin type
AGAP005174	0.81	0.45		Misc	Nucleoporin, WD-40 repeat containing
AGAP002284-RA		0.44		misc	Tetratricopeptide region, found in a wide variety of proteins, mediates protein-protein interactions/formation of multiunti complexes
AGAP011178-RA	0.61	0.44		Misc	Zinc/iron permease
AGAP011250-RA	0.63	0.43		Misc	calmodulin binding
AGAP003279-RA	0.60	0.43		misc	TPR (tetratrico peptide repeat region) protein-protein interactions, assembly of mutliprotein complexes
AGAP009507-RA	0.66	0.43		Misc	deubiquitinating-like enzyme
AGAP009659-RA		0.42		Misc	TPR (tetratrico peptide repeat region) protein-protein interactions, assembly of mutliprotein complexes
AGAP008136-RA	0.65	0.42		Misc	RNA binding, cyclophillin-like (protein folding?)
AGAP006925	0.64	0.42		misc	n-acetylgalactosaminyltransferase, oligosaccharide biosynthetic process
AGAP006949	0.85	0.41		misc	Metalloendopetidase activity
AGAP002387-RA	0.60	0.40		Misc	Histidine acid phosphatise
AGAP011742-RA		0.39		Misc	GCN5-related N-acetyltransferase
AGAP001423-RA	0.73	0.38		Misc	BIFUNCTIONAL PURINE BIOSYNTHESIS
AGAP002409-RA	0.55	0.36		Misc	pseudouridylate synthase activity, pseudouridine is a modefied nucleoside found in RNA, may offer protection from radiation
AGAP001522-RA	0.68	0.35		Misc	Chaperonin clpA/B, AAA ATPase - core, Peptidase S16, Lon protease, aedes ortho thyroid hormone receptor interactor
AGAP012005-RA		0.27		Misc	ABC transporter (import/exportof a wide variety of substrates)
AGAP010303-RA	1.45	2.23		protein degradation	ubiquitin activating enzyme, autophage associated (orthologs)
AGAP007721		2.57		protein transport/secretion	Synaptobrevin, membrane protein from vesicles, part of the SNARE complex involved in exocytosis
AGAP003576-RA		0.43		Protein transprot/secretion	Transportin, transport of protein into the nucleus
AGAP002836-RA		0.49	Q86MA9_ANOGA (Dicer-1)	RNA degradation	Dicer-1 (RNAi pathway, miRNA generation for gene expression regulation

AGAP011627-RA	0.71	0.45			RNA degradation	RNA metabolic process
AGAP005672	0.61	0.34		TSN	RNA degradation	Tudor-SN d.mel ortho, RNAi pathway
AGAP011204-RA	0.57	0.18			RNA degradation	PIWI, contains PIWI and PAZ domains, and PRO rich - d.mel ortho of PIWI, diverse roles but not classical siRNA
AGAP006941		0.44			RNA helicase	Probable ATP dependent RNA helicases
AGAP011145-RA	0.65	0.42			RNA helicase	DEAD box ATP dependent RNA helicases
AGAP012655-RA		0.42			RNA helicase	helicase?
AGAP011084-RA	0.51	0.40			RNA helicase	DEAD box ATP dependent RNA helicase
AGAP009600-RA	0.51	0.37			RNA helicase	ATP-dependent RNA/DNA helicase, DEAD box
AGAP003508-RA		0.34			RNA helicase	ATP-dependent RNA/DNA helicase, DEAD box
AGAP011322-RA		2.82			Signalling	Fibulin like - EGF-like calcium binding domains (ECM associated)
AGAP010766-RA		2.51			signalling	protein serine/threonine kinase activity, nucleotide binding, transferase activity, unknown function
AGAP012666-RA		2.20			Signalling	C-type lectin, sugar binding
AGAP005719	0.58	0.48	0.61		Signalling	Transcription initiation factor
AGAP002035-RA	0.65	0.47			Signalling	Putative transcription factor
AGAP007050	0.69	0.46			Signalling	GTP binding, porbable nucleolar GTP binding protein
AGAP005289		0.45			Signalling	Transforming growth factor ortholog, inhibitin beta unit
AGAP003121-RA	0.68	0.45			Signalling	PI-4 kinase/cadherin
AGAP001953-RA	0.71	0.42			Signalling	Rab GTPase activator activity
AGAP000627-RA	0.67	0.42			Signalling	Ser/thre or tyrosine kinase, ortho Casein kinase 1 isoform alpha (d.mel)(though to phos desheivilled in the WNT signalling pathway)
AGAP007118		0.42			Signalling	Tyrosine specific protein phosphatise
AGAP005362	0.64	0.39			Signalling	Nuclear Transcription factor, X-box binding
AGAP002902-RB	0.54	0.36		Q6PUC1_ANOGA (medea)	Signalling	Regulator of transcription in the dpp signalling pathway in drosophila (when bound to another Smad)
AGAP002123-RA	0.55	0.36			Signalling	negative regulator of the WNT pathway (D.mel Axin)

