UNIVERSITY OF CRETE
DEPARTMENT OF BIOLOGY
DEPARTMENT OF MEDICINE

FORTH
INSTITUTE OF MOLECULAR BIOLOGY
AND BIOTECHNOLOGY

Postgraduate program:

«Molecular Biology and Biomedicine (MBB) »

"Characterization of Stitcher-Cadherin 96Ca and the transcription factor Grainyhead-Grh in *Anopheles gambiae*"

"Χαρακτηρισμός του Στίτσερ-Καντερίνη 96Ca και του μεταγραφικού παράγοντα Γκρέινιχεντ-Grh στον Ανωφελή κώνωπα"

Spanou Chara

Supervisor: Dr Siden-Kiamos Inga
Committee: Dr Siden-Kiamos Inga
Assoc. Prof. Tzamarias Dimitrios
Prof. Vontas Ioannis

Heraklion 2016

Author's Summary

Malaria is a deadly infectious disease caused by the parasitic protozoo *Plasmodium* that cost the lives of nearly 438,000 people in 2015 alone. Due to the current socioeconomic situation, an approach that addresses the problem more globally needs to be employed. To that end, much attention has received the study of malaria transmission through the mosquito host. New findings could help elucidate the mechanisms involved in mosquito host defense that could subsequently be utilized in downstream applications to limit disease propagation through the mosquitoes. When an *Anopheline* mosquito ingests infected blood, the parasite is taken up and then fertilization takes place in the mosquito midgut where the resulting zygote proceeds to the next developmental stage, the ookinete. The ookinete traverses the midgut epithelium by penetrating 1-2 midgut cells in the process that then undergo apoptosis. In that way, the parasite wounds the midgut epithelium. Recently, a wound-healing mechanism has been dissected when *Drosophila melanogaster* embryos are subject to aseptic injury. In this signaling cascade, Stitcher, the receptor, receives the stress signals that are then transduced to intermediate kinases that activate Grainyhead to drive the transcription of barrier repair genes. In this work, it was investigated whether the homologous proteins in *Anopheles gambiae* midgut signal for a similar response during and right after ookinete traversal.

Περίληψη

Η ελονοσία είναι μία μολυσματική θανατηφόρα ασθένεια που προκαλείται από το παρασιτικό πρωτόζωο πλασμώδιο και στοίχισε τη ζωή περίπου 438.000 ανθρώπων μόνο το 2015. Λόγω της υπάρχουσας κοινωνικοοικονομικής κατάστασης, μία μέθοδος που αντιμετωπίζει το πρόβλημα πιο γενικευμένα θα πρέπει να χρησιμοποιηθεί. Για αυτό το λόγο, αρκετή έρευνα έχει γίνει για την μετάδοση της ελονοσίας στο κουνούπι. Νέα ευρήματα θα μπορούσαν να βοηθήσουν να αποσαφηνιστούν οι μηχανισμοί που εμπλέκονται στην άμυνα του κουνουπιού-ξενιστή με σκοπό να χρησιμοποιηθούν συνεπακόλουθα σε εφαρμογές που θα περιορίσουν την μετάδοση της ασθένειας από τα κουνούπια. Όταν ο Ανωφελής κώνωπας πίνει μολυσμένο αίμα, τα παράσιτα εισέρχονται στο κουνούπι και η γονιμοποίηση λαμβάνει χώρα στο μεσέντερο του κουνουπιού με τον ζυγώτη που προκύπτει να περνάει στο επόμενο αναπτυξιακό στάδιο, τον οοκινέτη. Ο οοκινέτης διεισδύει στο επιθήλιο του μεσεντέρου διαπερνώντας 1-2 επιθηλιακά κύτταρα τα οποία έπειτα αποπίπτουν. Με αυτόν τον τρόπο, το παράσιτο πληγώνει το επιθήλιο του μεσεντέρου. Πρόσφατα, ένας μοριακός μηχανισμός που επουλώνει τραύματα αποσαφηνίστηκε έπειτα από πρόκληση τραυμάτων σε έμβρυα Δροσόφιλας. Σε αυτόν τον καταρράκτη σήματος, το Στίτσερ, ο υποδοχέας, λαμβάνει σήματα στρες τα οποία μετάγονται σε ενδιάμεσες κινάσες που ενεργοποιούν το Γκρέινιχεντ ώστε να ξεκινήσει την μεταγραφή των γονιδίων που σχετίζονται με την επούλωση. Σε αυτήν την εργασία, ερευνήθηκε αν οι ομόλογες πρωτεΐνες στο μεσέντερο του Ανωφελούς κώνωπα σηματοδοτούν μία παρόμοια απόκριση κατά τη διάρκεια της διείσδυσης του οοκινέτη αλλά και αμέσως μετά από αυτή.

Characterization of Stitcher and Grainyhead in *Anopheles gambiae*

INTRODUCTION

Malaria is an infectious disease caused by the parasitic protozoon *Plasmodium* that is transmitted to humans by a single mosquito bite. In the time period 2000-2015, 57 countries have achieved a 75% reduction in malaria cases and mortality rate has been reduced by 60% globally. The greatest malaria burden is still on African countries (Fig. 1) however there too a significant progress has been made since 66% less people died of malaria among all age groups and 71% less people died under the age of five. However, there still 3.2 billion people that are at risk. More specifically, there were 214 million new cases resulting in 438,000 deaths. 80% of these new cases were all coming from 15 countries in Africa with 35% coming only from Nigeria and the Democratic Republic of Congo. Methods for malaria treatment include targeting the parasite whereas malaria control relies largely on targeting the mosquito host and much less on the generation of new vaccines. Current methods for treatment provide only modest protection as the parasite has already developed resistance to the drug artemisin in 5 countries (Cambodia, Lao People's Democratic Republic, Myanmar, Thailand and Vietnam) (WHO 2015 malaria report). Resistance to drugs can be attributed to the antigenic variation region that changes the proteins in the parasite surface thus rendering pharmacological interventions inefficient (Peter *et al.*, 2015; Su, 1995). For the same reason, many attempts to generate a vaccine against malaria have failed. The only exception is the newly developed vaccine "Mosquirix" that successfully passed phase III of clinical trials and was suggested for clinical use in young children. However, this vaccine is only 50% effective and its effectiveness declines with age (Morrison, 2015). The greatest problem for the people at risk arises from the fact that they do not have access in health facilities for prophylaxis (nets dipped in mosquito pesticides, diagnostic tools) and treatment (artemisin drug). Another consideration that should be taken into account is the fact that mosquitoes, the vectors for malaria transmission, show resistance to the pesticides applied on mosquito nets thus adding a further complexity to the problem (WHO 2015 malaria report). However, it should be stressed that targeting mosquitoes remains the most cost-effective alternative since this could prevent malaria transmission without accessibility to healthcare services for diagnosis and treatment being a prerequisite. This practically means that a new pesticide against mosquitoes that could be sprayed on malaria prevalent regions or the production and release of transgenic mosquitoes (mosquitoes that have been genetically manipulated to block transmission) could reduce disease occurrence for all people at risk regardless of their ability to access healthcare services and with less money compared to pharmacological interventions and the generation of new vaccines (Guillet *et al.*, 2001; Killeen *et al.*, 2007). A new study shows promising results for transgenic mosquitoes expressing human antibodies against the parasite without compromising mosquito viability (Isaacs *et al.*, 2012).

The major mosquito vector of malaria transmission to humans is the adult female *Anopheles gambie*. The male mosquitoes are fed exclusively on nectar whereas the females are fed on both nectar and blood. Blood is important for adult females since they need it for egg-laying. The nectar goes to the anterior midgut whereas the blood goes mainly to the midgut. The midgut is a single layer epithelium that consists of the basal lamina that faces the hemocoel and the midgut epithelium cells that are attached on the basal lamina with their basal part. These cells also have microvilli on their lateral part that face the midgut lumen (Fig. 2). After a meal, the mosquito midgut excretes gut proteases for digestion and the nutrients are absorbed with the aid of the midgut microvilli. The midgut also expands rapidly since the volume of the meal is large, to return back to its normal size 3 days after feeding (Vlachou *et al.*, 2004; Kotsifakis, 2004).

When a mosquito ingests blood infected with the parasite, the parasite is taken up in the form of gametocytes, which are blood cells containing either the female or male form of the parasite. In the mosquito midgut, the sexual stages begin as the appropriate conditions, pH=8 and xanthurenic acid activate these gametocytes so that fertilization can take place. More specifically, the appropriate

conditions in the mosquito midgut trigger gametocyte egress from the host cell. The parasite then divides to produce gametes. The male gamete releases flagella that attract female gametes so that fertilization can take place in the midgut. The two gametes form the zygote which later gives rise to an elongated cell with gliding motility, called the ookinete. The ookinete penetrates first the peritrophic matrix by secreting chitinases to digest it and then traverses the midgut epithelium by gliding between or through 1-2 midgut cells. These undergo apoptosis and are expelled towards the midgut lumen by neighboring healthy cells. Lamillipodial protrusions of the neighboring cells seal the generated wound inflicted on the midgut by the parasite in order to maintain the epithelium barrier. Then the parasite attaches to the basal lamina and progresses to the next developmental stage, a round cell called the oocyst (Fig. 3). The ookinete traversal and oocyst formation trigger mosquito innate immune responses in the midgut that kill the parasites either by cell lysis or melanization (Agrisano *et al*., 2012; Voltz *et al*., 2006). Melanization is a highly efficient defense response in invertebrates that is linked with wound-healing and entrapment of pathogens in a dense melanin coat produced by crystal cells (Bier and Guichard, 2012). As a result of the previous, the parasitic population undergoes a major bottleneck and only few parasites survive (Blandin *et al.*, 2004). The oocysts of the surviving parasites grow in a time period of approximately 13 days by gaining volume and size. In this life stage, the parasitic DNA replicates and the oocyst undergoes many mitotic cell divisions yielding individual nuclei that together with newly synthesized organelles give rise to many parasites, called sporozoites. Later, the oocyst ruptures and the sporozoites are released into the mosquito hemocoel. The sporozoites then travel through the hemocoel and traverse the salivary gland epithelium to result in the mosquito saliva ready to be transmitted back to the mammal host with the next mosquito bite. In the mammal host, the parasite completes a second life cycle undergoing two different sets of developmental stages; these are the exo-erythrocytic liver stages and the asexual erythrocytic stages (Agrisano *et al*., 2012; Banister and Sherman, 2009; Kotsifakis, 2004).

Recently, a wound-healing mechanism triggered by the proteins Stitcher (Stit), a receptor and Grainyhead (Grh), a transcription factor was identified after aseptic injury of *Drosophila melanogaster* embryos (Tsarouhas *et al.*, 2014). It is suggested that there is a close link between wound-healing and innate immunity in insects since a wound-opening can be the source of bacterial or parasitic infection (Davis and Engstrom, 2012; Pare *et al*., 2012). Since the protein Grainyhead (Grh) is known to mediate innate immune responses and wound-healing, an immediate question that arises is whether Grh and its immediate target, Stit can be involved in a wound-healing mechanism inducing midgut epithelium barrier repair after infection (septic wounding) with the rodent malaria parasite *P.berghei*.

Grh belongs to the family of Elf-1 transcription factors that have a common ETS DNA binding domain with the typical helix-turn-helix structure that recognizes a GGA(A/T) DNA sequence (Sharrocks *et al.*, 1997). *Grh* gene (flybase:: grainy head, FBtr0299705) is located on chromosome 2R in *D. melanogaster* and has eight alternative transcript variants (grh-RH, grh-RI, grh-RJ, grh-RK, grh-RL, grh-RM, grh-RN, grh-RO, grh-RP). Grh protein has an activation domain closer to the N-terminus. Closer to the C-terminus, Grh has a DNA binding and a dimerization domain. Through this dimerization domain, it is thought that Grh molecules form dimers to stabilize their interaction with DNA (Uv *et al.*, 1994).

Grh is conserved from insects to mammals and participates in early embryonic development, neural proliferation and cell fate determination, epidermal epithelium development, innate immunity, tubular size control and wound-healing in *D.melanogaster* regulating these processes as an activator or a repressor (Hemphala *et al*., 2003; Wang and Samakovlis, 2012). Grh deletion during development is lethal (Wang and Samakovlis, 2012) and results in fragile cuticle for *D. melanogaster* a phenotype resembling *C. elegans* Grh null mutants (Venkatesan *et al*., 2003). The previous suggests conserved functions for Grh in cuticle development between diptera and the nematode.

In the context of aseptic wound-healing in *D.melanogaster*, Grh acts as the downstream effector and enters the nucleus to drive the transcription of barrier repair genes *msn* (misshapen), *ddc* (dopa decarboxylase) and *ple* (pale) to achieve effective wound closure (Mace *et al*., 2005; Pearson *et*

al., 2009; Wang and Samakovlis, 2012). DDC and TH (TH:: Tyrosine Hydroxylase encoded by the *ple* gene) are important for hardening and coloration of the cuticle; TH catalyzes the conversion of tyrosine to DOPA and DDC catalyzes the conversion of DOPA to dopamine that after additional reaction steps yields melanin (Tang, 2009; True *et al*., 1999; Wang and Samakovlis, 2012). *Msn* encodes for a kinase that participates in a signaling cascade related with dorsal closure in the *Drosophila* embryo (flybase:: Msn).

In the context of septic wounding with the gram-negative bacteria *Erwinia carotovora carotovora 15* (Ecc15) or the gram-positive bacteria *Micrococcus luteus* in *D.melanogaster*, Grh RNAi mediated knock down compromised adult fly survival, implicating therefore a protective role for Grh after septic wounding. Furthermore, according to microarray experiments, abrogated Grainyhead signaling in stages 16-17 of *D.melanogaster* embryogenesis resulted in misregulation of 3 major gene clusters implicating therefore functional redundancy: 1) Innate immunity 2) Defense and Stress response 3) Detoxification whereas Grh downregulation led to downregulation of genes involved in cuticle melanization and wound-healing including Stitcher and *ddc* which are thought to be direct Grh targets. Among others, the expression of Peptidoglycan Recognition Protein LC (PGRP-LC) gene, a receptor of the Imd pathway was down in these experiments (Pare *et al*., 2012). To enhance the previous, Grh is suggested to mediate innate immune responses by activating the expression of PGRP-LC gene, which is required for the production of antimicrobial peptides (AMPs) after infection in *D. melanogaster* (Takehana *et al*., 2004; Wang *et al*., 2009). Furthermore, wound-healing per se leads to Grh dependent transcriptional activation of the genes *ddc*, *msn* and *ple* that promote melanization as previously described. *A.gambiae* mosquitoes also use melanization reaction to kill or dispose of dead parasites (Voltz *et al*., 2006).

Grh has also been detected in *A.gambiae* 6.5 h pregastrulating embryo using in situ-hybridization assays. It was also observed that *A. gambiae* embryos obtain resistance to desiccation between 8-14 h after egg laying implicating therefore that *grh* could be one of the genes responsible for this phenomenon. Resistance to desiccation allows *Anopheles* mosquitoes to survive in arid environments and transmit vector borne diseases such as malaria (Goltsev *et al*., 2009).

The *grh* gene (vectorbase:: *AGAP005564*) is localized on 2L chromosome in *A. gambiae* and is 165.77 kb long. Transcription initiates from the reverse strand giving rise to four annotated alternative transcript variants, the *grh* transcripts RA, RB, RC and RD. RA *grh* transcript consists of 12 exons and is 3412bp long, RB *grh* transcript consists of 12 exons and is 1960 bp long, RC *grh* transcript consists of 4 exons and is 644 bp long and RD *grh* transcript consists of 13 exons and is 2257 bp long. *Grh* exons are interrupted by large intronic regions that account for the extensive length of the gene. The exons among these transcripts differ significantly near the 5' end with the first exon of the RA transcript being unique. Closer to the 3'end, the exons differ only slightly with those indicated in the blue box being identical. All annotated *grh* transcripts are presented in Fig.4 as they appear in Vectorbase: https://www.vectorbase.org/Anopheles_gambiae/Gene/Summary?g=AGAP005564;r=2L:17101402-17247167 from 5' to 3'.

Drosophila Stit (flybase:: Cad96Ca-RA, FBtr0084874) is a Ret family Tyrosine Kinase (RTK) receptor. Its N-terminal half is extracellular and possesses structural features typical of cadherin proteins. The C-terminus is cytoplasmic and possesses structural features of a Protein Tyrosine Kinase (PTKc); it has the activation loop, an ATP binding site and a polypeptide substrate binding site. During signal transduction, Stit gets autophosphorylated and then phosporylates downstream mediators on serine or threonine residues (Wang, 2010; Tsarouhas *et al.*, 2014). The protein has a transmembrane domain that could dock Stitcher in the plasma membrane or intracellular vesicles. From a functional perspective, Stit participates in the formation of cuticle serosa, epidermis and wound-healing. Stit deletion results in pupae fatality in *D. melanogaster* (Wang and Samakovlis, 2012).

Stit gene (vectorbase:: *AGAP011648*) is located on chromosome 3L in *A. gambiae* and is 22.3 kb long. Transcription initiates from the forward strand giving rise to one annotated *stit* transcript. The total length of the annotated *stit* transcript is 2171 bp (Fig.5).

In the context of wound-healing, Stit and Grh signaling has been studied extensively when *D. melanogaster* embryos are subject to epidermal aseptic injury. In this signaling pathway, Stit receives the stress signal from the outer-cell environment that are then mediated and enhanced via kinases that ultimately communicate with the downstream effector, Grh, which in turn regulates gene expression to respond to wounding (Tsarouhas *et al.*, 2014; Sopko and Perrimon, 2013; Wang *et al.*, 2009). Stit activation signals is hypothesized to come from various sources; either growth factors are released at the site of injury (Martin, 1997) or during epithelium rupture, tissue cells become exposed to such factors or Ca^{2+} circulating through the hemolyph that can both act as Stit ligands (Woolley and Martin, 2000). A third alternative suggests that since disturbance of an epithelial layer can generate electric signals, this could possibly activate Stit in a similar manner that activates phosphatidylinositol-3-OH kinase-γ (PI(3)Kγ) and phosphatise and tensin homolog (PTEN) that regulate wound-healing (Zhao *et al*., 2006). Another hypothesis suggests that since this protein localizes in the plasma membrane as well, it may form trans homodimers between neighbouring cells in its inactive state (Tepass *et al*., 2000). Upon wounding however, if the connection between neighbouring cells is lost, Stit could form cis homodimers along the membrane of the same cell thus triggering the wound-healing mechanism by autophosphorylation. Lastly, Stit may not need a ligand for activation at all; wounding may relocalize Stit thus increasing its local concentration either on plasma membrane or in intracellular vesicles thus triggering the wound-healing mechanism in the absence of a ligand (Wang, 2010).

The Stit-Grh interplay in aseptic wound-healing can be described as follows; upon activation, Stit interacts physically with the downstream Drk kinase and Src kinases that mediate the response independently (Tsarouhas *et al.*, 2014). The Drk kinase triggers Erk phosphorylation that in turn leads to Grh transcriptional activation. The Src kinases apart from inducing transcriptional activation by Grh, they further facilitate wound closure by 2 other means; 1) they induce the formation of a contractile actin ring that surrounds the wound opening and 2) they promote re-epithelization. Furthermore, Stit enhances its own transcription whereas Grh activation establishes a positive feedback loop by direct interaction with the Stit regulatory domain composed of Grainyhead binding elements (Gbes) that further upregulate Stit. As a result, its transcripts accumulate at wound sites and the response is further enhanced to achieve efficient wound-closure without the need of extra stimuli (Fig. 4) (Wang and Samakovlis, 2012; Tsarouhas *et al*., 2014). In addition, Grh in the nucleus also activates the wound-healing reporters *msn*, *ddc* and *ple* in synergy with the transcription factor AP-1 (Mace *et al*., 2005; Pearson *et al*., 2009; Wang *et al*., 2009) (Fig. 6).

Since Grh and Stit are known to mediate innate immune responses and wound-healing of the epithelia, in this study, it was investigated whether the two proteins participate in a wound-healing response triggered in the mosquito midgut during and after ookinete traversal.

Figure 1

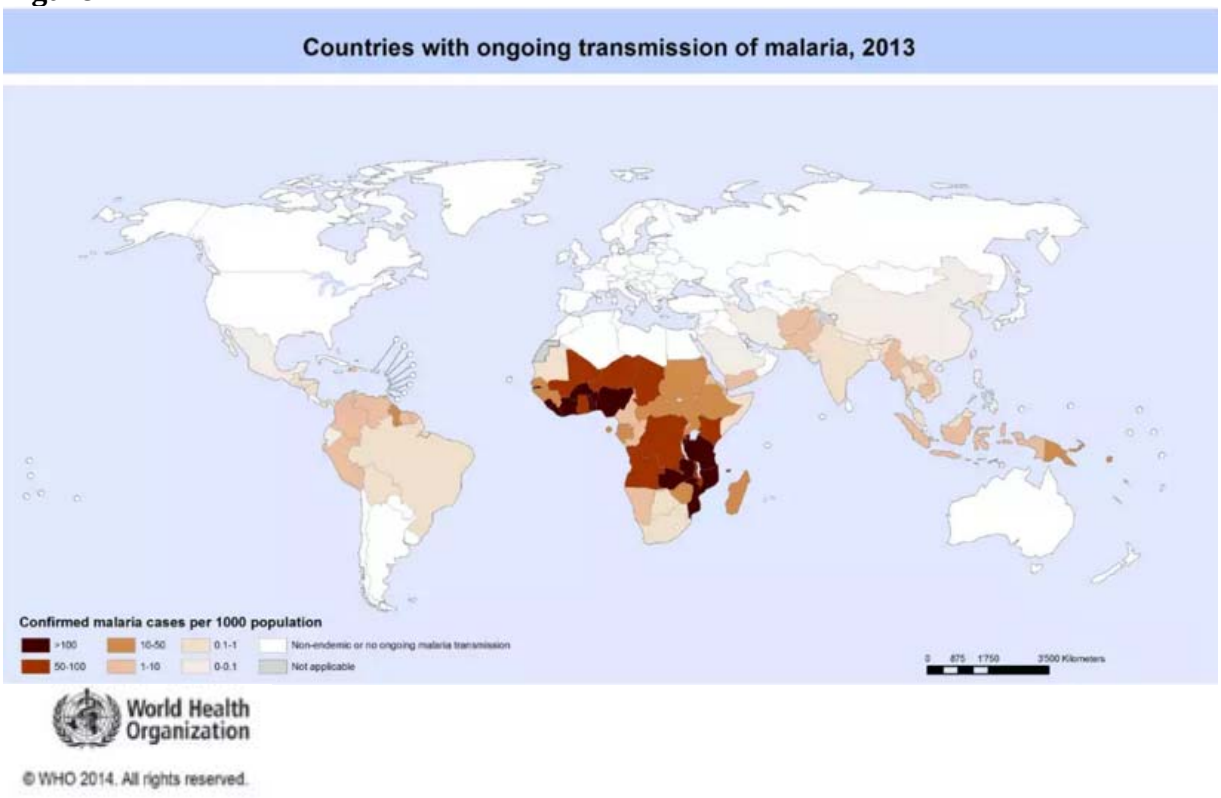


Fig.1: Global map that shows malaria cases for all countries in 2013. The cases of malaria have been reduced in Latin America and Asia. However, for most of the countries in Africa, malaria is still prevalent with over 50 cases per 1000 individuals.

Figure 2

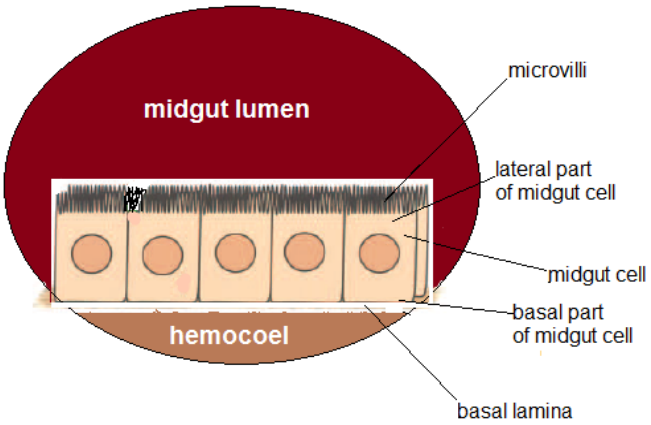


Fig.2: The mosquito midgut epithelium. The midgut cells are attached on the basal lamina with their basal part facing the hemocoel. The midgut cell microvilli on the lateral part, face the midgut lumen.

Figure 3

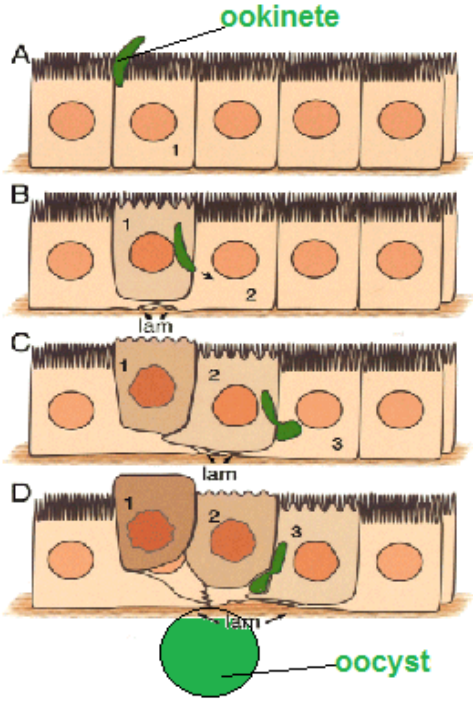


Fig.3: Schematic model describing parasite traversal and oocyst formation through the single-layer mosquito midgut epithelium; initially, the parasite infects midgut epithelium cells that then undergo apoptosis and are expelled towards the midgut lumen. Lamelipodial protrusions by the nearby uninfected cells are generated in order to seal the wound and maintain the integrity of the epithelium. Eventually, the parasite attaches itself to the basal lamina of the midgut epithelium and further maturates to the oocyst 24 hrs post mosquito feeding on infected blood.

Figure 4

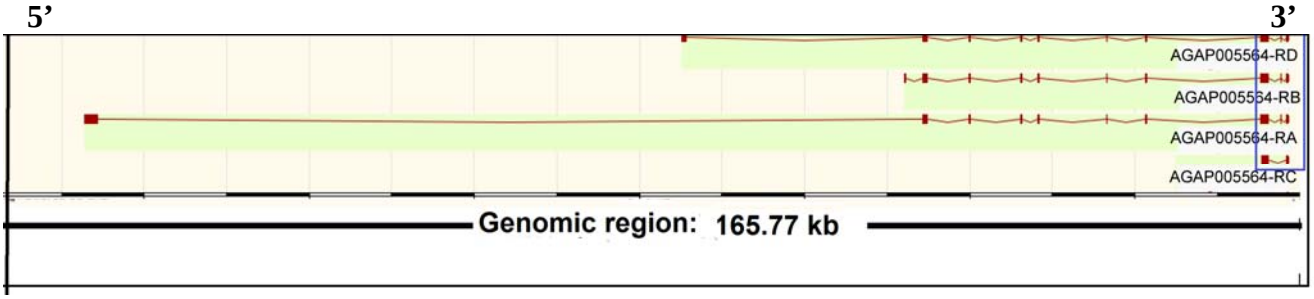


Fig.4: Schematic representation of all Grh transcripts as they appear in vectorbase for *A.gambiae*. The exons are interrupted by large intronic regions that further add to the length of Grh gene (165.77 kb). Shown in blue is the common region of all 4 Grh transcripts. Transcription initiates from the reverse strand of 2L chromosome.

Figure 5

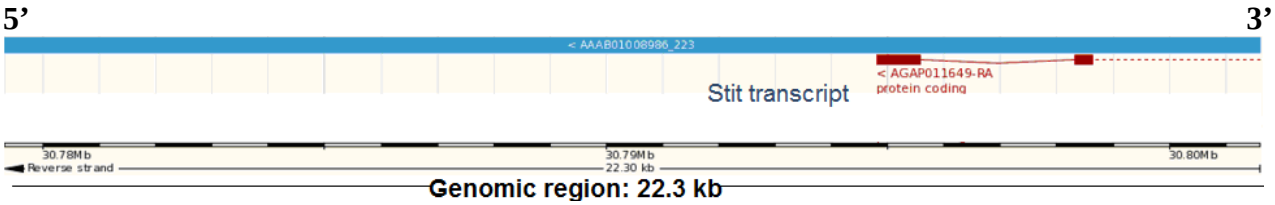


Fig.5: Schematic representation of the Stit transcript as it appears in vectorbase for *A.gambiae*. Stit gene is 22.3 kb long. Transcription initiates from the forward strand of 3L chromosome.

Figure 6

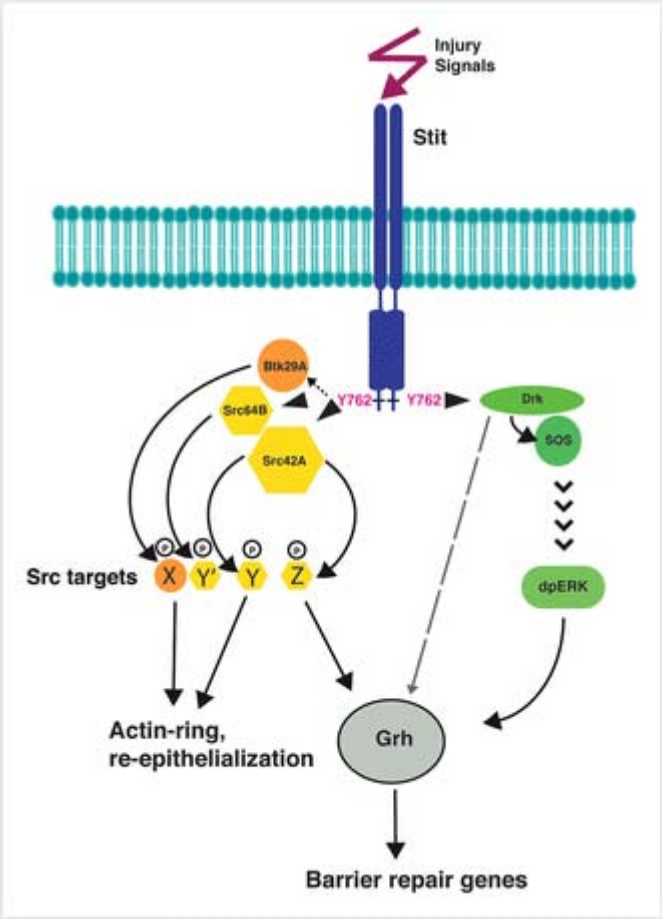


Fig.6: Schematic representation of the Stit-Grh interplay in aseptic wound-healing in *D.melanogaster*. Stit receives the stress signals and activates the kinase mediators Src and Drk that signal wound-closure by triggering cytoskeleton rearrangements and melanization through Grh.

MATERIALS AND METHODS

Sample preparation for Western blotting

Anopheles gambiae protein extracts:
15 mosquitoes were placed on 200 μl extraction buffer (0.3M NaCl, 100 mM Tris-HCl pH=7-7.5, 1% Triton, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM EDTA, 1 mM PMSF) supplemented with a general purpose protease inhibitor cocktail (SIGMA P2714) on ice. The insects were squeezed with a sterile pestle and sonication followed for 40 secs, 2-3 amplitude, 7 times while the samples were on ice at all times. The samples were then centrifuged for 10 mins, 10,141 xg at 4^{o}C. The supernatant was transferred to a new tube and the pellet was resuspended in 200 μl extraction buffer. The samples were analyzed on SDS-PAGE followed by transfer to nitrocellulose membrane filter and processing for Western blot.

Midgut protein extracts:
The midguts from 5-6 mosquitoes were dissected and placed in 20 μl PBS supplemented with 1 mM PMSF on ice and were subsequently stored at -80 ^{o}C until loaded on SDS-PAGE to be processed for Western blot.
D. melanogaster protein extracts:
35 female Oregon R flies and 35 male Oregon R flies were washed twice with PBS and protein extracted as described above. Triton-X was not always added to the extraction buffer and the protocol was modified as follows; after squeezing the insects, the samples were centrifuged for 5 min, 600 xg at RT. The supernatant was then sonicated and centrifuged at 15,800 xg for 30 min at 4 ^{o}C. The pellet was resuspended in extraction buffer without Triton-X. Samples were then loaded on SDS-PAGE and analyzed with Western blot. Grh antibody was used at dilution 1:1000 or at dilution 1:5000 to avoid the background. Integrin antibody was used at dilution 1:1000.

Immunofluorescence

The antibodies used for Stit and Grh were directed against the respective *Drosophila* proteins and were provided by Professor Samakovlis. Female midguts were dissected in PBS, fixated in 4% PFA, washed for 15 min 2x in PBS and blocked in the blocking solution (PBS, 0.2% Saponin, 0.5% BSA) for 1 h. Then the guts were incubated with the primary antibodies (1:800 for Grh, 1:500 for Stit diluted in the blocking solution), O/N at 4^{o}C shaking. To stain the plasma membrane, cadherin antibody (cad) was diluted 1:500 in the same blocking solution. The midguts were washed for 15 min 3 times and were subsequently incubated with the secondary antibodies diluted in blocking solution (anti-rabbit for Grh, anti-guinea pig for Stit) and washed as described previously. The midguts were then incubated with RNAse for 30 min on ice with occasional light shaking incubated for 10 min in TOPRO stain diluted 1:5000 in PBS, washed for 15 min 3 times in PBS and mounted in Vectashield.

Wounded midguts:
In order to elicit the response in conditions of aseptic wounding, uninfected female midguts were pricked with an insulin needle and incubated for 30 min in Ringer solution (150 mM NaCl, 3.4 mM KCl, 1.7 mM $CaCl_2$, 1.8 mM $NaHCO_3$, 1 mM $MgSO_4$, 5 mM sucrose and 25 mM Hepes pH=7.1).

Confocal microscopy:
Samples were viewed in a Biorad confocal. The same settings were used for both experiment and control midgut samples. Image J (http://rsb.info.nih.gov/ij/download.html) was used for image processing.

RT-PCR

RNA extraction
RNA was extracted from 25 infected or 65 uninfected female midguts using Trizol, obtained from Sigma, according to manufacturer's instructions. RNA was purified using the RNA Clean-up MinElute kit *cat.no*.74204, (Qiagen).

cDNA synthesis
cDNA was synthesized using Thermoscript Reverse Transcriptase (Invitrogen) with the oligodT method according to manufacturer's suggestions.

Primers for detecting the grh transcripts

Grh
RA transcript: 5'-CGAGGGTAACAAAACGGTCA-3' (RAF-forward)
RD transcript: 5'- GCATCAGCAAATTCATCAGC-3' (RDF-forward)
RC transcript/Common region: 5'- GTCGATCGGTCGGAGTTTTA-3' (RCF-forward)

For all the *grh* transcripts a common reverse primer (RBR) was used:
5'- GTCATCAATTCTCGCCGTAA-3'

Grh RA (RAF/RBR PCR fragment), RD and RC transcripts were detected with Go-Taq (Promega, *cat.no*.M3001) on larval cDNA.
Grh RA (RAN/RAP and RAF/RBR PCR fragments), RD and *stit* transcripts were detected with PCR using the Phusion enzyme (NEB) in GC buffer supplemented with 3% DMSO according to manufacturer's instructions (*cat.no*. M0530S).

Primers Grh	PCR conditions
RAF/RBR (Go-Taq)*	98 ^{o}C Initial denaturation for 2 mins
	98 ^{o}C denaturation for 30 secs
	50 ^{o}C annealing for 30 secs
	72 ^{o}C extension for 2 mins, 35 cycles
	72 ^{o}C Final extension for 10 mins
	4 ^{o}C indefinitely
RDF/RBR (Go-Taq)*	95 ^{o}C Initial denaturation for 2 mins
	95 ^{o}C denaturation for 30 secs
	50 ^{o}C annealing for 30 secs
	72 ^{o}C extension for 2 mins, 35 cycles
	72 ^{o}C Final extension for 10 mins
	4 ^{o}C indefinitely
RCF/RBR (Go-Taq)*	95 ^{o}C Initial denaturation for 2 mins
	95 ^{o}C denaturation for 30 secs
	50 ^{o}C annealing for 30 secs
	72 ^{o}C extension for 35 secs, 35 cycles

	72 °C Final extension for 10 mins
	4 °C indefinitely
RAN/RAP (Phusion)*	98 °C Initial denaturation for 30 secs
	98 °C denaturation for 10 secs
	61.5 °C annealing for 15 secs
	72 °C extension 1.16 mins, 40 cycles
	72 °C extension for 10 mins
	4 °C indefinitely
RAF/RBR (Phusion)*	98 °C Initial denaturation for 30 secs
	98 °C denaturation for 15 secs
	50 °C annealing for 15 secs
	72 °C extension for 3 mins, 40 cycles
	72 °C extension for 10 mins
	4 °C indefinitely
RDF/RBR (Phusion)*	98 °C Initial denaturation for 30 secs
	98 °C denaturation for 10 secs
	51 °C annealing for 15 secs
	72 °C extension for 1.16 mins, 40 cycles
	72 °C extension for 10 mins
	4 °C indefinitely

Primers for detecting the stit transcript

5'-GTAGATTCTTCCCCCACGAC-3' (forward-STF)
5'-CCAGTCCGTTCACAGCTTCT-3' (reverse-STR)

PCR conditions Stit (phusion)*

98 °C Initial denaturation for 30 secs
98 ° C Denaturation for 10 secs
54 °C annealing for 15 secs
72 °C extension 1.16 mins for 40 cycles
72 °C Final extension for 10 mins
4 °C indefinitely

Cloning

The PCR products of *grh* RA transcript and *stit* were purified and cloned in the pGEMem-T-easy vector (Promega). Clones were verified by sequencing.

RNAi

PCR product of total length 468 bp was prepared from the 2 kb *grh* RA transcript cloned in t-vector pGEMem-T-easy. The PCR product was flanked with the T7 promoter sequence added to the 5' end of the primers;
Primers for the transcript:
5'- GAATTAATACGACTCACTATAGGGAGAATGTCTGCATCGCCTGAAAT-3' (forward primer-FGK)
5'- GAATTAATACGACTCACTATAGGGAGAGCTGCTGCTGCCCACCGTGT-3' (reverse primer-RGK)
The resulting DNA amplicon was purified using the PCR purification kit (Qiagen) and eluted in DEPC-treated ddH_20. Purified DNA template for GFP was provided by Vontas lab. Both DNA templates (*grh*, GFP) were used for in vitro transcription using T7 Megascript kit (Ambion) according to manufacture's instructions at 20 µl (experiment 1) or 30 µl (experiment 2) final volume. The reaction was incubated at 37 ^oC for 24 h and dsRNA (442 bp or 650 bp for *grh* or GFP respectively) was purified using phenol/chloroform. The samples were precipitated with ethanol and resuspended in injection buffer (0.2 mM Na_2HPO_4, 10 mM KCl calibrated to pH~7 with NaH_2PO_4. In experiment 1, 1.5 day old mosquitoes were injected with Grh or GFP dsRNA (1000 ng/µl). Silencing was assessed with immunofluorescence 5 days post injection. For silencing experiment 2, 3 days old mosquitoes were injected with dsRNA (3000 ng/µl) and silencing was assessed 3 days post injection.

PCR conditions for Grh DNA template for in vitro transcription (phusion)*

98 ^oC Initial denaturation for 30 secs
98 ^{o} C denaturation for 15 secs
72 ^oC annealing for 15 secs
72 ^oC extension for 10 secs for 35 cycles
72 ^oC final extension for 10 mins
4 ^oC indefinitely

*Note: For all reactions indicated with an asterisk (*) a mix of DNA, primers and ddH_2O was preheated at 85 ^{o} C for 5 mins prior PCR reaction.