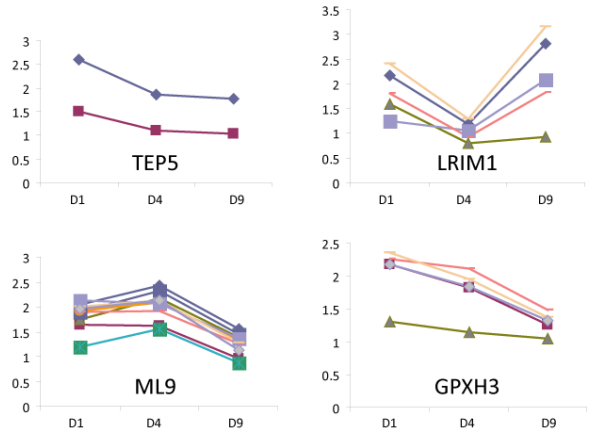
AGAP008149-RA	0.58	0.35		Signalling	Protein phosphatase 2C
AGAP010984-RA	0.65	0.33		Signalling	Transcription factor Tfb2
AGAP010118-RA		2.32		Transcription	Histone acetyltransferase, promotes transcription
AGAP001900-RA		0.49		Transcription	negative regulation of transcription
AGAP009024-RA	0.63	0.28		Transcription	RNA polymerase I specific transcription initiation factor RRN3
AGAP001879-RA	0.66	0.49		Translation	Regulation of alternative splicing? D.mel ortho crokked neck protein
AGAP005640	0.60	0.49		Translation	Pre mRNA spicing factor, RNA processing
AGAP007326	0.76	0.47		Translation	Part of the splicosome, regulation of mRNA maturation/splicing
AGAP000325-RA	0.84	0.46		Translation	Catalyss attachement of aa to tRNA during translation
AGAP002337-RA	0.68	0.43		Translation	Eukaryotic translation initiation factor 3, subunit 7
AGAP004336-RA	0.51	0.40		Translation	rRNA processing
AGAP012283-RA	0.64	0.34		Translation	Cysteinyl-tRNA synthetase, class Ia
AGAP006200		4.28		unknown	Unknown
AGAP003083-RA	1.31	2.98		unknown	Alpha-beta hydrolasefold-1 - common fold in hydrolytic enzymes
AGAP006506	1.97	2.82		unknown	Unknown
AGAP010365-RB		2.82		unknown	Unknown
AGAP004784	1.69	2.80		unknown	Male sterility domain, function unknown
AGAP007946-RA		2.64		unknown	Unknown
AGAP002582-RA	1.35	2.60		unknown	PA fragment
AGAP010111-RA		2.52		unknown	Unknown
AGAP002853-RA		2.46		unknown	Unknown
AGAP005614	1.90	2.34		unknown	unknown
AGAP012432-RA	1.37	2.30		unknown	membrane protein of unknown function
AGAP004549-RA		2.29		unknown	major royla jelly protein domain
AGAP010502-RA		2.29		unknown	unknown
AGAP010385-RA		2.27		unknown	unknown
AGAP003095-RA	1.92	2.26	Q8MZM5_ANOGA	unknown	major royla jelly/calcium binding EF-hand domains
AGAP008013-RA	1.67	2.25		unknown	unknown

AGAP005247		2.24		unknown	unknown
AGAP004695	1.48	2.24		unknown	unknown
AGAP006259	1.77	2.23		unknown	unknown
AGAP011784-RA	1.80	2.22		unknown	unknown
AGAP008301-RA		2.22		unknown	Cystine knot, c-terminal, growth factor? (common motif in several growth factors)
AGAP004674-RA	1.51	2.13		unknown	conotoxin
AGAP003636-RA		2.11		unknown	unknown
AGAP009145-RA	1.48	2.09		unknown	PA fragment
AGAP001989-RA		2.07		unknown	unknown
AGAP012436-RA		2.05		unknown	PA fragment
AGAP006275	1.54	2.04		unknown	unknown
AGAP005259	0.62	0.48		unknown	unknown, antifreeze domain
AGAP011333-RA		0.48		unknown	AMP-dependent synthetase and ligase
AGAP003086-RA	0.73	0.47		unknown	unknown, TM
AGAP002014-RA	0.65	0.46		unknown	unknown
AGAP003646-RA	0.62	0.46		unknown	unknown
AGAP005033	0.62	0.46		unknown	unknown
AGAP004468-RA	0.65	0.45		unknown	Phosducin-like, unknown function
AGAP012103-RA	0.57	0.45		unknown	RNA binding
AGAP009424-RA		0.44		unknown	tubby domain, unknown function
AGAP012281-RA		0.42		unknown	protein binding, function unknown
AGAP009307-RA		0.40		unknown	bromodomain, KV14 domain, functin unknwon
AGAP008879-RA		0.39		unknown	PA fragment, unknown function
AGAP004463-RA	0.56	0.34		unknown	unknown
AGAP000270-RA	0.63	0.32		Unknown	Tropomyosin (mucle contaction in muscle cells, unknown function in other cells), P60-like (protein found to interact with hereps virus proteins, funtion unknown)
	170.00				
Mid to late onset					
AGAP005693		5.88	2.76	Immunity	LRR protein
AGAP001768-RA		3.16	2.35	metabolism	Gamma interferon inducible lysosomal thiol reductase GILT domain

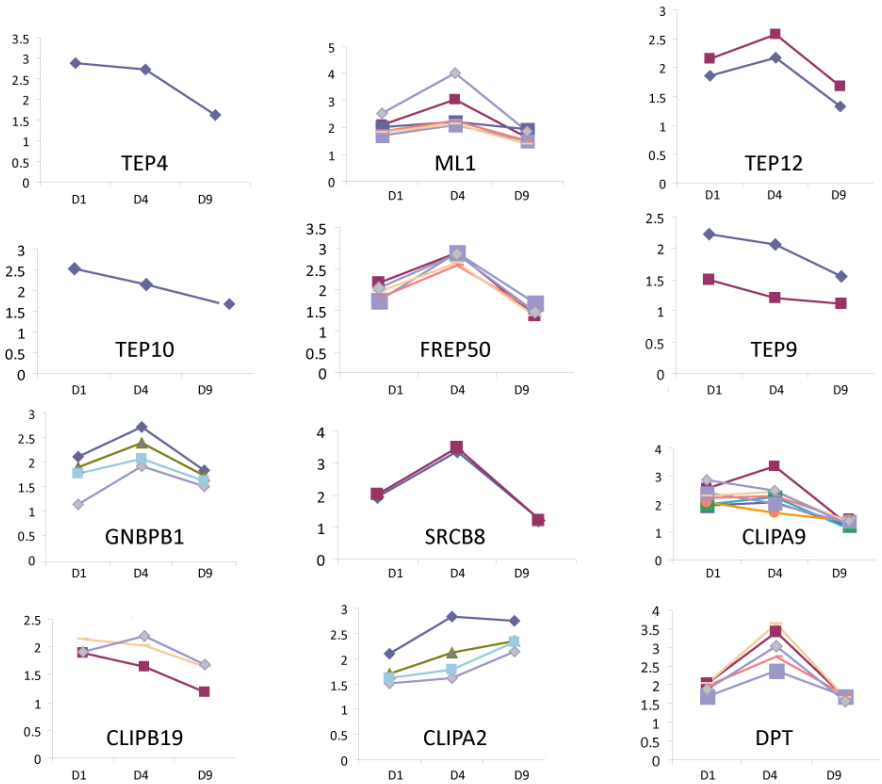
AGAP008141-RA		3.25	2.03		Misc	Fumerate lyase
AGAP003776-RA		8.33	2.81		unknown	unknown
	5.00					
Late onset						
AGAP003502	1.45		2.53	HPX6	Immunity	Peroxidase
AGAP012037			2.26	CLIPB20	Immunity	Clip domain serine protease
AGAP003246			2.18	CLIPB2	Immunity	Clip domain serine protease
AGAP004316	1.91		2.72		immunity?	CLIP
AGAP004880			3.43		metabolism	L-lactate dehydrogenase
AGAP012561	1.62		2.98		misc	iron binding, peroxidase
AGAP009551			2.26		misc	Sulfotransferase, transfers sulphate groups to specific compounds
AGAP005458			0.43		misc	5'nucleotidase, putative
AGAP008862			0.50		RNA degradation	Piwi and Paz domain containing (not d.mel Piwi)
AGAP007965			0.49		RNA degradation	tudor/maternal tudor domain containing (not d.mel Tudor-SN)
AGAP005079	1.43	1.66	2.20	AOEJE7_A NOGA	Signalling	G protein alpha subunit, q2
AGAP001610		1.83	2.44		unknown	unknown
AGAP004133			2.41		unknown	unknown
AGAP012604			2.41		unknown	unknown
AGAP001096	1.21	1.59	2.38		unknown	unknown
	16.00					
Early and late onset						
AGAP005848	2.05	1.85	2.59	FREP44	Immunity	Fibrinogen-like
AGAP010781	2.64		2.46		unknown	unknown
AGAP000183	2.45		2.07		unknown	unknown
	3.00					
Broadly responsive						
AGAP007385	2.03	2.54	2.07	LYSC4	Immunity	Lysozyme
	1.00					

Figure S1. Immunity probe set plots for all immunity related genes identified as being differentially regulated