Parasites and mosquito methods

Mosquito infections were carried out using the rodent malaria parasite *Plasmodium berghei* ANKA $PbGFP_{CON,}$ which expressed GFP under a constitutive promoter in all life stages (Ref Franke-Fayard *et al*, 2004). Two-month old TO/Ola mice were infected by intraperitoneal injection. Parasitemia was monitored on Giemsa stained blood smears. For the exflagellation test, a blood sample from the tail of the infected mouse was taken and incubated in RPMI, pH=8.0 at 18 ^oC for 10-12 mins. To transmit the parasite to the mosquito host, the mice were anesthetized and offered to mosquitoes for 30-45 min. *Anopheles gambiae* NGousso strain mosquitoes were kept at 37 ^oC and fed on 10% sugar.

RESULTS

ANALYSIS OF GRH AND STIT PROTEINS

Alignment of Grh and Stit in *Drosophila* and *Anopheles*

D.melanogaster Grh protein has an activation domain near the N-terminus and a DNA binding and a dimerization domain near the C-terminus. Of the eight Grh isoforms produced from *grh* gene in *D.melanogaster*, grh-RI (flybase:: grh-RI, FBtr0299705) was used for Grh antibody production provided by Samakovlis lab. Hence, the specific protein isoform was aligned with the Grh isoforms in *A.gambiae* (vectorbase:: AGAP005564) using ClustalW. The first half of *D.melanogaster* grh-RI protein isoform including the activation domain is not very conserved for *A.gambiae*. The second half of this protein isoform closer to C-terminus includes the antibody binding region, the DNA binding domain and the dimerization domain and shares 80% similarity with *A.gambiae* Grh isoforms. The similarity for the dimerization domain in particular is over 90% for the two species. Only the alignment of the most similar *A.gambiae* Grh isoforms (PA and PD) to *D.melanogaster* grh-RI is shown in Supplemental Fig. 1 and is summarized in Fig. 7.

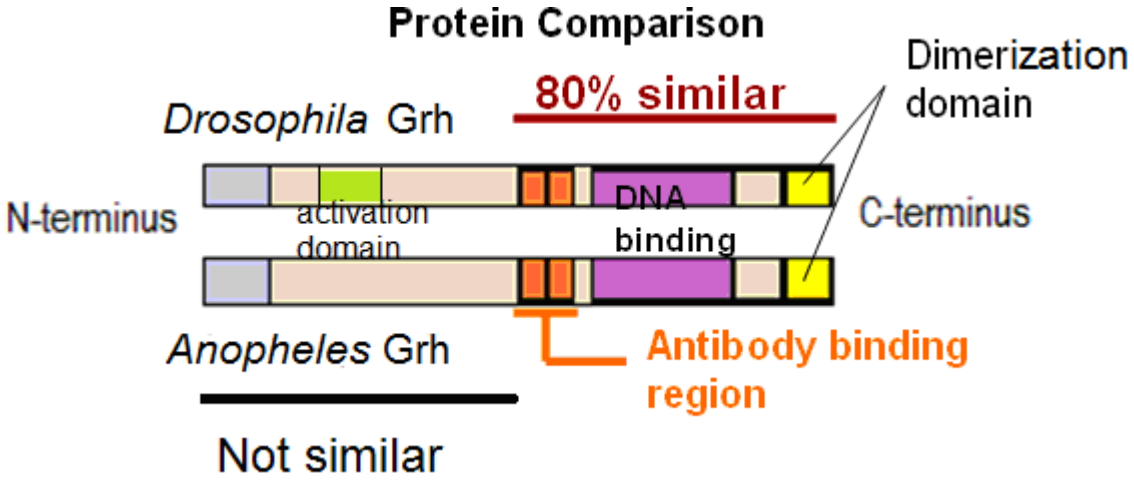


Fig.7: Alignment of *D. melanogaster* grh-RI with *A. gambie* Grh PA isoform and *A. gambie* Grh PD isoform. The antibody binding region (orange), the DNA binding domain (violet) and the dimerization domain (yellow) are indicated with 80% similarity between the two species. On the contrary, the activation domain (green) closer to the N-terminus of the *D.melanogaster* grh-RI protein isoform is not conserved for *A.gambiae* Grh.

Protein blast of the peptide sequence used for the Grh antibody production shows high scores for *A.gambiae* Grh isoforms PA and PD implicating therefore that the antibody could recognize both of these Grh isoforms (Fig.8A). The alignment of the isoforms PA and PD against the peptide sequence used for Grh antibody production is shown in Fig.8B and Fig.8C. *A.gambiae* PB isoform also came out with a high score, however since PB protein isoform is not that similar to *D.melanogaster* grh-RI, this was not investigated further.

Figure 8A: Protein blast of the sequence used for the production of the antibody recognizing *D.melanogaster* Grh.

Hit	Desc	Query	Aln Length	E-value	Score	Identity	Query Hit	DB Sequence Hit
AGAP005564-PD	transcription factor CP2 and r		59	2e-04	116	41.9%		
AGAP005564-PB	transcription factor CP2 and r		59	2e-04	116	41.9%		
AGAP005564-PA	transcription factor CP2 and r		39	3e-04	115	47.6%		
AGAP003539-PA	Armadillo repeat-containing pr		29	0.22	82	44.8%		
AGAP011138-PA	myosin XVIII |protein_coding|3		35	0.97	76	48.6%		
AGAP011337-PA	|protein_coding|3L:21374065:21		68	1.1	75	31%		
AGAP003282-PB	histone-lysine N-methyltransfe		24	1.7	74	45.8%		
AGAP003282-PA	histone-lysine N-methyltransfe		24	1.6	74	45.8%		
AGAP001901-PA	DNA helicase MCM9 |protein_cod		30	2.7	71	33.3%		
AGAP001255-PA	|protein_coding|2R:2046301:205		41	3.8	70	34.1%		

Fig.8A: The protein sequence used for Grh antibody production was blasted against all *Anopheles* protein databases. For *A.gambiae*, the 3 hits detected with a significant score correspond to the Grainyhead (AGAP005564) protein isoforms PD, PB and PA.

Figure 8B: Sequence alignment for the isoform PD:

```
Query 18  PKNGNIAGAATANGPGSVITQKSFDYTELCQ-PGTLIDAN--GSIPVSVNSIQQRTAVHG 74
          P +G  A A T     +V  Q+ FDY E CQ P  ++D N  G IP  V S+Q+  A++G
Sbjct 216 PDSGIGADAITPRDQNNV--QQQFDYAEPCQAPIGMVDPNAAGHIPACVASLQRNLAING 273


Query 75  SQ                                                        76
          SQ
Sbjct 274 SQ                                                        275

Score = 116
Expect = 2e-04
Identity = 41.9355%
Strand = Plus/Plus
Aln Length = 59
```

Fig.8B: Protein sequence alignment of the peptide sequence used for Grh antibody production and the PD Grh isoform, *A.gambiae*.

Figure 8C: Sequence alignment for the isoform PA:

```
Query 38  QKSFDYTELCQ-PGTLIDAN--GSIPVSVNSIQQRTAVHGSQ 76
          Q+ FDY E CQ P  ++D N  G IP  V S+Q+  A++GSQ
Sbjct 618 QQQFDYAEPCQAPIGMVDPNAAGHIPACVASLQRNLAINGSQ 659

Score = 115
Expect = 3e-04
Identity = 47.619%
Strand = Plus/Plus
Aln Length = 39
```

Fig.8C: Protein sequence alignment of the peptide sequence used for Grh antibody production and the PA Grh isoform, *A.gambiae*.

D.melanogaster Stit (flybase:: Cad96Ca-RA, FBtr0084874) has an extracellular domain closer to N-terminus that includes Cadherin (Cad) repeats and a cytoplasmic domain closer to the C-terminus that includes domains indicative of a kinase such as: 1) Protein Tyrosine Kinase domain (PTKc) responsible for kinase activity, 2) the ATP binding site, 3) the polypeptide substrate binding site and 4) the activation loop. Between the extracellular and the cytoplasmic domains, there is a predicted transmembrane domain. *D.melanogaster* extracellular domain that was used for Stit antibody production provided by Samakovlis lab, shares 56% similarity with *A.gambiae* Stit whereas the cytoplasmic part is highly conserved between the two species sharing 94% similarity. The complete alignment is shown in Supplemental Fig.2 and the schematic summary in Fig.9.

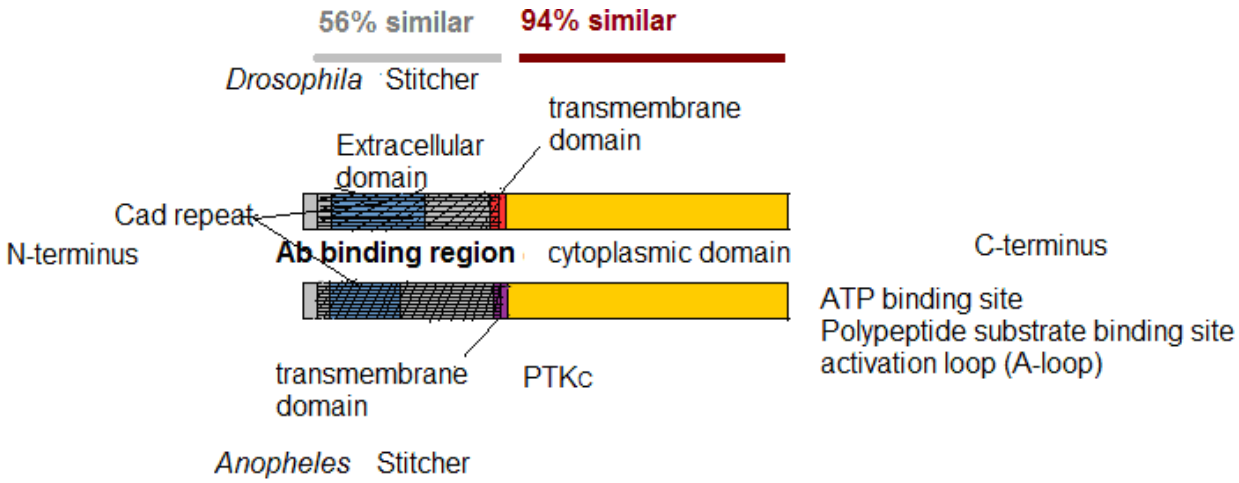


Fig.9: Schematic representation of Stitcher comparison between *D.melanogaster* and *A.gambiae*. The cadherin repeats (blue and black) and the cytoplasmic domain (yellow) are highly conserved between the two species. The cytoplasmic domain includes the following conserved structural features; the PTKc , the ATP binding site, the polypeptide substrate binding site and the activation loop. The grinded area represents the antibody binding region that shares 56% similarity between the two species and corresponds to Stit extracellular domain. The cytoplasmic domain shares 94% similarity between the two species.

Western blot of Grh and Stit

In order to detect Grh and Stit in adult *A. gambiae*, whole mosquito protein extracts and midgut protein extracts were analyzed by Western blot. Two bands were detected at the expected size for Grh isoforms PA, PD and PB (PA isoform: 124 kDa, PD isoform: 83 kDa, Fig.10, PB isoform: 71 kDa, Fig.11A) Isoform PC was not detected. Due to the high background these experiments were not conclusive, but indicate the presence of these two isoforms in females and in midguts derived from females, while only PD was detected in the male sample. No bands were detected for Stit at the expected molecular size (data not shown).
As a control, protein extracts were prepared from *D. melanogaster*. Several bands were detected between the 80-175 kDa molecular size range consistent with the expected molecular sizes for *Drosophila* Grh protein isoforms. Using the Stit antibody, no bands at the expected molecular size were detected (data not shown). An antibody detecting mosquito integrin was used as control, which gave a signal at the expected molecular size, (Fig.11B, only the integrin Western blot for the filter used for Grh detection is shown), thus excluding the possibility that the protein extracts had undergone proteolytic processing giving rise to the many bands.

Figure 10

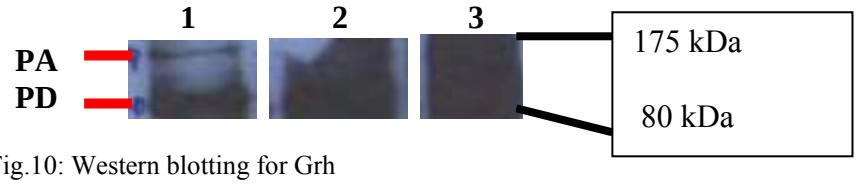


Fig.10: Western blotting for Grh
Lane 1: female supernatant, Lane 2: male supernatant Lane 3: female midguts.

Figure 11A

Figure 11B

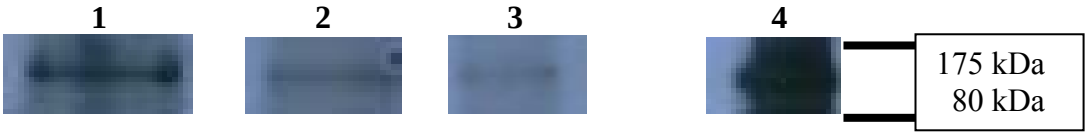


Fig.11A: Western blot with for Grh antibody
Lane 1-3: Samples derived from *D. melanogaster*. Lane 1. Female sample extracted with Triton X-100., Lane 2. Female samples prepared without detergent. Lane 3. Male sample prepared without detergent. Many bands are detected ranging from 80-175 kDa consistent with the expected molecular sizes of the Grh protein isoforms. Lane 4: female mosquito sample extracted with Triton X-100. One band was detected of the expected size for Grh protein isoforms PB (71kDa)
Fig.11B: Western blot using an antibody against mosquito integrin.
Integrin was detected at the expected molecular size (100 kDa) in all samples shown in Fig.11A.

GRH AND STIT TRANSCRIPTS

Detection of *grh* transcripts in 1st instar larvae and adult mosquito midgut

As described in Introduction the *grh* gene consists of a number of exons and four different transcripts have been annotated. In order to determine the identity of the transcript(s) in larvae and adult RT-PCR (reverse transcription PCR) was carried out. Specific primer pairs were designed to identify each annotated transcript. Forward primers were designed binding to specific exons of each transcript (RAF and RAN for RA, RDF for RD, RBF for RB transcripts). The same reverse primer (RBR) was used to detect all Grh transcripts.

Figure 12: The transcripts as shown in vectorbase

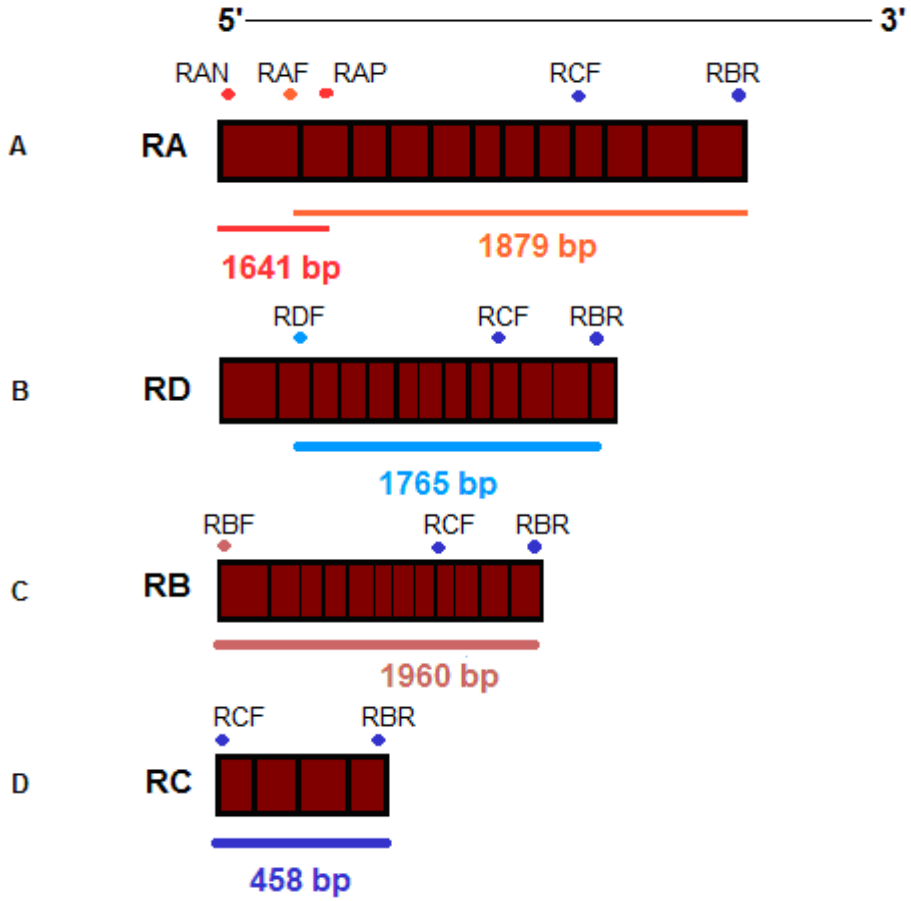


Figure 13: The transcripts as detected with RT-PCR on 1st instar larva and uninfected female mosquito midguts

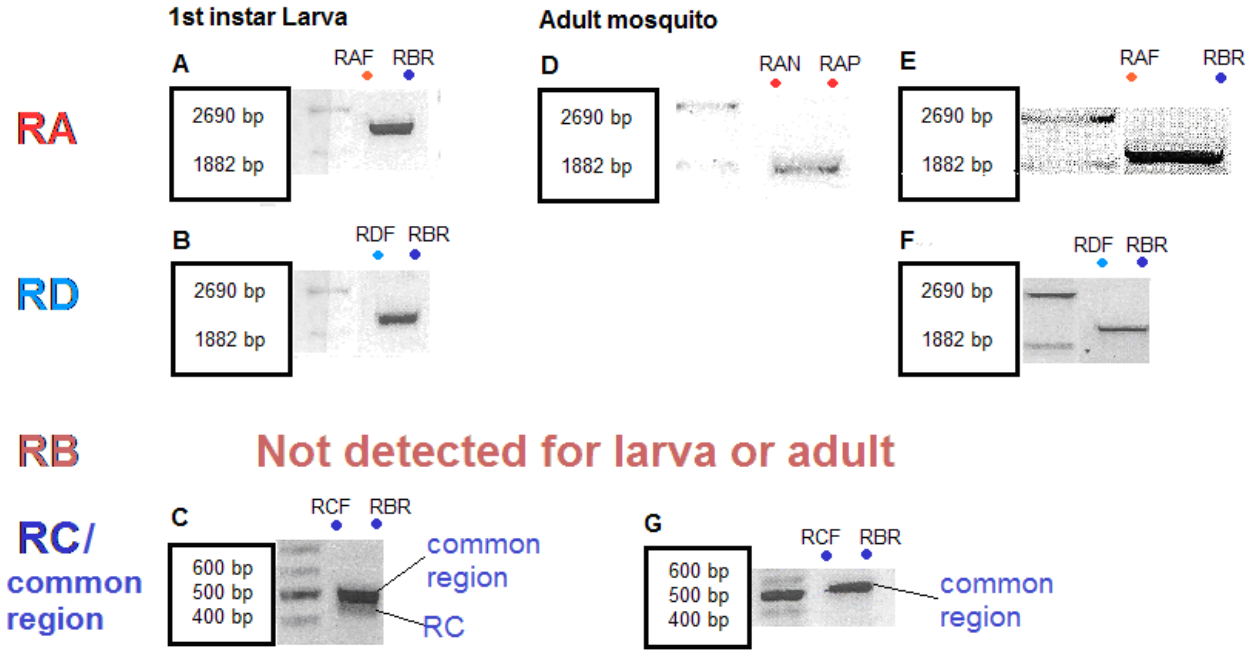


Fig.12A: The primer pairs RAF/RBR and RAN/RAP were used to detect the indicated regions of the *grh* RA transcript. The RAF/RBR PCR fragment includes the sequence encoding the antibody binding region for Grh antibody (exon 2). Primer pair RCF/RBR detects the region near the 3' of RA that is common among all Grh transcripts. **Fig.12B:** The primer pair RDF/RBR detects the indicated region of RD. **Fig.12C:** The primer pair RBF/RBR detects the indicated region of RB, **Fig.12D**: The primer pair RCF/RBR also detects the indicated region of RC and a region that is common between RA, RD and RB transcripts. The brown boxes for all transcripts show individual exons.
Fig.13G: The common region of all *grh* transcripts was detected for the adult mosquito at the expected molecular size (520 bp) using the primer pair RCF/RBR. The RC transcript was not detected for the adult mosquito. **Fig.13C:** The common region and the RC transcript were detected at the expected molecular sizes (450 bp for RC) using the same primer pair as in Fig.10G in 1st instar larva. **Fig13F:** The RD transcript was detected using the primer pair RDF/RBR at a molecular size different from the expected (expected: 1765 bp, detected: ~2kb)**. Fig.13B:** The same primer pair as in Fig.10F was used to detect RD in the larva. The molecular size of the transcript detected was consistent with the molecular size of the transcript detected in the adult mosquito. **Fig.13D:** The primer pair RAN/RAP was used to detect the region of the *grh* RA transcript indicated. The detected region significantly differed from the expected (detected: 1890 bp, expected: 1641 bp). **Fig.13E:** The primer pair RAF/RBR was used to detect the indicated region of *grh* RA at the expected molecular size (1879 bp) in the adult mosquito. **Fig.13A:** The same primer pair as in Fig.10E was used to detect the region of RA in the larva at the expected molecular size**.**

Transcripts in 1st instar larva

First, mRNA isolated from 1st instar larvae was used to identify the common exons present in all transcripts. A single primer pair was used which was expected to result in a 458 bp fragment for the RC transcript (12D) and 520 bp for the other three transcripts (Fig.12A, 12B, 12C). Both fragments were amplified in the larval sample (Fig. 13C). Next, I used the specific pairs for the RA, RB, and RD transcript to determine which of these transcripts were present in larvae (Fig.12A, 12B, 12C). Only RA and RD were amplified (Fig. 13 A, B). Notably, the molecular size of RD transcript differed from the expected (2 kb, expected size 1765). This was not further investigated.