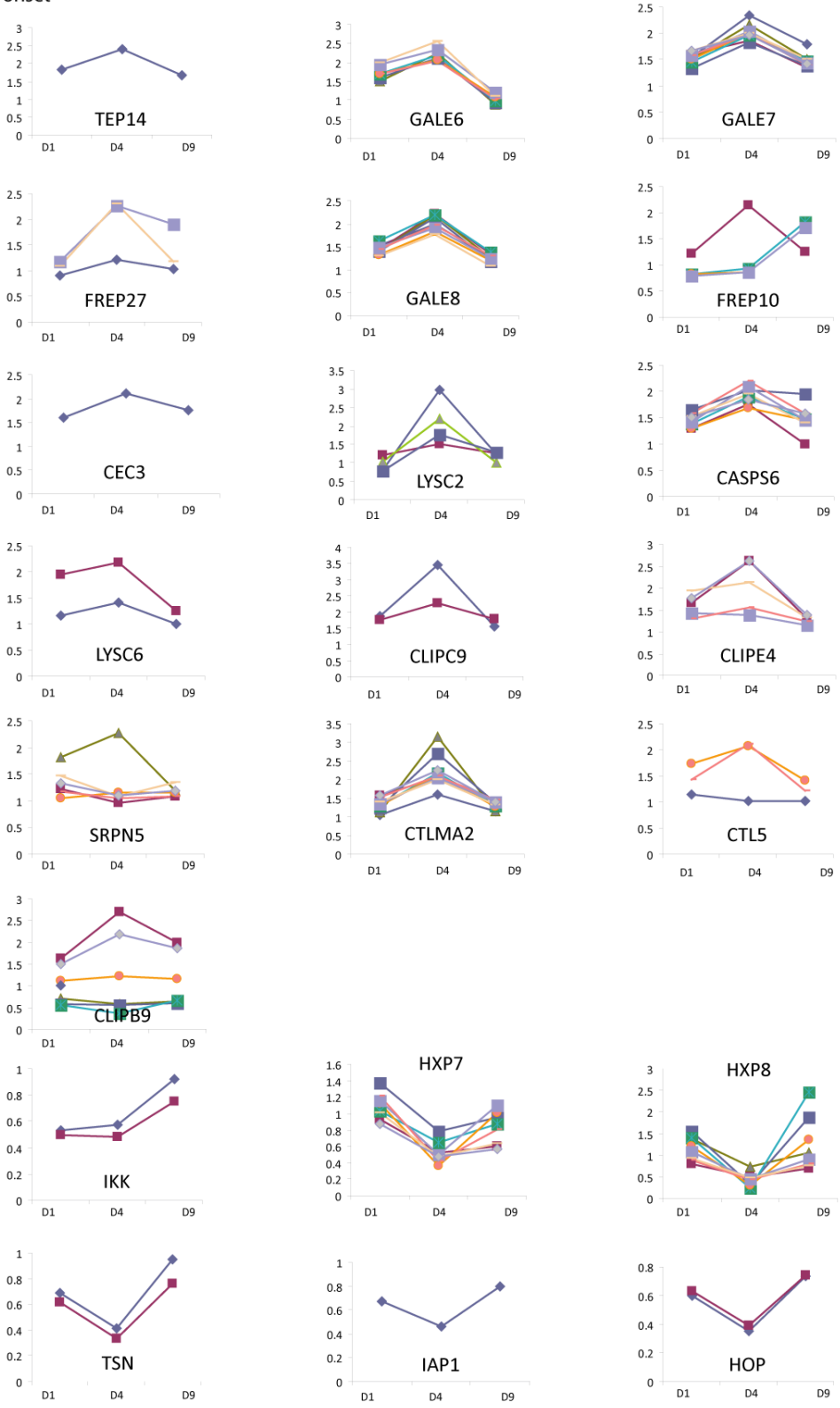
Early onset



Early-mid onset



Mid onset


TEP14
GALE6
GALE7
FREP27
GALE8
FREP10
CEC3
LYSC2
CASPS6
LYSC6
CLIPC9
CLIPE4
SRPN5
CTLMA2
CTL5
CLIPB9
HXP7
HXP8
IKK
TSN
IAP1
HOP
D1
D4
D9