Transcripts in adult mosquito

The same primer pairs used for the larva were also used to detect the common region, the RC transcript (Fig.12D), the RD transcript (Fig.12B) and the RB transcript (Fig.12C) in uninfected adult female mosquito midguts. PCR fragments were generated only for the common region (Fig. 13G) and the RD transcript (13F). Their molecular sizes were consistent with the results in the larva using the same primer pairs. The same primer pair used to detect RA in the larva (Fig.12A) amplified a region spanning exon 1 and exon 12 (1879 bp) in the adult (Fig.13E) with a molecular size consistent with the result in the larva (13A). This fragment includes the sequence encoding the antibody binding region in exon 2. A different primer pair (RAN/RAP, Fig.12A) was used to amplify a fragment spanning exon 1 and exon 2 of RA (1641 bp). However, a 1890 bp PCR fragment was generated instead (Fig.13D). Sequencing analysis revealed that this fragment contains sequences corresponding to three novel exons located between exon 1 and exon 2. Therefore, this is an alternative transcript variant, hereinafter called *grh* RX transcript (Fig.14A and see Suppl. Fig.4). Using the same primer pair, RX was detected in cDNA from infected adult female mosquito midguts as well (Fig.14B). The calculated length of the RX transcript is 3661 kb. (See Suppl. Fig.3). The protein coming out of this transcript is called Grh PX isoform and its sequence is given in Fig.15. A nuclear localization signal has been identified for this isoform and for Grh PD protein isoform consistent with Grh nuclear localization (Suppl. Fig.5A, Fig.5B).

Figure 14A

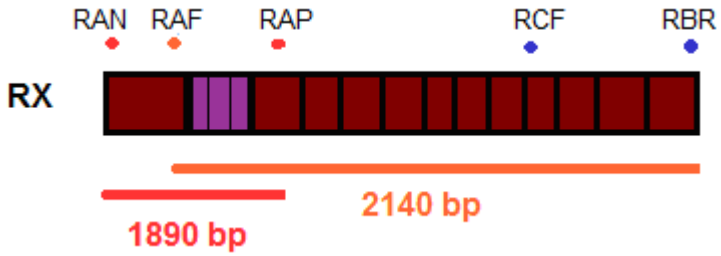


Figure 14B

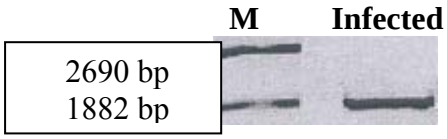


Fig.14A: The RX transcript has 3 mini exons indicated in purple between Exons 1 and 2 of the annotated RA transcript.
Fig.14B: A band consistent with the molecular size of the RX transcript was detected in infected female mosquito midguts as well using primer pair RAN/RAP.

Figure 15: The protein sequence for *Anopheles* Grh PX isoform is given below:

MSASPEMHHQHQQLQQEANAPLEMKSNSAEGTPPPELATMTTVSVLDLHKDYNGGGGGGG
GGTAESGATAGAVTSPHIVHEGATDMSLPDDGTTEKVYDKDTNTVYVYTTAAGVAGHKLV
VNPHHHQLTTIVHGGQQQQQQQQQQQQQQQMASPDQLHPSEHHAVAEQNLLHARLIQQQQ
QAAEEQQQQQQQQQQHLQRMSPGDPHQQHHQQHSQVHPDDSGIIDGHRLLPATINGTDAS
DSQQHQQQQHHHLGRLSPEDQQQQQAHQQGGVRLLEDSHIQRLLGNQEIISRDIINGEHH
IITGNENGETILTRIAISTADQLLNRMDNGIIYTTTGGSTGVIGAGPQEQLPTTVLQYEK
DVEDKHQPQQQQQQHHGHAHAAHQPQTIYATAGAAPDQTGQTKQIVYALGGGEPKNVIYG
DPKAAMPHFEAVSGAGSGAGSGGPGSVEEEKPQIDYVYNEGNKTVIYTDQKGLESLYANN
ELGLMDGTQIVVQSNLYTQQQGPDGTTVYVVSSDMNPEDINGLQQSTNAGAKLNGQTLQA
MDLLLGAHPSSQAINVKREPEDLRKEPKNPRNQKGPSHQNSTAATASSNVNTNSPSPSSY
AQYDMYPPNRLGPGGTTFITEPYTYREYFDNQGYAPARTIYGTAADSEGPQPATTYEGRF
TKTGSIYTKTITSAGLTVDLPSPDSGIGADAITPRDQNNVQQQFDYAEPCQAPIGMVDPN
AAGHIPACVASLQRNLAINGSQPSPTTSLGGSSTAAAVAVAGAAAAPRSRPWHDFGRQND
ADKVQIPKIYTDVGFKYYLESPISSSQRREDDRITYINKGQFYGITLEYVHDPDKPLKNQ
TVKSVIMLLFREEKSPEDEIKAWQFWHSRQHSVKQRILDADTKNSVGLAGCIEEVSHNAI
AVYWNPLESSAKINVAVQCLSTDFSSQKGVKGLPLHLQIDTFEDPRDTSVFHRGYCQIKV
FCDKGAERKTRDEERRAAKRKMTATGRKKLDELYHPVVDRSEFYGMSDLMKPPVLFSPSE
DIDKLTSMDMQFYGHDADSLSGTSDNVKSPFLLHANKPATPTLKFHNHFPPDVPTSDKKD
PSIIMDGSMVTNSMVDFTPQIKRQRMTPPLSERVMLYVRQDNEDVYTPLHVVPPSTVGLL
NAIENKFKISSSRINTIYRKNKKGITARIDDDMIRHYCNEDIFILEVQRYEEDLYDITLT
ELPTH

Fig.15: The protein sequences indicated in purple, turquoise and green correspond to the novel exons discovered. The antibody binding region is highlighted in yellow. The cDNA sequence that corresponds to the protein region indicated in bold and italics was sequenced.

Detection of the *stit* transcript

Two primers were used to bind to exon 1 and 3 amplifying a fragment of 1500 bp (Fig.16).

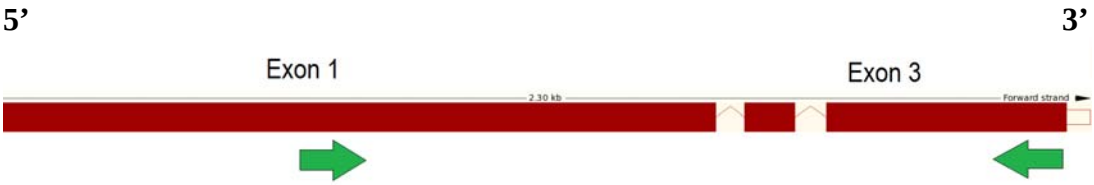


Fig.16: The primers designed to detect *stit* correspond to the 1500 bp cytoplasmic region of the respective protein with the forward primer binding to Exon 1 and the reverse primer binding to Exon 3.

RT-PCR was conducted on RNA extracted from uninfected and infected female adult mosquito midguts. The transcript detected was in accordance with the expected result (Fig.17) and was also confirmed by sequencing (See Supl. Fig.6).

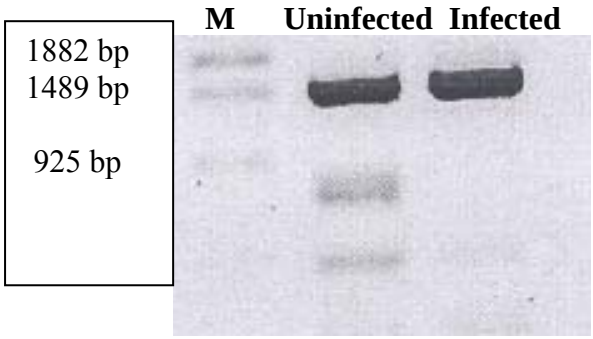


Fig.17: *stit* transcript was detected at the expected molecular size (1500 bp) in uninfected and infected adult female mosquito midguts.

GRH AND STIT LOCALIZATION IN MOSQUITO MIDGUTS

Grh has nuclear localization in uninfected midguts.

In order to detect Grh and Stit proteins in *A.gambiae,* midguts were immunolabeled using the antibodies recognizing each protein. Uninfected mosquito midguts were dissected and labeled with antibodies directed against Grh (green) and mosquito epithelial cadherin (Cad), a cell membrane marker (red). Nuclei were stained with TOPRO (blue). Grh is localized in the nucleus of uninfected female midguts consistent with the fact that a potential NLS has been detected for the Grh protein isoforms PX and PD (Fig.18A, 18B and Suppl. Fig.5A and Fig.5B).

Figure 18A

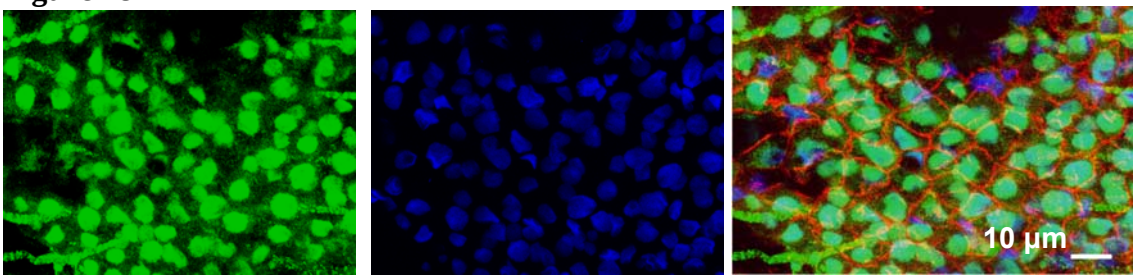
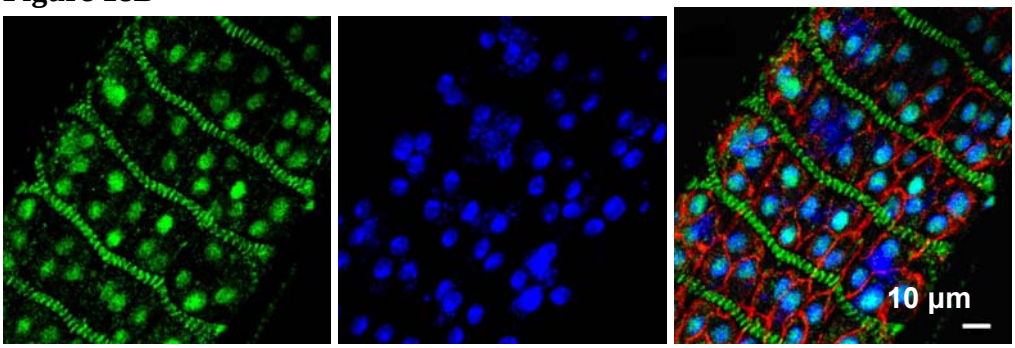


Figure 18B



Grh (green) is localized in the nucleus (blue) for uninfected female *A.gambiae* midguts.
Cad (red) is the cell membrane marker.**Fig.18A** shows midgut cells whereas **Fig.18B** shows anterior midgut cells.

Grh translocates from the nucleus to the cytoplasm after wounding of the midgut tissue in uninfected midguts.

In order to test whether aseptic epidermal injury of the midgut can elicit the response, sugar-fed female midguts were wounded. The midguts were injured with an insulin needle, then incubated in Ringer's solution for 30 mins to trigger the response and stained for Grh.
Surprisingly, Grh acquired a ring-like localization 30 mins after midgut tissue injury, suggesting cytoplasmic localization.
In order to investigate this further, the experiment was repeated and TOPRO stain was applied to stain the nuclei. This confirmed that Grh is localized in the cytoplasm after wounding the midgut cells. The response was systemic and not limited only in cells near the site of injury (Fig.19A, 19B and 19C).

Figure 19A

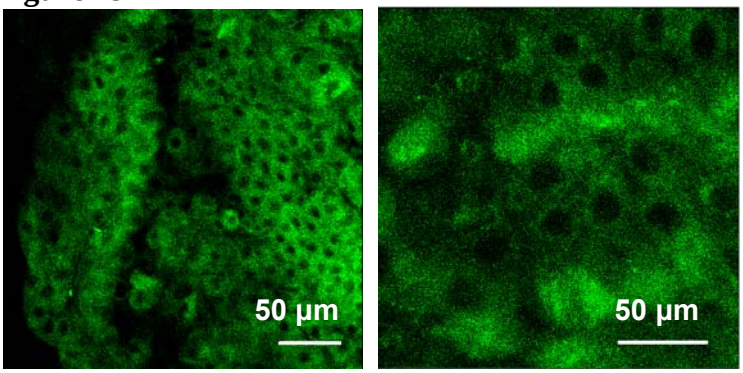


Figure 19B

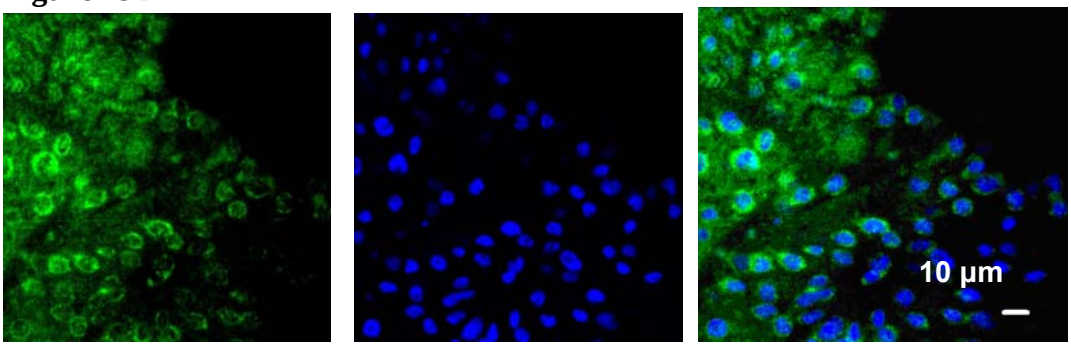


Figure 19C

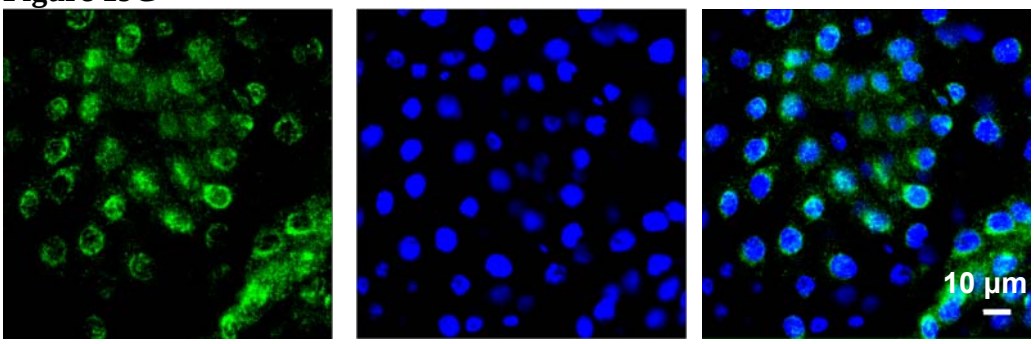


Fig.19A: *A. gambiae* midguts were dissected and wounded with an insulin needle in order to elicit the wound-healing response. Grh (green) obtained a ring-like localization 30 mins after midgut tissue injury. **Fig.19B, Fig.19C**: The midguts were wounded and stained for Grh (green). Cell nuclei were stained using TOPRO (blue). Grh was localized in the cytoplasm in the midgut cells. The response was systemic and not limited only to cells near the site of injury.

Grh translocates from the nucleus to the cytoplasm 24h after infection with the rodent malaria parasite *P.berghei.*

At the ookinete life-stage of the parasite, the parasite traverses the midgut epithelium thus wounding it. In order to clarify whether ookinete midgut traversal can elicit the wound-healing response, mosquitoes were fed on a mouse infected with a *P. berghei* parasite strongly expressing GFP. Female midguts were dissected 24 h post-feeding, when the ookinete is penetrating the epithelium. The blood was removed and the midguts were stained for Grh. In this experiment, a heavy infection was established with many ookinetes traversing the epithelium of a single midgut. Again, Grh was found in the cytoplasm of the midgut cells (Fig.20A and 20B). The cytoplasmic localization was also seen in cells at a distance from the ookinetes implicating that the response is systemic since all midgut tissue cells participated irrespectively of their proximity at the site of parasitic infection. Hence, it seems that Grh translocation is not limited to sites near ookinete wounding alone (Fig.20C, 21A and 21B).