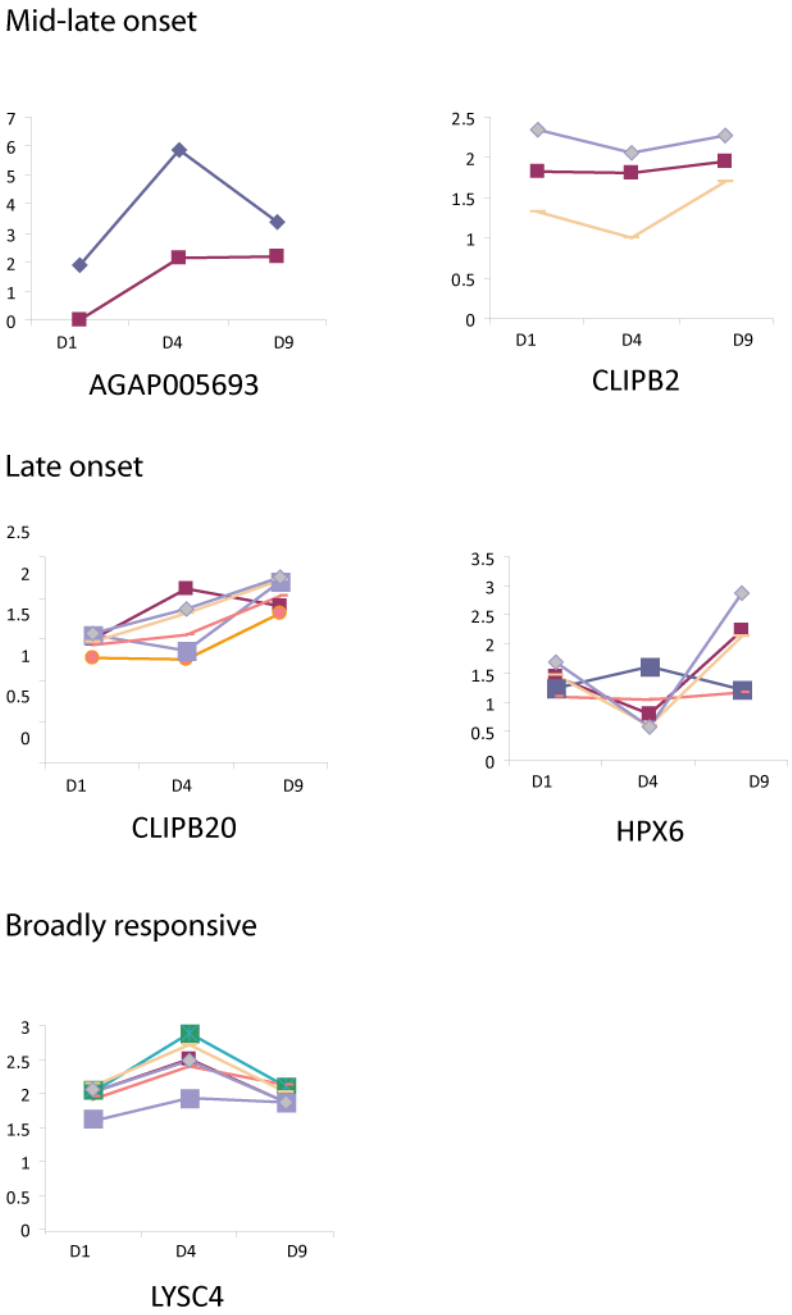


Figure S1. Immunity probe set plots for all immunity related genes identified as being differentially regulated. For all immunity genes where at least one immunity probe passed all the filtering criteria outlined in figure 20, the entire probe-set, including data with poor P-values, was plotted.

Appendix 2

Tables of genes differentially regulated by ds/ssRNA. Red text denotes a P value of >0.05. 'Lost' denotes no data for the probe due to low intensity or flagged as absent in GeneSpring. ssRNA/dsRNA denotes the fold change ratio for ss/dsRNA treated compared to a water control. Genes have grouped based on functional groups Immunity, Metabolism – general, Metabolism – Redox, Other, Unknown. Genes are ordered by fold change ratio within each functional group.

Table 1. Genes regulated by ds and ssRNA

AGAP	Name	ssRNA	dsRNA	Putative function
Immunity				
AGAP011780	CLIPA4	3.525	4.988	Serine protease homolog
AGAP005234	SOD2	3.289	4.7045	superoxide dismutase
AGAP004148	CLIPB5	2.741	4.61	Serine protease
AGAP012529	GALE8	3.866	4.57	galectin (sugar binding)
AGAP004855	CLIPB13	2.611	4.059	Serine protease
AGAP010731	CLIPA8	2.303	3.423	CLIP-SPH
AGAP007411	CTLMA1	2.165	2.631	sugar binding
AGAP011788	CLIPA14	2.176	2.149	Clip domain serine protease
AGAP002825	PPO1	0.377	0.496	prophenoloxidase
AGAP004917	FREP12	0.454	0.464	FREP
AGAP012539	FREP15	0.461	0.419	FREP
AGAP011224	FREP57	0.43	0.41633	FREP
AGAP011197	FREP13	0.384	0.41171	FREP
AGAP010775	FREP26	0.415	0.41133	FREP
AGAP010814	TEP5	0.493	0.406	TEP5
AGAP012616	PPO5	0.479	0.4025	prophenoloxidase
AGAP010819	TEP10	0.423	0.401	Thioester containing protein
AGAP004975	PPO3	0.445	0.398	prophenoloxidase
AGAP011225	FREP28	0.409	0.391	FREP
AGAP011231	FREP59	0.421	0.3898	FREP
AGAP000693	CEC1	0.428	0.38922	ceropin
AGAP011230	FREP58	0.4625	0.3889	FREP
AGAP007343	LYSC2	0.424	0.3873	lysozyme
AGAP011276	FREP29	0.407	0.3676	FREP
AGAP011226	FREP5/61	0.402	0.363	FREP
AGAP010763	FREP35/21	0.406	0.354	FREP
AGAP007346	LYSC5	0.42	0.35325	lysozyme

AGAP010815	TEP1	0.427	0.353	Thioester containing protein
AGAP000694	CEC3	0.398	0.34633	Cecropin
AGAP010830	TEP9	0.4095	0.345	Thioester containing protein
AGAP010773	FREP22	0.427	0.342	FREP
AGAP010762	FREP4	0.375	0.3346	FREP
AGAP009184	FREP7	0.396	0.3296	FREP
AGAP006743	FREP65	0.435	0.3265	FREP
AGAP007347	LYSC1	0.374	0.3005	Lysozymes
AGAP007041	FREP39	0.374	0.254	FREP
AGAP010760	FREP32	0.322	0.317	FREP
AGAP010811	FREP19	0.37	0.297	FREP
AGAP011239	FREP60	0.345	0.279	FREP
Metabolism - general				
AGAP010733		6.724	8.65	Vanin-like (nitrogen metabolism)
AGAP011750		3.53	4.187	n-acetylglucosaminidase
AGAP006121		3.063	3.704	carboxylic ester hydrolase
AGAP012400		2.231	3.372	amino acid transporter/Glycosyl hydrolase
AGAP012818		2.112	3.2	ATPase, V0/A0 complex, 116-kDa subunit
AGAP005846		3.144	3.124	Carbohydrate metabolic process
AGAP004954		2.215	2.622	Haloacid dehalogenase-like hydrolase
AGAP006821		2.014	2.238	Thiolase (acetyltransferase activity)
AGAP004399		2.089	2.074	phosphoglycerate mutase
AGAP008837		0.459	0.411	ubiquitin-dependent protein catabolic process
AGAP003350		0.429	0.373	Phosphoenolpyruvate carboxykinase (GTP)
Metabolism - redox				
AGAP006576		4.028	4.9445	Oxidoreductase activity
AGAP003408		3.309	4.627	Oxidoreductase activity
AGAP010255		2.785	3.666	Oxidoreductase activity
AGAP011051		3.088	3.6	Oxidoreductase activity
AGAP011325		2.424	3.567	Chymotrypsin