Figure 20A

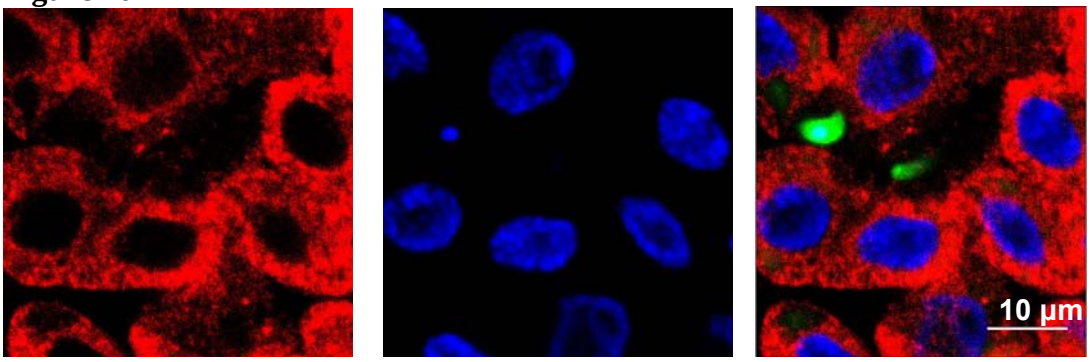


Figure 20B

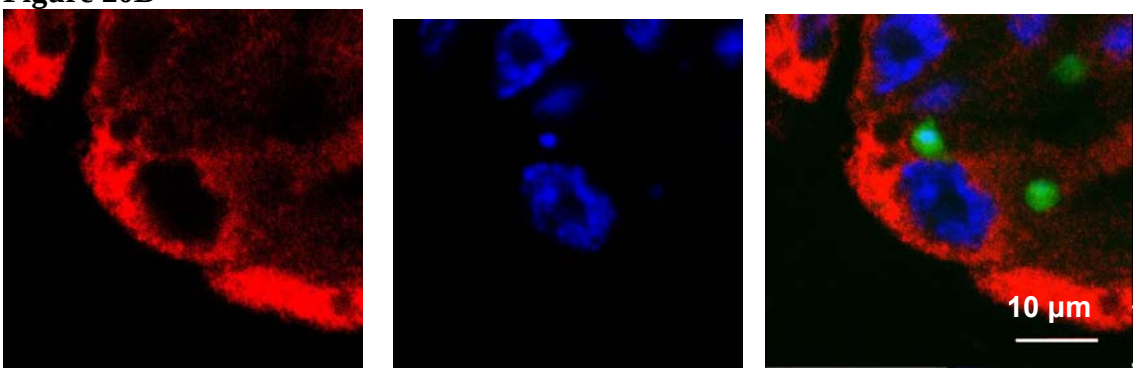


Figure 20C

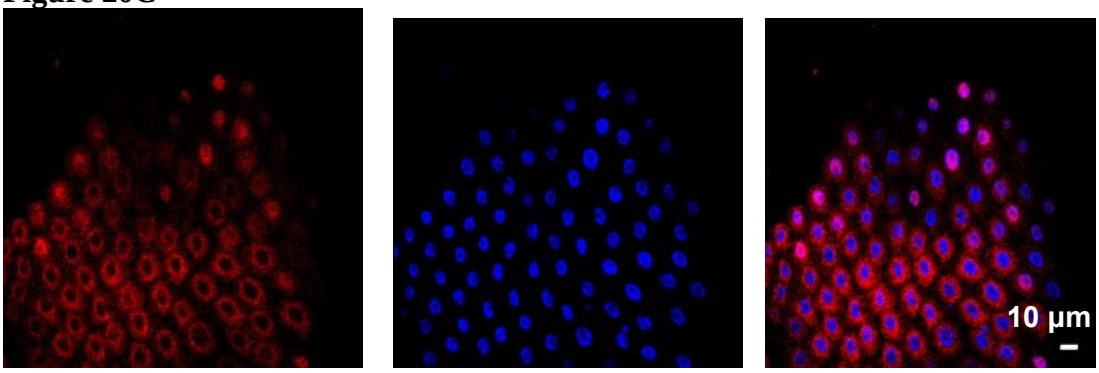


Fig.20A, Fig.20B: Infected blood-fed midguts were dissected 24 h post-feeding and the midguts were stained for Grh (red). DNA (cell nuclei and parasitic DNA shown in blue) was stained using TOPRO stain. Grh translocated to the cytoplasm during ookinete (green) midgut traversal.

Fig.20C: The midgut was stained as described in Fig.16A and Fig.16B. The response elicited 24 h post-feeding was systemic and not limited at sites near parasitic infection alone. Grh is shown in red and nuclei are shown in blue.

Figure 21A

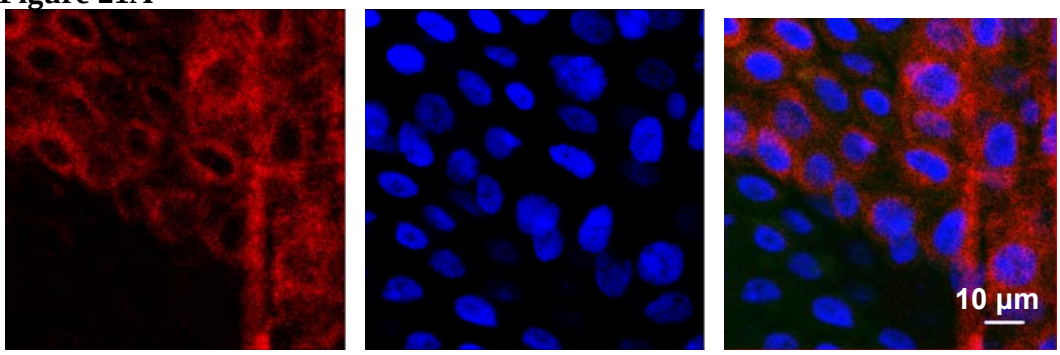


Fig.21A: The midgut was stained as described in Fig.20A and Fig.20B. The response elicited 24 h post-feeding was systemic and not limited at sites near ookinetes. Grh is shown in red and nuclei are shown in blue Parasites were detected in other fields in the same gut.

The previous experiment was repeated but this time a lighter infection was established with fewer ookinetes traversing the midgut epithelium. Again, Grh was found in cytoplasmic localization (Fig.22A) that was not limited at sites near ookinetes, implicating therefore that even in the event of a light infection, the response seems systemic (22B and 22C). Hence, it seems that the wound-healing mechanism has a global mode of action for both septic and aseptic midgut tissue injury since the response was systemic for the midguts wounded with the insulin needle as well (Fig.19A).

Figure 22A

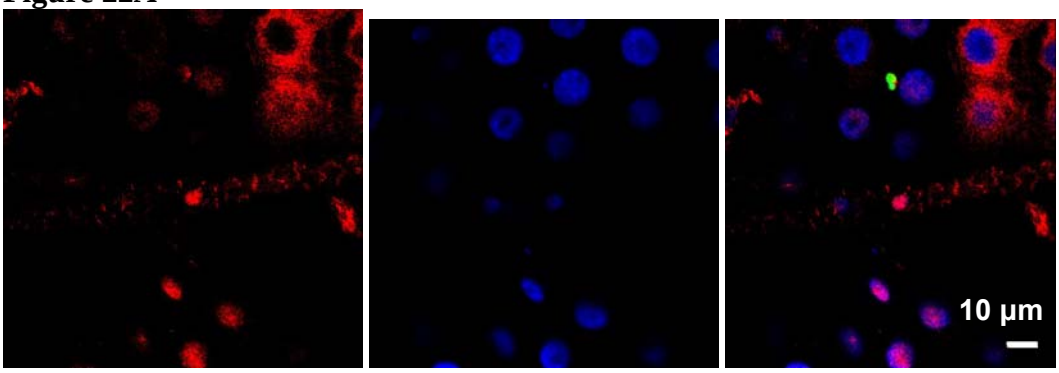


Figure 22B

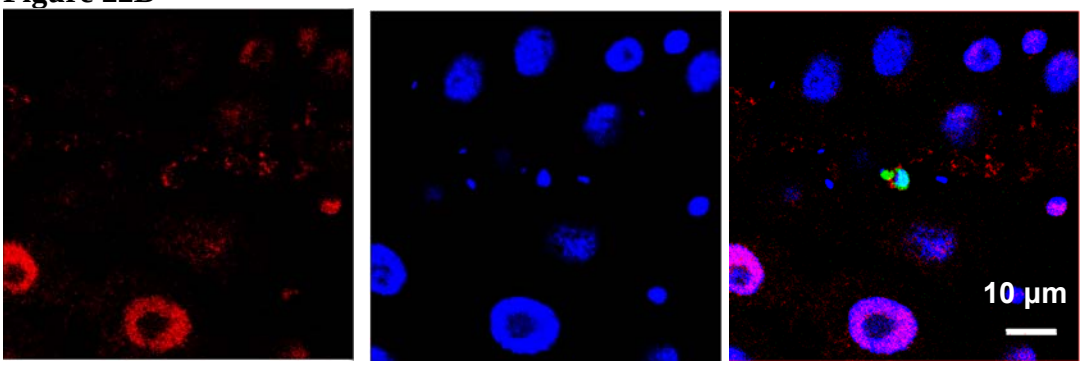


Figure 22C

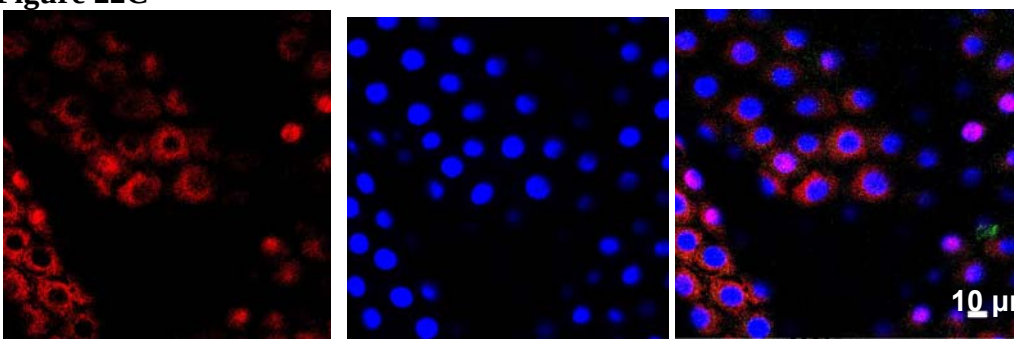


Fig.22A: The midgut was stained as described in Fig.20A and Fig.20B. In this experiment, fewer ookinetes traversed the midgut epithelium. Grh (red) acquired cytoplasmic localization near the parasite (green). DNA (cell nuclei and parasitic DNA) is shown in blue. **Fig.22B and 22C:** Images from the same midgut, different stacks. **Fig.22C**: Cytoplasmic localization for Grh (red) is observed at a different section level from ookinetes (green) shown in **Fig.22B**. Nuclei are shown in blue.

Grh is localized in the nucleus in blood-fed uninfected midguts.

When *Anopheline* mosquitoes take a blood-meal, their midguts expand and dilate, sometimes to such an extent, that one could possibly assume that minor tissue injuries are provoked during blood-feeding or shortly afterwards. Hence, in order to check whether blood-feeding alone can trigger Grh translocation and not parasitic infection per se, mosquitoes were fed on a healthy mouse and midgut dissection followed ~24 h post feeding. The blood was removed from the midguts as was done for the infected midguts and staining for Grh followed. Grh was observed in the nuclei of midguts cells suggesting therefore that blood feeding alone cannot trigger the translocation (Fig. 23A and 23B). Hence, Grh translocation in infected midgut cells was specifically triggered by the parasite.

Figure 23A

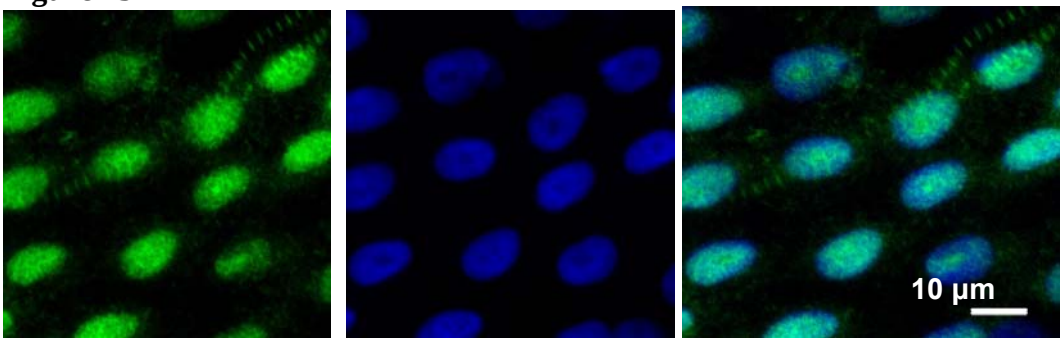


Figure 23B

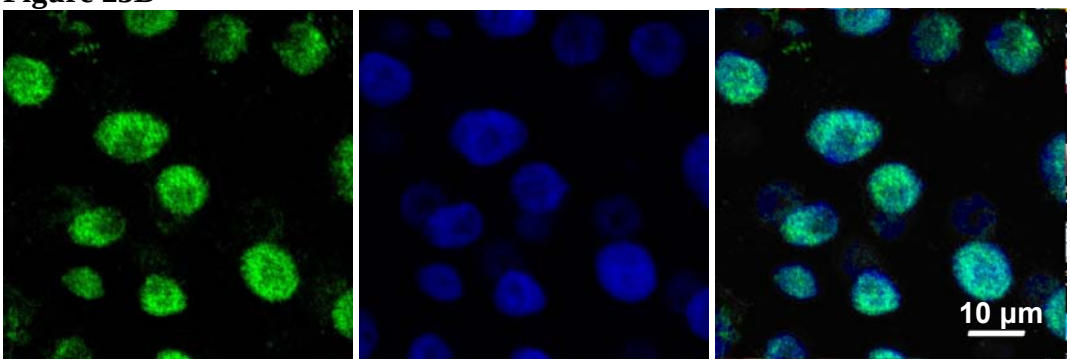


Fig.23A, Fig.23B: *A.gambiae* midguts were dissected 24 h after feeding on a uninfected, healthy mouse. The blood was removed and staining followed for Grh (green). Nuclei are shown in blue after TOPRO staining. Grh localized in the nucleus 24 h post-feeding with not infected blood.

Detecting the time window for Grh translocation

In order to clarify the time window during which Grh resides in the cytoplasm, infected blood-fed midguts were dissected 27 h post-feeding. The blood was removed and the midguts were stained for Grh (red) and DNA (blue). In these experiments Grh is partly nuclear and partly cytoplasmic, suggesting that Grh translocates back to the nucleus at this time point (Fig.24A, 24B and 24C).

Figure 24A

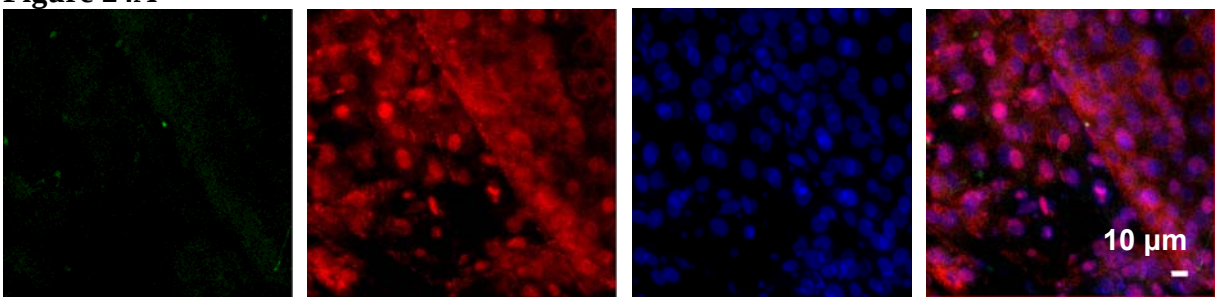


Figure 24B

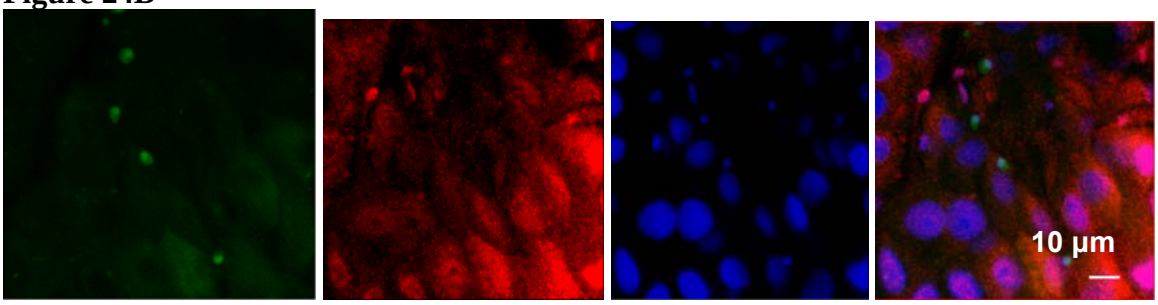


Figure 24C

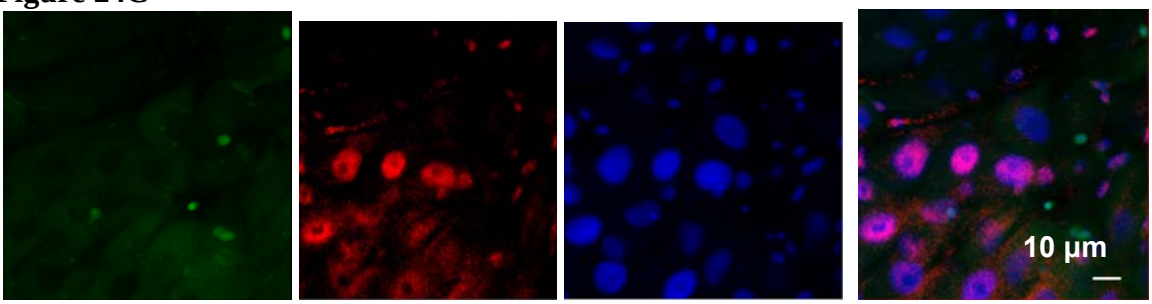


Fig.24A, 24B and 24C: Mosquito midguts infected with the GFP parasite (green) were dissected 27 h post-feeding and were stained for Grh (red) and DNA (blue). Grh acquires partly nuclear and partly cytoplasmic localization at this time point.

Grh is localized in the nucleus 13 days after an infectious bloodmeal.

During oocyst growth the midgut could be perturbed as the occyst gains volume to reach its most mature form, 13 days post feeding with infected blood. In order to examine whether this could affect Grh translocation, midguts containing mature oocysts were dissected and stained for Grh. Grh is clearly localized in the nucleus of the midgut cells implicating therefore that the response is not triggered as this time point (Fig.25A and 25B).

Figure 25A

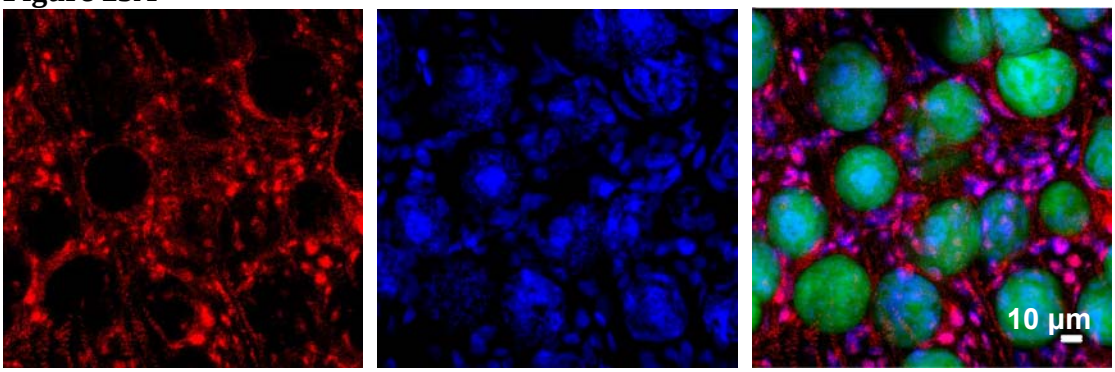


Figure 25B

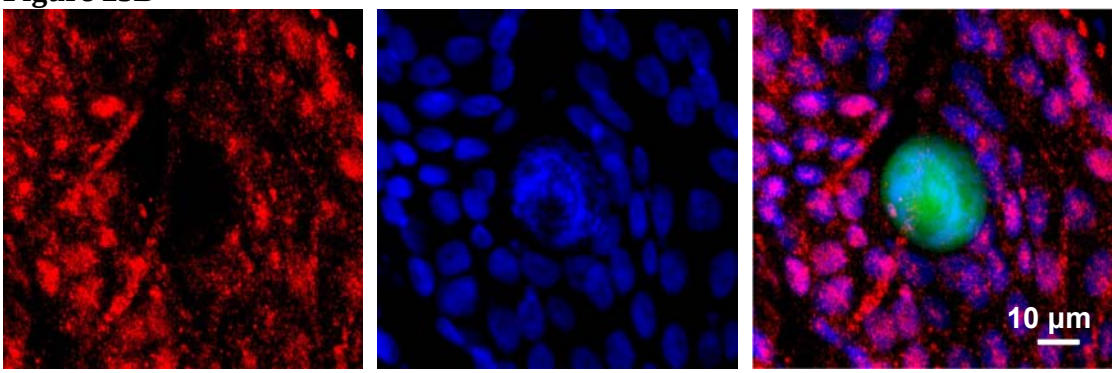


Fig.25A, Fig.25B: *A. gambiae* midguts were dissected 13 days after feeding with infected blood, at the mature oocyst life stage (green). The midguts were stained for Grh (red) that localizes in the cell nuclei (blue). The parasites (13-day oocysts) are shown in green.