AGAP010400		2.576	3.478	Flavin-containing monooxygenase FMO
AGAP002866		2.693	3.035	Cytochrome P450 (oxidoreductase activity)
AGAP009197	GSTE3	2.442	3.007	Glutathione s transferase
AGAP002867	CYP6P4	3.322	2.737	Cytochrome P450 (oxidoreductase activity)
AGAP009225		2.001	2.372	Oxidoreductase
AGAP003030		2.052	2.103	dehydrogenase E1 - oxidoreductase activity
Other				
AGAP010361		6.581	10.79	Chitin binding
AGAP004785		4.394	7.105	Transport - carrier protein?
AGAP002810		3.618	4.975	Calcium binding - supercoiling factor?
AGAP000628		2.569	3.792	tudor - RNA binding/ssDNA binding?
AGAP012197		2.277	3.777	Histone H3
AGAP001823		2.166	2.692	Vacuolar (H+)-ATPase G subunit
AGAP000740		2.503	2.667	Histidine triad protein (HIT)
AGAP008403		2.106	2.303	serine-type endopeptidase activity
AGAP006669		2.879	2.094	multiple domains - DNA binding/atrophin/antifreeze/proline rich/eggshell protein/voltage gated K channel/ELM2
AGAP006948		0.393	0.5	P25-alpha tubulin polymerisation (?)
AGAP007237		0.484	0.462	Peroxidasin (with LRR repeats)
AGAP012029		0.495	0.45	WD-40 repeat containing protein
AGAP012067		0.463	0.435	calcium ion binding
AGAP001392		0.487	0.388	Enhanser of split-like
AGAP007349		0.435	0.348	Tropoelastin/annexin
AGAP000279	OBP8	0.461	0.337	oderant binding protein
Unknown				
AGAP003358		6.674	9.718	Unknown

AGAP008923	7.193	7.605	Unknown
AGAP003713	4.086	6.599	Twist protein family - unknown function
AGAP004787	2.568	5.295	Unknown
AGAP006207	2.879	4.814	Unknown
AGAP010266	2.615	4.207	Unknown
AGAP003764	3.683	3.997	Unknown
AGAP012536	2.167	3.606	Unknown
AGAP006398	3.279	3.507	Unknown
AGAP006449	2.413	3.467	Unknown
AGAP012127	2.801	3.372	Unknown
AGAP007021	2.394	3.326	Unknown
AGAP007641	2.164	3.251	unknown
AGAP008192	2.381	3.104	Unknown
AGAP004587	2.489	2.895	Unknown
AGAP011518	2.43	2.714	ATP binding cassette
AGAP007149	2.585	2.706	Unknown
AGAP003834	2.119	2.662	Unknown
AGAP000103	2.206	2.486	Unknown
AGAP000669	2.104	2.183	Unknown

Table 2. Genes regulated by dsRNA

AGAP	Name	dsRNA	ssRNA ratio	ssRNA p value	Putative function
Immunity					
AGAP006645		4.22	2.11006667	poor P value	LRR
AGAP011503		3.662	1.558	0.412	LRR
AGAP002857	ML5	2.629	1.489	0.094	MD-2-related lipid-recognition
AGAP009213	SRPN16	2.179	1.432	0.110	serine protease inhibitor
AGAP009221	SRPN5	2.127	1.384	0.258	serine protease inhibitor
AGAP007407	CTLMA4	2.117	1.69311111	good p value	sugar binding
AGAP001212	PGRPLB1	2.032	1.406	0.409	peptidoglycan recognition protein
AGAP006348	LRIM1	0.495	0.642	0.063	LRR
AGAP011294	DEF1	0.4945	0.71425	poor P value	Defensin
AGAP009212	SRPN6	0.45125	0.65211111	poor P value	inhibitory serine protease
AGAP010869	FREP30	0.44375	0.46216667	poor P value	Fibrinogen-like
AGAP004981	PPO4	0.441	lost	#N/A	prophenoloxidase
AGAP011974	SCRC1	0.4342	0.65988889	good p value	Scavenger receptor
AGAP010774	FREP27	0.415	0.84757143	poor P value	Fibrinogen-like
AGAP006258	PPO2	0.387	0.671	0.320	prophenoloxidase
AGAP011307	FREP3	0.206	0.318	0.066	Fibrinogen-like