Stit localization in uninfected and infected midguts.

Uninfected midguts

Uninfected female midguts were dissected and stained for Stit. Stit was detected in the cytoplasm of uninfected midgut cells consistent with Samakovlis group observations for cytoplasmic *Drosophila* Stit localization (Fig.26A, 26B, 26C).

Figure 26A

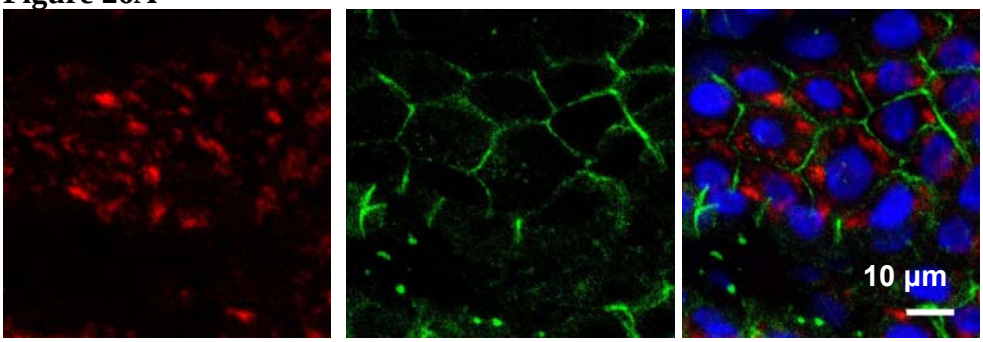


Figure 26B

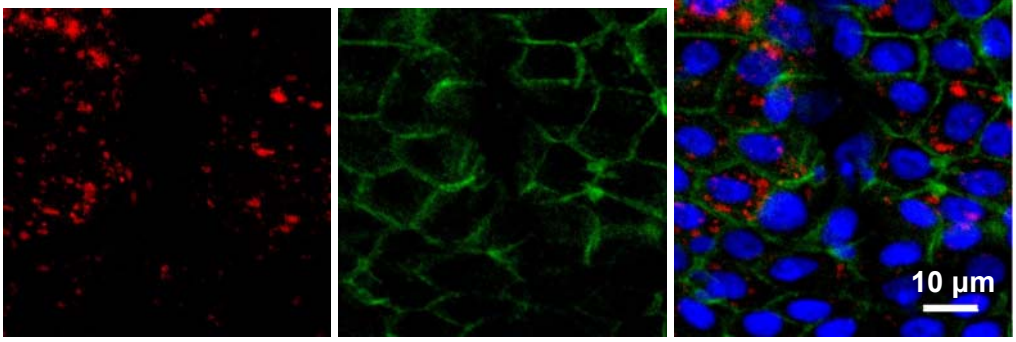


Figure 26C

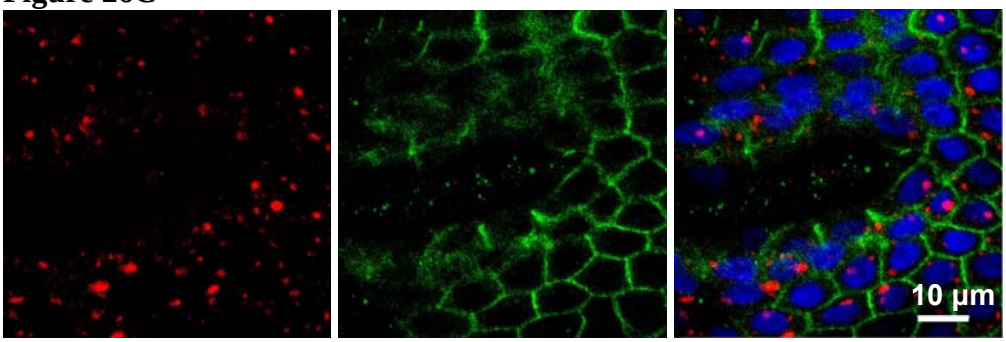


Fig.26A, 26B, 26C: Uninfected *A. gambiae* midguts were stained for Stit (red) and Cad (green). The nuclei were stained in blue using TOPRO stain (blue). Stit is detected in the cytoplasm of the midgut cells.

Infected midguts

Infected midguts were dissected ~24h post feeding, the blood was removed and the midguts were stained for Stit. Stit was detected in the cytoplasm of midgut cells. Stit signal seemed clustered in some infected midguts with a cytoplasmic assymetrical distribution (Fig.27A and Fig.27B). Then Stit progressively was observed to relocalize in the entire cytoplasm of midgut cells (Fig.27C and Fig.27D).

Figure 27A

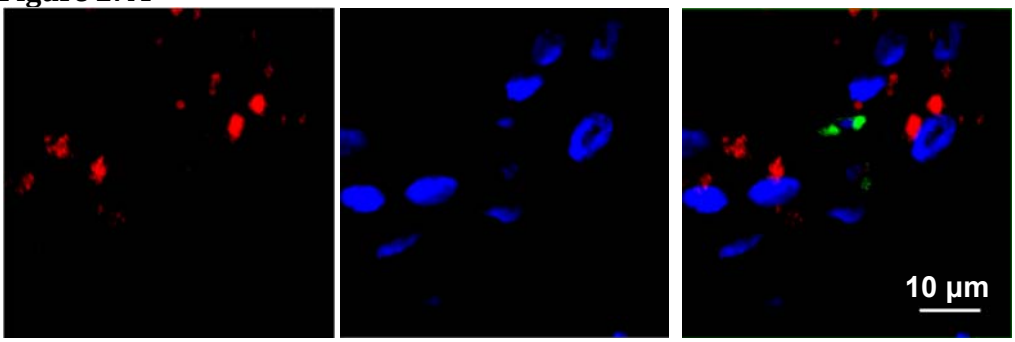


Figure 27B

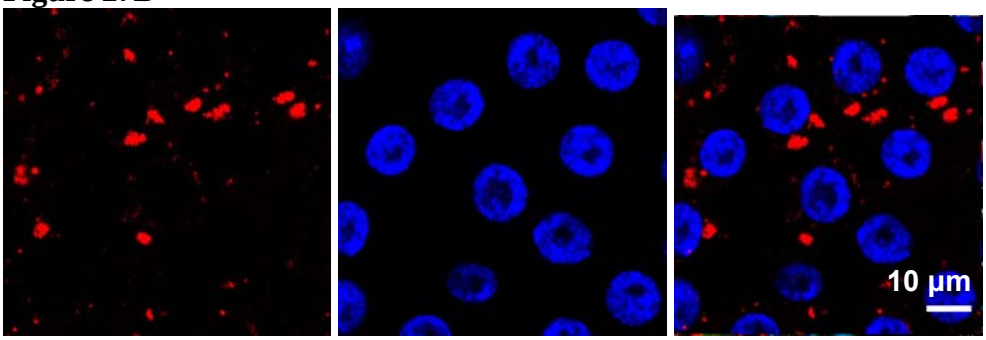


Figure 27C

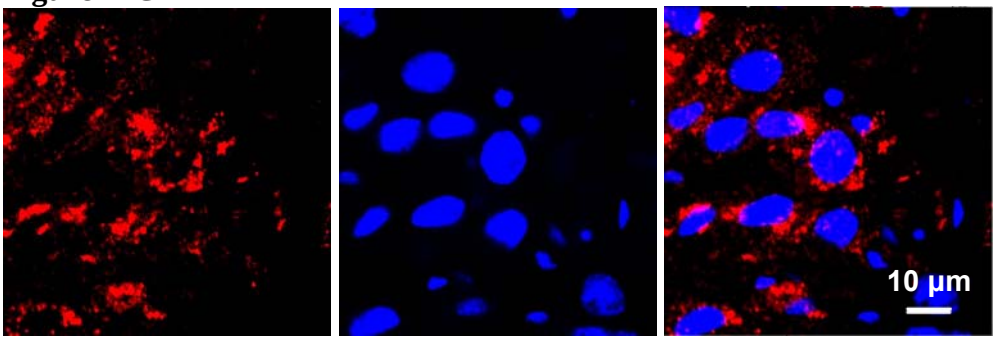


Figure 27D

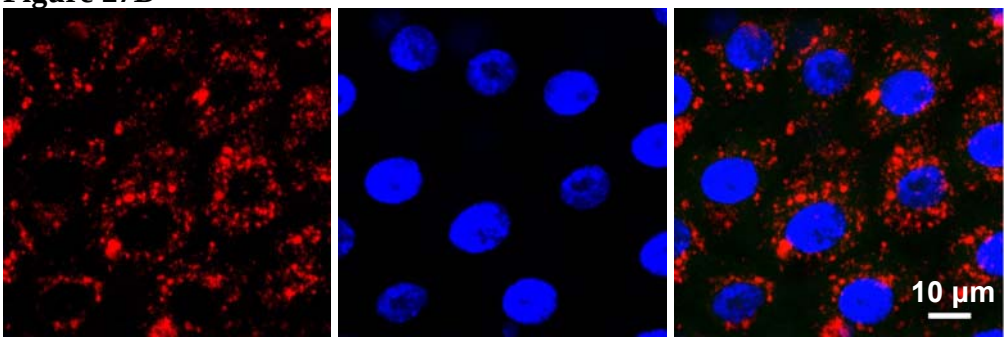


Fig.27A, 27B, 27C, 27D: *A.gambiae* midguts were stained for Stit (red). Cell nuclei were stained in blue. The images show pieces of different midguts. Stitcher signal seems to relocalize in the cytoplasm after infection. In **Fig.27A** and **Fig.27B**, Stit is clustered and has asymmetric distribution near the parasite (green). In **Fig.27C**, Stit maintains the asymmetric distribution and starts to relocalize in the entire cytoplasm. Its signal is again clustered. In **Fig.27D** Stit has reached equal distribution in the entire cytoplasm.

Stit in midguts containing day 13 oocysts

13 days after infected blood meal midguts were dissected and labeled for Stit. Stit was detected in these midguts in the cytoplasm of the cells. Its signal was similar to uninfected midguts (Fig.28A, 28B, 28C)

Figure 28A

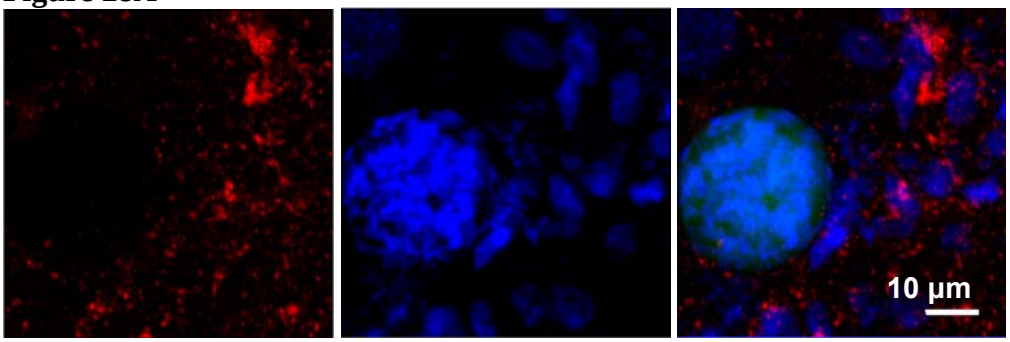


Figure 28B

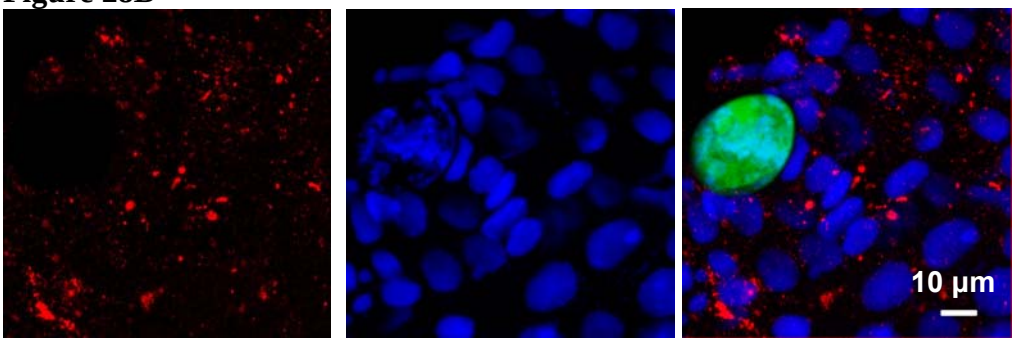


Figure 28C

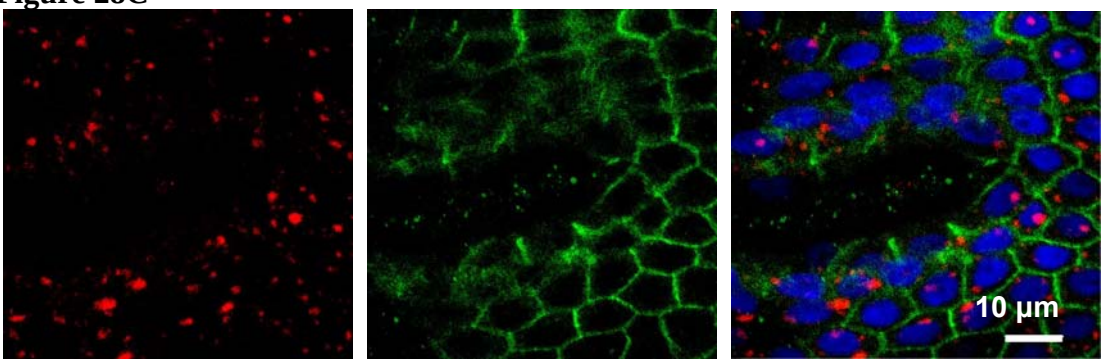


Fig.28A, Fig.28B: *A.gambiae* midguts were dissected 13 days after the ookinete life stage, at the oocyst life stage. The midguts were stained for Stit shown in red. DNA (cell nuclei and parasitic DNA) was stained in blue using the TOPRO stain. The pattern for Stit signal resembles the pattern observed for uninfected mosquito midguts. **Fig.28C**: A representative sample for Stit staining in uninfected midguts is shown for comparison.

SILENCING *grh* EXPRESSION USING dsRNA

Silencing experiments

In order to clarify whether the signal observed from the Grh antibody corresponds indeed to the respective protein in *Anopheles* and also to check whether Grh depletion might have a possible effect on ookinete traversal in the mosquito midgut, silencing experiments were attempted using RNAi. Two independent silencing experiments are presented; for experiment 1, 1000ng/ul dsRNA targeting *grh* RX or GFP was injected in 1.5 day old mosquitoes and midguts were stained 5 days post injection to assess signal depletion of the Grh antibody. For experiment 2, 3000 ng/ul dsRNA were injected in 3 day old mosquitoes, silencing was assessed 3 days post injection and the remaining mosquitoes were used in order to assess whether *Anopheles* Grh could play a role during parasite infection.

Experiment 1

DsRNA was generated using a PCR fragment (Fig.29) corresponding to a 470 bp region of *grh* RX transcript exon 1 (for the protein sequence targeted refer to Supplementary Fig.7). DsRNA for both Grh and GFP was successfully produced. DsRNA molecules migrate slightly differently on agarose gels than DNA molecules (marker) so the gel image below shows an approximation of their true molecular sizes (Fig.30).

Figure 29

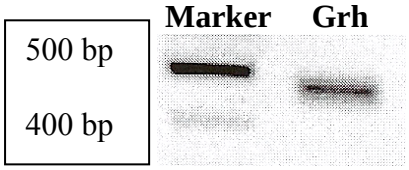


Figure 30

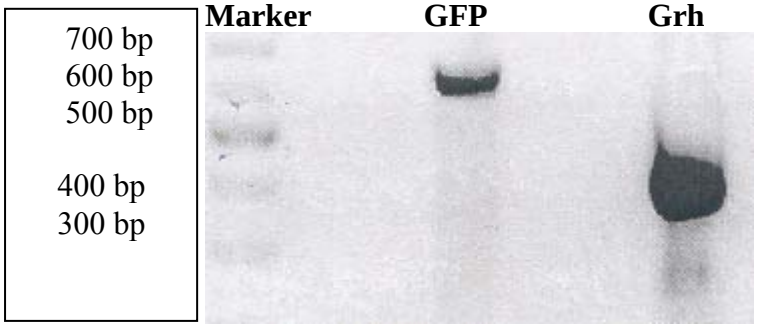


Fig.29: The DNA probe for *grh* ran at the expected molecular size after amplifying a 470 bp region of *grh* RX exon 1.
Fig.30: DsRNA for *grh* or GFP ran at the expected molecular sizes after purification. (442 bp for *grh* dsRNA and 650 bp for GFP dsRNA)

The dsRNA was injected into uninfected *A.gambiae* mosquitoes and silencing was assessed with immunofluorescence and confocal microscopy using mosquitoes injected with dsRNA against GFP as a control. 1.5 day old mosquitoes were used for this silencing experiment as it was suspected that Grh protein turnover occurs just before sexual maturation is complete (before mosquitoes reach the 2nd day of adulthood). Seven midguts were stained for Grh (green), Cad (red) and DNA (blue) (Fig.31A, 31B, 32A and 32B). Intensity of the Grh signal was measured integrating data from 50 cell nuclei of 5

silenced midguts and 50 nuclei from 2 control midguts. A statistically significant downregulation was measured (t-test result: p value=0.0001, Fig.33).