Signalling				
AGAP010233	2.691	1.995	0.068	multiple domains - integrin beta subunit/EGF-like region/von Williebrand factor, Type A
AGAP007654	2.658	lost	#N/A	GTPase (Ras/Rab/Rho)
AGAP002119	2.336	1.948	0.058	Dual specificity tyrosine-phosphorylation-regulated kinase
AGAP008235	2.284	1.584	0.003	Zinc-finger, DNA binding
AGAP006355	2.281	1.642	0.020	transcription initiation factor IIE, alpha subunit
AGAP000107	2.063	1.473	0.155	ion transport/ankyrin domains
AGAP008708	0.474	0.592	0.051	tyrosine phosphatase (de-phosphorylation)
Metabolism - general				
AGAP000219	3.238	1.852	0.193	microtubule cytoskeletal organisation
AGAP010229	4.968	4.237	0.061	nitrogen metabolism - aliphatic nitrilase
AGAP006962	4.556	3.408	0.097	acid phosphatase activity
AGAP001600	3.61	2.643	0.065	5'nucleotidase / hydrolyase activity
AGAP003664	3.293	2.673	0.063	metalopeptidase activity
AGAP007028	2.943	2.377	0.071	transferase activity, transferring glycosyl/hexosyl groups
AGAP006670	2.421	2.168	0.062	gamma-glutamyl hydrolase
AGAP012399	2.283	1.728	0.212	carbohydrate metabolism, cation binding
AGAP008686	2.275	1.873	0.084	epoxide hydrolase activity/response to toxin/microsome/membrane
AGAP003688	2.226	1.905	0.005	Aromatic amino acid beta-eliminating lyase/threonine aldolase
AGAP010164	2.188	1.967	0.008	CARNITINE O PALMITOYLTRANSFERASE
AGAP011061	2.167	1.704	0.028	Phosphotriesterase-related protein
AGAP007590	2.124	1.841	0.003	carbohydrate kinase
AGAP002055	2.124	1.841	0.003	BETA GALACTOSIDASE PRECURSOR
AGAP011183	2.101	1.585	0.039	OXYSTEROL BINDING RELATED 11
AGAP003148	2.083	1.966	0.024	peptide-modifying enzyme component
AGAP005750	2.066	1.431	0.061	UDP-glucuronosyl/UDP-glucosyltransferase
AGAP006371	2.059	1.37	0.061	Glycoside hydrolase/Alpha-amylase
AGAP003730	2.035	1.458	0.099	Neutral/alkaline nonlysosomal ceramidase (phospholipid generation (?))

AGAP011859		0.49	0.626	0.012	Glycoside hydrolase, family 2
AGAP005066	Q6WLH5_ANOGA	0.484	0.566	0.018	Carbonic anhydrase (carbonate dehydratase activity)
AGAP012656		0.48	0.686	0.123	Carbonic anhydrase (carbonate dehydratase activity)
AGAP000162		0.469	0.607	0.035	cysteine synthase
AGAP011939	Q17032_ANOGA	0.454	0.644	0.065	Glycosyl hydrolase, family 13, catalytic region
AGAP008596		0.422	0.564	0.005	AMP-dependent synthetase and ligase
AGAP002198		0.421	0.548	0.064	Glycine N-methyltransferase
AGAP011569		0.41	0.517	0.032	Phospholipase_A2_met
AGAP011984		0.391	0.505	0.026	Glycosyl transferase, family 2
Metabolism - redox					
AGAP002865		3.667	2.246	0.082	Cytochrome P450 (oxidoreductase activity)
AGAP004383	GSTD10	2.275	1.873	0.084	Glutathione S-transferase
AGAP005948		2.171	1.871	0.015	kynurenine 3-monooxygenase
AGAP006132		2.006	1.431	0.061	Glutathione S-transferase, N-terminal
AGAP012291		0.47	0.606	0.010	Cytochrome P450, E-class, CYP3A
Other					
AGAP010861		2.918	1.975	0.009	neurotransmitter transport/sodium dependent aa transporter
AGAP008335		2.846	2.401	0.065	general substrate transporter
AGAP005653		2.496	1.561	0.191	amino acid transmembrane transporter
AGAP003295		2.352	1.942	0.070	DNAse II
AGAP008738		2.231	1.55	0.099	(from drosophila) RNA binding/inhibition of translation
AGAP003733		2.217	1.448	0.029	Cellular retinaldehyde binding/alpha-tocopherol transport
AGAP006934		2.196	1.98	0.002	Tropomyosin/voltage gated K channel domains
AGAP009381		2.162	1.857	0.017	Cellular retinaldehyde binding/alpha-tocopherol transport
AGAP008578		2.122	1.82	0.014	Vasa-like ATP dependent RNA helicases
AGAP005776	PDH	2.034	1.722	0.104	pigment dispersing hormone
AGAP007119		2.03	1.4	0.258	sideroflexin/ion carrier
AGAP000633		2.023	1.691	0.034	testis specific protein
AGAP010344		2.006	1.431	0.061	sulphate transporter
AGAP012107		2.052	1.495	0.013	unbiquitin conjugating enzyme