Figure 31A

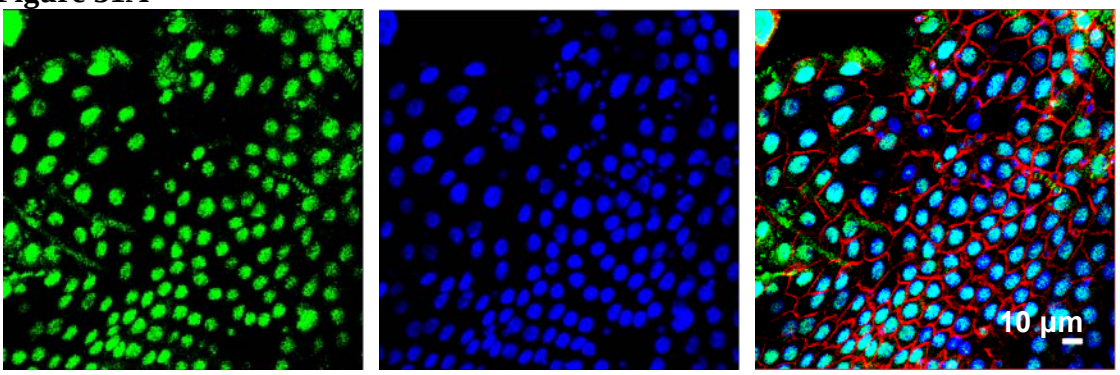


Figure 31B

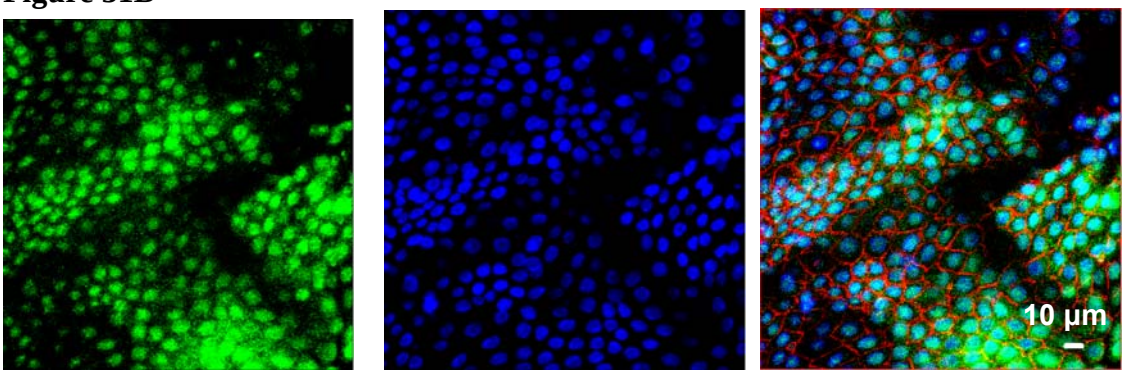


Figure 32A

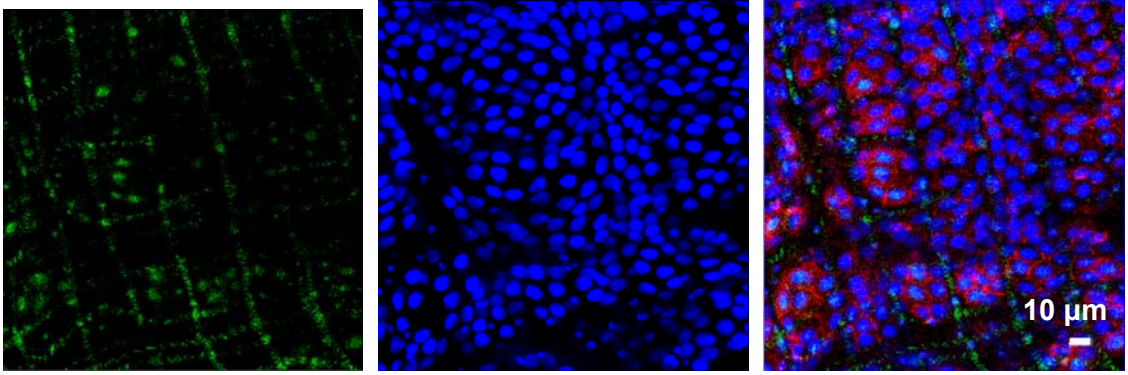


Figure 32B

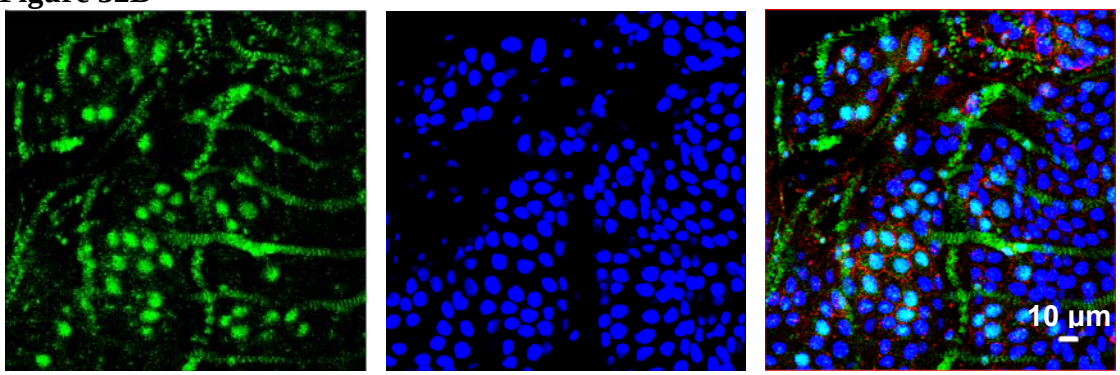


Fig.31A, Fig.31B: *A.gambiae* mosquitoes were injected with dsRNA against GFP and their midguts were stained for Grh (green), Cad (red) and cell nuclei (blue). Again, Grh obtains nuclear localization for uninfected midguts.
Fig.32A, Fig.32B: *A.gambiae* mosquitoes were injected with dsRNA against *grh* RX transcript and their midguts were stained for Grh (green), Cad (red) and cell nuclei (blue). Again, Grh obtains nuclear localization for uninfected midguts.

Figure 33

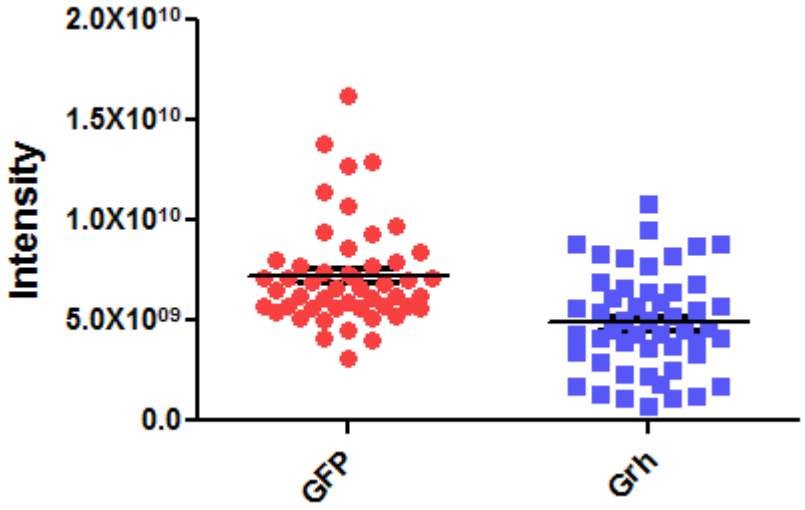


Fig.33: The average intensity of Grh signal was measured from 50 midgut cells of 5 silenced midguts and was compared with the average intensity of 50 midgut cells from 2 control midguts. A statistically significant downregulation was measured (t-test result: P<0.0001)

.

Experiment 2

Next, 3000ng/ul dsRNA was generated to maximize the RNAi effect. (Fig.34) Three days old mosquitoes (sexually mature mosquitoes) were used. The mosquitoes were fed on an infected mouse on day 2 post injection and were stained on day 3 post injection to assess silencing. Grh signal is more intense and robust in the *grh* silencing experiment (Fig.36A and 36B) compared to the control (Fig. 35A and 35B). This may be due to the fact that the mosquitoes in the control experiment ingested less blood than the *grh* KD mosquitoes. This could indicate that *grh* transcripts increase after blood-feeding. In support of an upregulation of transcription, microarray experiments for *A. gambiae* demonstrated an upregulation of *grh* transcripts that lasts for several days after bloodfeeding (Marinotti *et al.*, 2006, See Supl. Fig.8A and Fig.8B), whereas another set of microarray experiments show expression levels of the unique exon of RX/RA transcripts (exon 1) returning to the non-blood-fed situation 24 hr post-feeding (Marinotti *et al.*, 2006, See Supl. Fig.9A and Fig.9B). It should also be noted that the Grh antibody possibly recognizes both Grh isoforms (Fig. 8A, 8B and 8C). If the Grh antibody recognizes both PD and PX, a potential downregulation of the RX transcript would not be detected with immunofluorescnce, since the RNAi effect for RX would be masked by a massive upregulation of the *grh* RD transcript 24 h after bloodfeeding (Fig.35A, Fig.35B and Fig.36A, Fig.36B). Microarray experiments showed that in the 24-27 h time window after blood-feeding, *grh* expression levels do not change significantly (See Suppl. Fig.8A and Fig.9A).

Figure 34

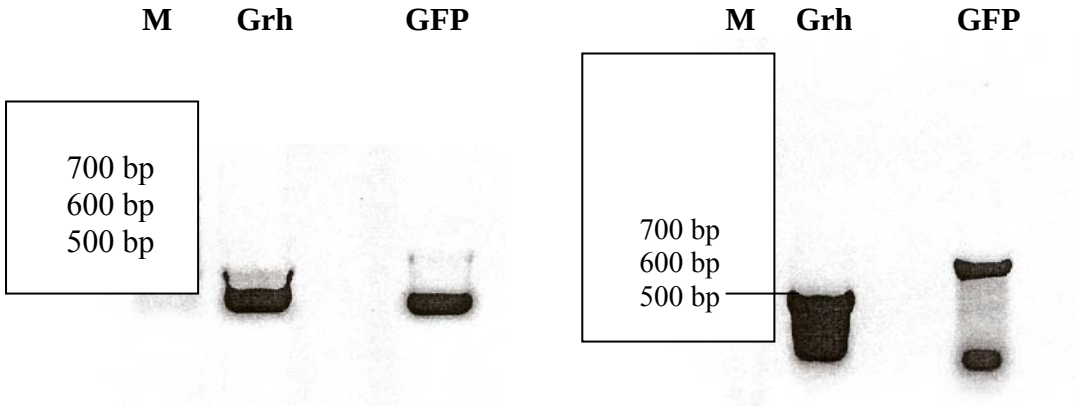


Fig.34: DsRNA was generated using the established procedure. DsRNA molecules run at the expected molecular sizes. The smear observed, reminiscent of a supercoiled plasmid, is the secondary structure of dsRNA molecules.

Figure 35A

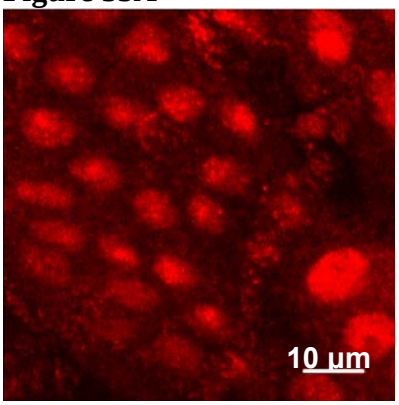


Figure 35B

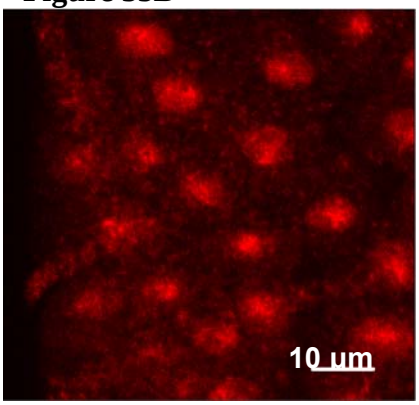


Figure 36A

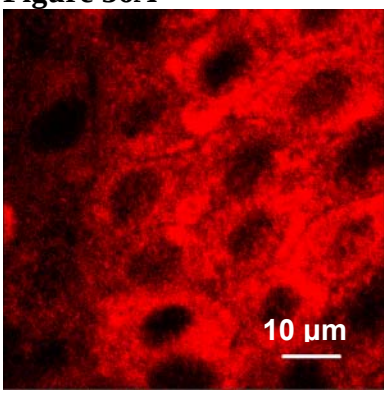


Figure 36B

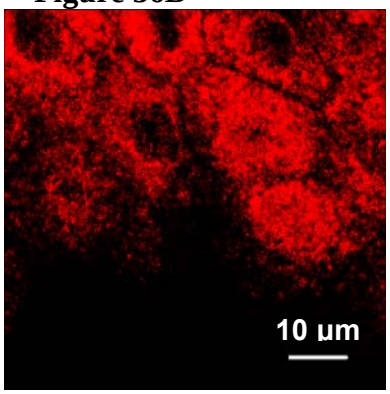


Figure 35A and 35B: Grh signal (red) as observed in GFP dsRNA injected midguts. These mosquitoes were only slightly fed on infected blood and were dissected at the 27 h timepoint so Grh is localized in the nucleus. **Figure 36A and 36B:** Grh signal (red) as observed in *grh* dsRNA injected midguts. These mosquitoes were adequately fed on infected blood and were dissected at the 24 h timepoint, so Grh has cytoplasmic localization.

Some of the mosquitoes injected with dsRNA for either Grh or GFP were not fed with infected blood and were stained for Grh (green) and DNA (blue) on day 3 post injection in order to clarify if silencing

is achieved in uninfected mosquitoes. Unfortunately, it was not possible to determine this because using these conditions the staining did not work well for the control experiment (Fig. 37A is the control experiment and Fig. 37B and 37C show the actual silencing experiment).

Figure 37A

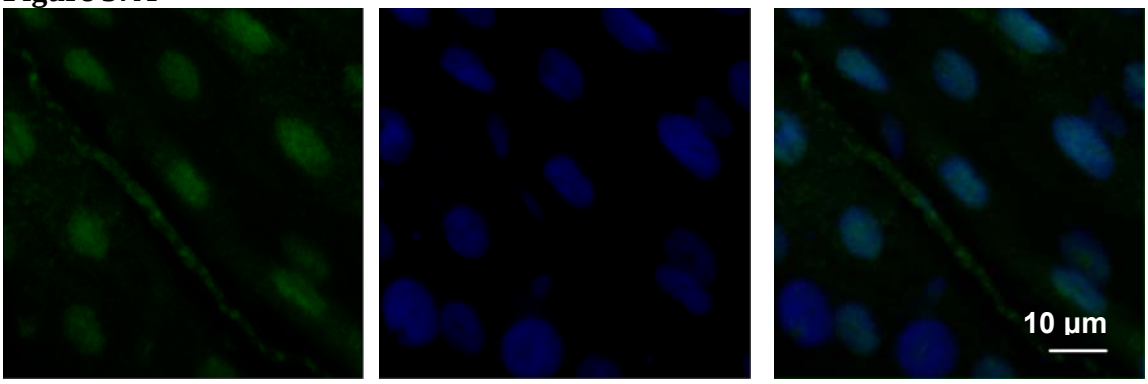


Figure 37B

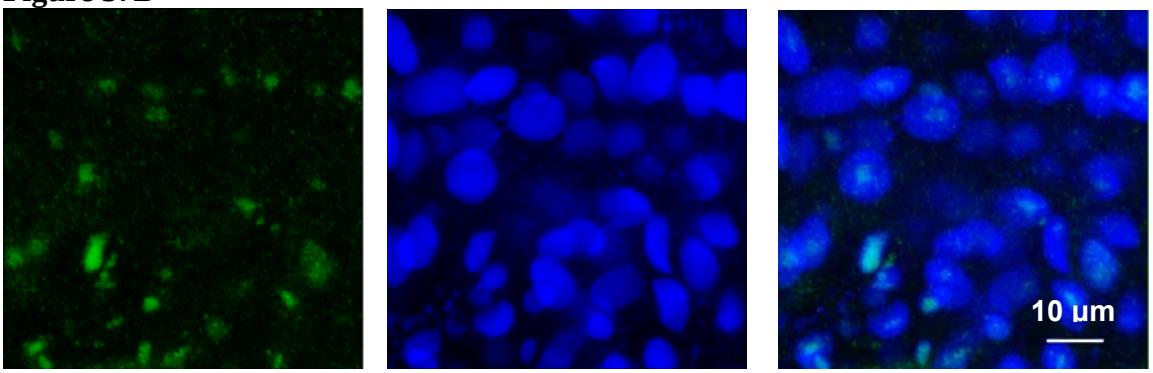


Figure 37C

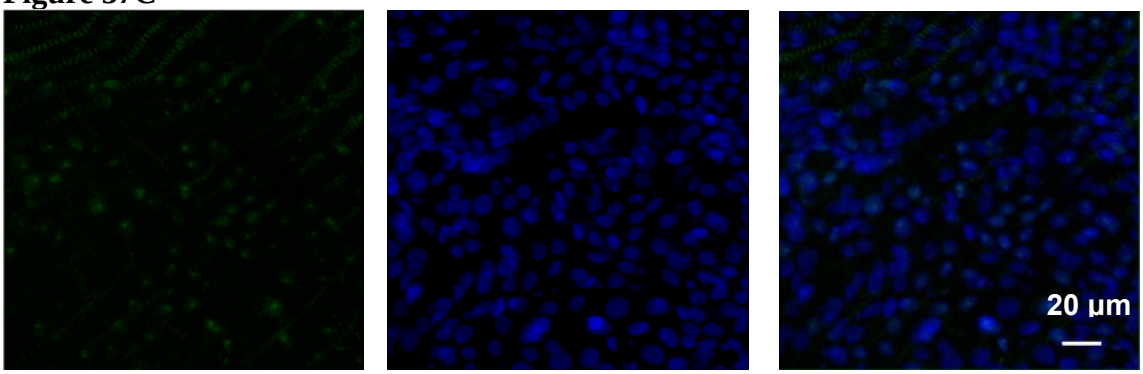


Fig. 37A: Grh signal (green) as observed in uninfected midguts injected with dsRNA for GFP (control experiment) Grh signal seems solid and covers the cell nuclei (blue). As in all uninfected midguts examined so far, Grh localizes in the nucleus.

Fig. 37B and 37C: Grh signal (green) as observed in uninfected midguts injected with dsRNA for *grh* RA transcript (actual experiment). It is not clear whether Grh was indeed downregulated in uninfected mosquitoes. Cell nuclei are shown in blue.

Oocyst development in *grh* silenced mosquitoes

10 infected mosquitoes were kept in order to count oocysts at day 6 post blood feeding. Only 2 were infected with 100 and 14 oocysts, respectively. Surprisingly, these oocysts were unusually large with an average diameter 18 μm (Fig.38A). Non-injected mosquitoes infected with the same parasite (conGFP) (Fig.38B) gave statistically significant smaller oocysts than the *grh* KD mosquitoes (t-test result: p value<0.0001, Fig.39). However, since only a small sample was examined and the experiment was performed only once, no strong conclusions can be deduced for an actual effect on infection.

Figure 38A

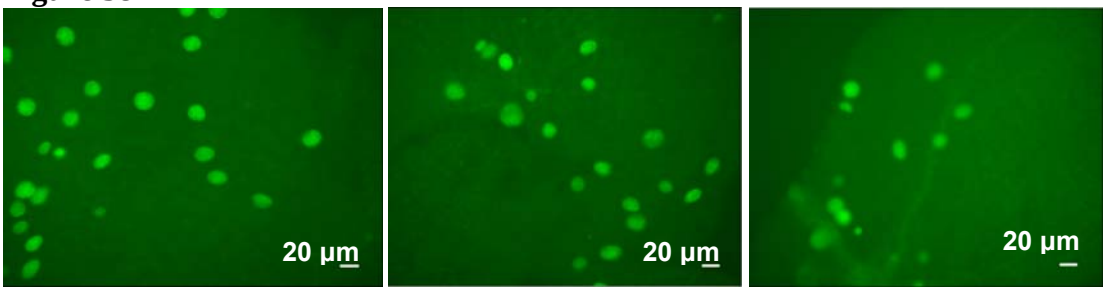


Figure 38B

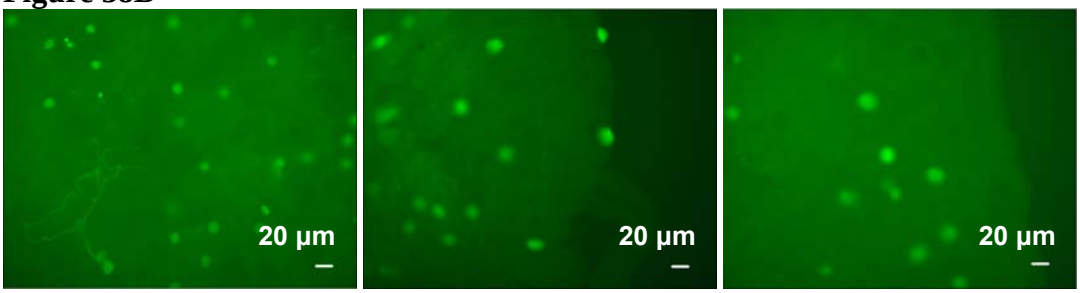


Fig.38A: The images above show 6-day oocysts from 2 midguts of *grh* RX KD mosquitoes that seem unusually large. The average diameter was found to be 18 μm. **Fig.38B**: Non-injected mosquitoes were infected with the same parasite (conGFP) and examined for 6-day oocysts. The 3 images show oocysts from 3 different midguts.