AGAP000793	2.044	2.331	0.065	E3 UBIQUITIN LIGASE SHPRH EC_6.3.2.- EC_3.6.-.- 1 SNF2 HISTONE LINKER PHD AND RING FINGER DOMAIN CONTAINING HELICASE
AGAP005064	2.012	1.419	0.057	protein family - E3 UBIQUITIN LIGASE EC_6.3.2.- RING FINGER
AGAP009182	0.489	0.513	0.071	Sec61 protein (Fragment)
AGAP006278	0.489	0.513	0.071	Insect pheromone/odorant binding protein PhBP
AGAP011870	0.481	0.618	0.018	Extracellular matirx metalloproteinase
AGAP003703	0.48	0.686	0.123	actin/microtubule binding
AGAP009791	0.478	0.594	0.026	metallopeptidase, neprilysin-like
AGAP002248	0.478	0.594	0.026	microtubule motor (kinesin-motor region containing), in Drosophila meiotic/meitotic spindle associated
AGAP010386	0.474	0.592	0.051	transmembrane, TRAM1-like, in drosophila phagocytosis/engulment
AGAP000376	0.467	0.577	0.019	Transferrin - Iron transport
AGAP001242	0.462	lost	#N/A	Immunoglobulin subtype 2
AGAP012391	0.445	lost	#N/A	Unknown
AGAP006718	0.407	0.512	0.005	sodium:dicarboxylate symporter activity
Unknown				
AGAP004988	3.54	2.553	0.055	Anti-freeze/unknown
AGAP009473	3.52	2.093	0.155	Unknown
AGAP007049	3.398	2.357	0.229	Defensin?
AGAP006067	3.233	2.351	0.087	unknown
AGAP006102	2.852	2.234	0.249	Unknown
AGAP007663	2.689	2.164	0.062	Unknown
AGAP010032	2.678	1.885	0.032	Unknown
AGAP007529	2.669	2.285	0.055	Unknown
AGAP001974	2.523	1.843	0.015	Unknown
AGAP010640	2.318	lost	#N/A	Anti-freeze/unknown
AGAP009146	2.245	1.61	0.003	Unknown
AGAP009145	2.156	1.376	0.461	PA fragment
AGAP007159	2.125	1.476	0.142	Unknown
AGAP008410	2.065	1.566	0.167	Unknown
AGAP005397	2.011	1.282	0.002	Unknown
AGAP010736	0.5	0.629	0.240	Unknown
AGAP000014	0.493	0.682	0.107	Unknown
AGAP005732	0.492	0.602	0.009	Unknown
AGAP012878	0.491	0.64	0.019	Unknown
AGAP010839	0.484	0.566	0.018	Unknown
AGAP000268	0.479	lost	#N/A	PA-fragment, unknown
AGAP002034	0.474	0.592	0.051	PA-fragment, unknown
AGAP010523	0.471	0.611	0.031	Unknown
AGAP004753	0.467	0.577	0.019	Unknown
AGAP007997	0.452	0.696	0.052	Unknown
AGAP012423	0.398	0.6	0.106	Unknown

Table 3. Genes regulated by ssRNA

AGAP	Name	ssRNA	dsRNA ratio	dsRNA p value	Putative function
Immunity					
AGAP006647		3.71	lost	#N/A	LRR
AGAP000572	CLIPC10	3.09	1.25	poor p value	Clip domain serine protease
AGAP004318	CLIPC3	2.99	lost	#N/A	Clip domain serine protease
AGAP012425		2.09	lost	#N/A	LRR
AGAP004918	Frep63	0.46	0.51	0.01290	Fibrinogen-like
AGAP009728	FREP42	0.44	lost	#N/A	Fibrinogen-like
AGAP012384		0.42	lost	#N/A	LRR
AGAP007758		0.37	lost	#N/A	LRR
AGAP007345	LYSC3	0.35	lost	#N/A	lysozyme
Metabolism - redox					
AGAP005645		3.76	lost	#N/A	Oxidoreductase
AGAP010250		3.03	lost	#N/A	Aldo/keto reductase
AGAP010399		2.68	lost	#N/A	oxidoreductase activity/FAD binding
AGAP004639		2.67	lost	#N/A	Peptidase S1A, chymotrypsin
AGAP009196	GSTE3/7 (probe maps to both genes)	2.20	lost	#N/A	Glutathione S-transferase
AGAP010966	Q5XNS9_ANOGA	0.45	lost	#N/A	Cytochrome P450
Other					
AGAP003168		2.14	1.99	0.04140	Isocitrate/isopropylmalate dehydrogenase
AGAP006400		2.12	lost	#N/A	alkaline phophatase
AGAP011333		2.05	1.99	0.07260	Disco-interacting protein/catalytic activity metabolism
AGAP004642	NPF	4.62	lost	#N/A	neuropeptide hormone activity
AGAP004281		3.30	lost	#N/A	dsRNA binding
AGAP010579		3.25	lost	#N/A	Tyrosyl-DNA phosphodiesterase/DNA repair
AGAP004018		2.75	lost	#N/A	protein binding
AGAP007753		2.65	lost	#N/A	sugar transporter integral to membrane
AGAP011026	Q9UB34_ANOGA	2.46	lost	#N/A	5'nucleotidase, putative (?)
AGAP012462	CPR147	2.31	1.85	0.01510	structural constituent of cuticle

AGAP007588	2.31	1.81	0.01330	glucosyl/glucuronosyl transferases
AGAP007645	2.25	lost	#N/A	protein binding
AGAP004169	0.50	lost	#N/A	Collagen triple helix repeat
AGAP001537	0.48	lost	#N/A	DEAD-like DNA (?) helicases
AGAP007631	0.46	lost	#N/A	adenylate cyclise
AGAP001657	0.30	lost	#N/A	hemocyanin oxygen transport?

Unknown

AGAP006093	5.09	lost	#N/A	unknown
AGAP001661	2.72	lost	#N/A	unknown
AGAP009955	2.70	lost	#N/A	unknown
AGAP008478	2.31	lost	#N/A	unknown
AGAP005492	2.30	lost	#N/A	unknown
AGAP007425	2.28	1.23	0.11700	unknown
AGAP002160	2.10	1.58	0.00233	unknown
AGAP005890	2.04	lost	#N/A	unknown
AGAP007365	2.03	lost	#N/A	unknown
AGAP010651	2.01	0.88	0.35500	unknown
AGAP007081	0.50	lost	#N/A	unknown
AGAP000944	0.50	lost	#N/A	PA fragment
AGAP010838	0.50	lost	#N/A	unknown
AGAP001104	0.47	0.50	0.01190	unknown
AGAP008878	0.44	lost	#N/A	unknown
AGAP011605	0.41	lost	#N/A	PA fragment