Figure 39

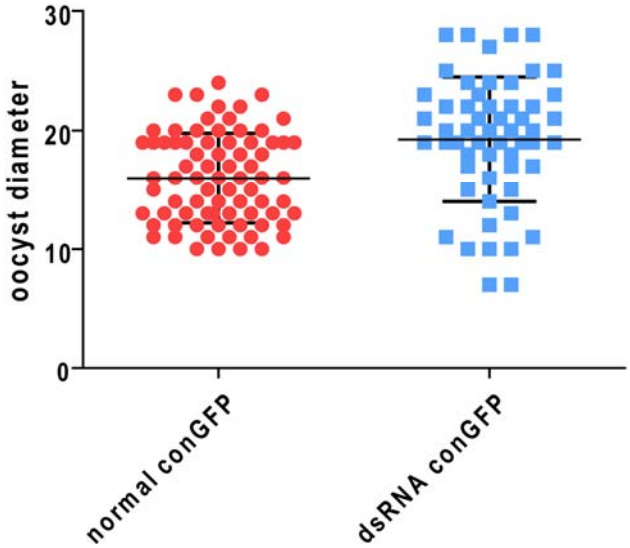


Fig.39: Diameter from 6-day oocysts was measured. The oocyst diameter of the *grh* KD infected mosquitoes was larger than the oocyst diameter from non-injected mosquitoes and the difference was statistically significant (t-test result: p-value<0.0001).

DISCUSSION

In this study it was examined whether two proteins, Stit, a Ret family Tyrosine Kinase (RTK) receptor and Grh, a transcription factor, respond to infection with the rodent malaria parasite *P.berghei* in *A. gambiae* midgut cells. Elucidating their potential role in triggering a wound-healing mechanism in the mosquito midgut that could either facilitate or hinder ookinete traversal would provide useful insight for malaria transmission in the mosquito host. Initial experiments using Western blot to detect the two proteins were not conclusive as it seems that these antibodies do not work on Western blot experiments.

In order to clarify which of the *grh* transcripts deriving from alternative splicing are expressed in the midgut, RNA was extracted from the midguts for RT-PCR. Only the transcripts RX and RD were detected in the adult mosquito corresponding to the Grh protein isoforms PX and PD. Even though a band consistent with the molecular size of the PB Grh protein was indeed detected using Western blot, the respective *grh* transcript RB could not be detected in larva or adult mosquito. An additional transcript, RC that corresponds to PC protein, was detected only in 1st instar larva, implicating therefore that as in *Drosophila*, *A.gambiae* Grh may have specific roles during the larva to adult mosquito developmental transition. What is more, using the exon sequence provided for *A.gambiae* strain, specific primers were designed in order to detect *grh* RA transcript in *A.gambiae*. Surprisingly though, sequencing confirmed a 2 kb PCR product as *grh* whereas the database indicated a 1.6 kb PCR product instead. This extra sequence corresponded to 3 mini exons with total sequence length of 261 bp, implicating therefore that the transcript detected is an alternative transcript variant for this gene that was not previously detected either by community annotation or prediction algorithms. The annotated *grh* RA transcript per se was not detected. One possible explanation for these results could be that RX is a tissue-specific transcript variant that is expressed preferably in the midgut. Furthermore, a potential NLS (Nuclear Localization Signal) was detected for the PX and PD Grh protein isoforms using the prediction algorithm Nucpred.

Regarding Stit, its transcript sequence is highly conserved for the cytoplasmic part between the species *Drosophila* and *Anopheles*. Sequencing confirmed the PCR fragment of the expected molecular size that corresponds to the cytoplasmic part as *stit*. The cytoplasmic part of this protein is responsible for signal transduction and possibly for signaling of the wound-healing mechanism.

Next, immunofluorescence was employed in order to detect Grh and Stit proteins in *A.gambiae* midguts during ookinete traversal (24 hrs post-feeding). Immunofluorescence technique maintains the native state of the proteins as the sample is fixated without spoiling the antigenicity of the proteins for antibody detection. Indeed, this approach was successful; more specifically, it was shown that Grh translocates from the nucleus to the cytoplasm 24 hrs after infection with the parasite (septic wounding) or after wounding with a sterile insulin needle (aseptic wounding), implicating therefore that contrary to its role in *Drosophila* wound-healing, Grh may have a different function in *Anopheles*. Grh localization in the nucleus is consistent with the detection of a potential NLS for both Grh protein isoforms PX and PD. In addition, Grh translocation is systemic with cells demonstrating Grh cytoplasmic localization in the proximity or away from ookinetes for both instances of heavy and light parasitic infection. Furthermore, a systemic Grh response was observed for aseptic midgut wounding with a sterile insulin needle implicating therefore that the wound-healing mechanism for *Anopheles* has a global mode of action for both septic and aseptic wounding. Grh translocation was not observed in uninfected blood-fed midguts suggesting therefore that the response was specifically triggered by the parasite. Furthermore, Grh was observed to acquire partly nuclear and partly cytoplasmic localization in cells near ookinetes 27 hrs post feeding with infected blood implicating therefore that Grh activity is tightly regulated and that there is a specific time window (approximately 3 hours) during which the wound-healing mechanism is active. The previous observation is in support with existing data from microarray experiments demonstrating a dynamic expression profile of *A.gambiae* midgut genes involved in parasitic infection before, during and after ookinete traversal (Vlachou *et al.,* 2005). What is

more, Grh was not observed in the cytoplasm in midguts infected with 13-day oocysts, implicating therefore that potential midgut perturbations caused by the increasing size of the oocyst does not trigger a similar response.

Signal from Stit antibody was observed in the cytoplasm. For uninfected mosquito midguts, Stit was localized in the cytoplasm whereas for infected mosquito midguts, Stit signal was observed in two variable patterns; it was either clustered for some of these midguts or observed in a symmetric distribution in the cytoplasm for some other midguts. The previous observations imply that Stit molecules may respond to infection by residing in clusters in the cytoplasm that progressively relocalize to reach equal distribution in the entire cytoplasm. The previous is in agreement with current hypotheses for *Drosophila* Stit response to aseptic *Drosophila* embryo wounding (Wang, 2010). Stit relocalization was not observed when staining midguts infected with 13-day oocysts. All in all, these data suggest that Stit relocalizes after infection with the parasite and also might suggest that Stit responds to parasitic infection (septic wounding) in a similar manner that *Drosophila* Stit might respond to aseptic embryo wounding (Wang, 2010).

Collectively, even though the Stit domain responsible for signal transduction (cytoplasmic part of the protein) is conserved between species, the downstream effector, Grh, seems to function differently from *Drosophila* Grh, possibly through repressing and not activating gene targets like in *Drosophila* wound-healing since *Anopheles* Grh was observed to leave the nucleus after wounding. The previous implies that if indeed the wound-healing cascade exists in the mosquito midgut, then it works differently from *D. melanogaster*.

In order to elucidate the role of Gh in infection and hence the nature of the Grh-mediated response, the RNAi technique was employed in order to downregulate the novel *grh* RX transcript. Grh, as the downstream effector, is expected to have a more specific role in the context of the wound-healing response in comparison to Stit that could receive other environmental stimuli to trigger signaling apart from stress signals. For this reason, targeting was attempted only for *grh*. In this study, results from two silencing experiments are provided; for experiment 1, the Grh signal depletion observed in uninfected mosquito midguts injected with dsRNA targeting the *grh* RX was statistically significant in comparison with the control. However, when the *grh* KD mosquitoes were exposed to infection to assess whether *grh* silencing affects parasitic infection, Grh signal was brighter 24hrs post-feeding in these mosquitoes rather than in the control population. Microarray experiments (Marinotti *et al.*, 2006) suggest that *grh* RD increases after blood feeding thus making assessment of silencing efficiency practically impossible during ookinete traversal using immunofluorescence. This is because Grh antibody could recognize both Grh isoforms and not just the PX isoform thus yielding non-specific signal from Grh PD (the protein produced from *grh* RD transcript) that masks the signal from Grh PX (protein produced from *grh* RX transcript). Hence, since downregulation for blood-fed midguts could not be assessed with staining experiments, the only possible way to clarify whether Grh PX isoform was downregulated was to observe an effect on infection with the parasite. Intriguingly, microarray experiments (Marinotti *et al.*, 2006) reveal that the expression of the unique *grh* exon that belongs to *grh* RX drops to a level similar with non-blood-fed mosquitoes 24hrs post-feeding, implicating therefore that to some extent, the RNAi mediated *grh* KD for RX stood chances to be successful in infected mosquitoes. Hence, upon examining 6-day oocysts in the *grh* injected midguts, 2 midguts were found infected with unusually large oocysts (average diameter: 18 μm). The increase in oocyst size was measured as statistically significant when examining 6-day oocysts from non-injected mosquitoes infected with the same parasite (conGFP). However, since only a small sample was examined and the experiment was performed only once, no solid conclusions can be deduced. *A.gambiae* Grh may hinder or facilitate ookinete traversal since it seems that it has a different mechanism of action than *Drosophila* Grh. For all these reasons, *grh* expression pattern of the two *Anopheles grh* transcripts (RX and RD) should be studied extensively with Real-time qPCR. Samples should include mosquitoes fed with infected or non-infected blood at different time points after feeding. Furthermore, transcript levels after injection with dsRNA targeting these transcripts and exposure to parasite infection should be determined. Combined, these experiments

should clarify whether injection of dsRNA has a significant effect on these transcripts in blood-fed midguts. The challenging part is that downregulation of the *grh* transcripts should be achieved within the time window (24-27 h) that Grh translocation is observed and the wound-healing mechanism is active. Then the truly silenced mosquitoes for Grh at the specific time window should be used in order to study Grh role in parasitic infection by examining oocyst number and morphology using a mosquito population injected with dsRNA against GFP and infected with the same parasite as the control population. It would be interesting to clarify Grh mode of action for *A.gambiae* as this could offer a more comprehensive understanding of malaria transmission in the mosquito host that could in turn lead to downstream applications to limit disease propagation through the mosquitoes.

Acknowledgments:

My supervisor, Dr Inga Siden-Kiamos for teaching me everything and for helping me to grow scientifically

The following people:

- Prof. Christos Samakovlis for the Grainyhead and Stitcher antibodies
- Prof. Christos Delidakis for the Oregon R flies, for guinea pig secondary HRP antibody and guinea pig red fluorescent antibody
- John Livadaras for the mosquito injections
- Lefteris Spanos for the NGousso mosquitoes, the cadherin and integrin antibodies and for the larval cDNA
- Prof. John Vontas for providing the reagents for RNAi
- Dr Elena Deligianni for providing the congfp parasite
- Dr Vassilis Douris for providing the GFP DNA template for in vitro transcription
- Dr Pantelis Topalis for the useful advice on accessing microarray data

REFERENCES

1. Agrisano, F., Tan, Y., Sturm, A., McFadden, G.I. and Baum, J. (2012). Malaria parasite colonisation of the mosquito midgut – Placing the Plasmodium ookinete centre stage. *Internatonal Journal for Parasitology* **42**, 519-27.
2. Banister, H.L. and Sherman, W.I. (2009). Plasmodium. *Encyclopedia of Life Sciences*.
3. Bier, E. and Guichard, A. (2012) Deconstructing host-pathogen interactions in *Drosophila. Disease Models & Mechanisms* **5**, 48-61.
4. Davis, M.M.and Engstrom, Y. (2012). Immune Response in the Barrier Epithelia: Lessons from the Fruit Fly *Drosophila melanogaster*. *Journal of Innate Immunity* **4**, 273-83.
5. Goltsev, Y., Rezende L.G., Vranizan, K., Lanzaro, G., Valle, D. and Levine, M. (2009). Developmental and evolutionary basis for drought tolerance of the *Anopheles gambiae* embryo. *Dev Biol*. **330**(2), 462–470.
6. Guillet, P., Alnwick, D., Cham, K.M., Neira, M., Zaim, M. and Heymann, D. (2001). Long-lasting treated mosquito nets: a breakthrough in malaria prevention. *Bulletin of the World Health Organization* **79**(10), 998.
7. Hemphälä, J., Uv, A., Cantera, R., Bray, S. and Samakovlis, C. (2003). Grainy head controls apical membrane growth and tube elongation in response to Branchless/FGF signalling. *Development* **130**, 249-258.
8. Isaacs, T.A., Jasinskieneb, N., Tretiakovb, M., Thieryc, I., Zettorc, A., Bourgouinc, C. and Jamesa, A.A. (2012). Transgenic *Anopheles stephensi* coexpressing single-chain antibodies resist *Plasmodium falciparum* development. *PNAS*, 1922–1930.
9. Killeen, F.G., Smith, A.T., Ferguson, M.H., Mshinda, H., Abdulla, S., Lengeler, C., and Kachur, P.S. (2007). Preventing Childhood Malaria in Africa by Protecting Adults from Mosquitoes with Insecticide-Treated Nets. *Plos Medicine* **4**(7), 1246-1258.
10. Kotsifakis, M. (2004). A study of the *Anopheles* annexin family as proteins that bind *Plasmodium berghei* parasites. University of Crete, Department of Biology and IMBB.
11. Mace, K. A., Pearson, J. C. and McGinnis, W. (2005). An epidermal barrier wound repair pathway in Drosophila is mediated by grainy head. *Science* **308**, 381-5.
12. Marinotti, O., Calvo, E., Nguyen, Q.K., Dissanayake, S., Ribeiro, J.M., James, A.A. (2006). Genome-wide analysis of gene expression in adult *Anopheles gambiae*. *Insect Mol Biol* **15**(1), 1-12.
13. Martin, P. (1997). Wound healing; aiming for perfect skin regeneration. *Science* **276**, 75-81.
14. Morrison, C. (2015). Landmark green light for Mosquirix malaria vaccine. *Nature Biotechnology* **33**, 1015-1016.
15. Pare, A., Myungjin, K., Juarez, T.M., Brody, S. and McGinnis, W. (2012). The Functions of Grainy Head-Like Proteins in Animals and Fungi and the Evolution of Apical Extracellular Barriers. *Plos one* **7**(5).
16. Pearson, J. C., Juarez, M. T., Kim, M., Drivenes, O. and McGinnis, W. (2009). Multiple transcription factor codes activate epidermal wound-response genes in Drosophila. *Proc Natl Acad Sci U S A* **106**, 2224-9.
17. Petter, M. and Duffy F.M. Antigenic variation in *Plasmodium falciparum*. Springer, 2015. Springer Link. Web. 05 Jan. 2016. http://link.springer.com/chapter/10.1007%2F978-3-319-20819-0_3
18. Sharrocks, D.A., Brown, L.A., Ling, Y. and Yates, R.P. (1997). The ETS-domain transcription factor family. *The International Journal of Biochemistry & Cell Biology* **29**(12), 1371–1387.
19. Takehana, A., Yano, T., Mita, S., Kotani, A., Oshima, Y. and Kurata, S. (2004). Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in Drosophila immunity. *EMBO J* **23**(23), 4690-700.

20. Tang, H. (2009) Regulation and function of the melanization reaction in Drosophila, *Fly*, **3**(1), 105-111.
21. Tepass, U., Truong, K., Godt, D., Ikura, M. and Peifer, M. (2000). Cadherins in embryonic and neural morphogenesis. *Nat Rev Mol Cell Biol* **1**, 91-100.
22. True, R.J., Edwards, A.K., Yamamoto, D. and Carroll, B.S. (1999). *Drosophila* wing melanin patterns form by vein-dependent elaboration of enzymatic prepatterns. *Current Biology* **9**, 1382-1391.
23. Tsarouhas, V., Yao, L. and Samakovlis, C. (2014). Src kinases and ERK activate distinct responses to Stitcher receptor tyrosine kinase signaling during wound healing in *Drosophila*. *Journal of Cell Science* **127**, 1829-39.
24. Sopko, R. and Perrimon, N. (2013). Receptor Tyrosine Kinases in Drosophila Development. *Cold Spring Harb Perspect Biol* **5**.
25. Su, X., Heatwoie, M.V.,Wertheimer, P.S., Guinet, F., Hertfeldt, A.J., Peterson, S.D., Ravetch, A.J. and Weilems, E.T. (1995). The Large Diverse Gene Family var Encodes Proteins Involved in Cytoadherence and Antigenic Variation of Plasmodium falciparum-Infected Erythrocytes. *Cell* **82**, 89-100.
26. Uv, E.A., Thompson, R.L.C. and Bray, J.S. (1994). The Drosophila Tissue-Specific Factor Grainyhead Contains Novel DNA-Binding and Dimerization Domains Which Are Conserved in the Human Protein CP2. *Molecular and Cellular Biology* **14**(6), 4020-4031.
27. Venkatesan, K., McManus, R.H., Mello, C.C., Smith, F.T. and Hansen, U. (2003) Functional conservation between members of an ancient duplicated transcription factor family, LSF/Grainyhead. *Nucleic Acids Research* **31**(15), 4304-4316.
28. Vlachou, D., Schlegelmilch, T., Christophides, K. G. and Kafatos, C.F. (2005). Functional Genomic Analysis of Midgut Epithelial Responses in *Anopheles* during *Plasmodium* invasion. *Current Biology* **15**, 1185–1195.
29. Vlachou, D., Zimmermann, T., Cantera, R., Janse, J.C., Waters, P.A. and Kafatos, C.F. (2004). Real-time, *in vivo* analysis of malaria ookinete locomotion and mosquito midgut invasion. *Cellular Microbiology* **6**(7), 671–685.
30. Voltz, J. Muller, H., Zdanowicz, A., Kafatos, C.F. and Osta, A.M. (2006). A genetic module regulates the melanization response of *Anopheles* to *Plasmodium. Cellular Microbiology* **8**(9), 1392-1405.
31. Wang, S. and Samakovlis, C. (2015). *Transcriptional Switches During Development*. 1st ed. [ebook] Elsevier, pp.41-42, 47. Available at: https://books.google.gr/books?hl=el&lr=&id=wjCMBPmAgVgC&oi=fnd&pg=PA35&dq=barrier+repair+genes+upregulated+by+Grh&ots=ElFzbJPb2N&sig=Q-Dl4QXEAKNU1pJV86vbBYROXOU&redir_esc=y#v=onepage&q=barrier%20repair%20genes%20upregulated%20by%20Grh&f=false [Accessed 7 Sep. 2015].
32. Wang, S. and Samakovlis, C. (2012). Grainy head and its target genes in epithelial morphogenesis and wound healing. *Curr Top Dev Biol.* **98**, 35-63.
33. Wang, S. (2010). Grainy head target genes in epithelial morphogenesis and wound healing (Doctoral dissertation). Distributed by the Department of Developmental Biology, Wenner-Gren Institute, ISBN (978-91-7447-004-8).
34. Wang, S., Tsarouhas, V., Xylourgidis, N., Sabri, N., Tiklová, K., Nautiyal, N., Gallio, M. and Samakovlis, C. (2009). The tyrosine kinase Stitcher activates Grainy head and epidermal wound healing in Drosophila. *Nat Cell Biol.* **11**(7), 890-5.
35. Who. World malaria report. 2015. World Health Organization. Web. 04 Jan. 2016. http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/

36. Woolley, K. and Martin, P. (2000). Conserved mechanisms of repair: from damaged single cells to wounds in multicellular tissues. *Bioessays* **22**, 911-9.
37. Xia, Y. and Karin, M. (2004). The control of cell motility and epithelial morphogenesis by Jun kinases. *Trends Cell Biol* **14**, 94-101.
38. Zhao, M., Song, B., Pu, J., Wada, T., Reid, B., Tai, G., Wang, F., Guo, A.,
39. Walczysko, P., Gu, Y. et al. (2006). Electrical signals control wound healing through phosphatidylinositol-3-OH kinase-gamma and PTEN. *Nature* **442**, 457-60.

Supplementary material

Figure 1: Multiple Protein Sequence Alignment of the Grh isoforms:

1. *D.melanogaster* Grh 1063 aa PI isoform (flybase:: grh-PI, FBtr0299705)
2. *A.gambiae* Grh protein PX isoform (novel)
3. *A.gambiae* Grh protein PD isoform

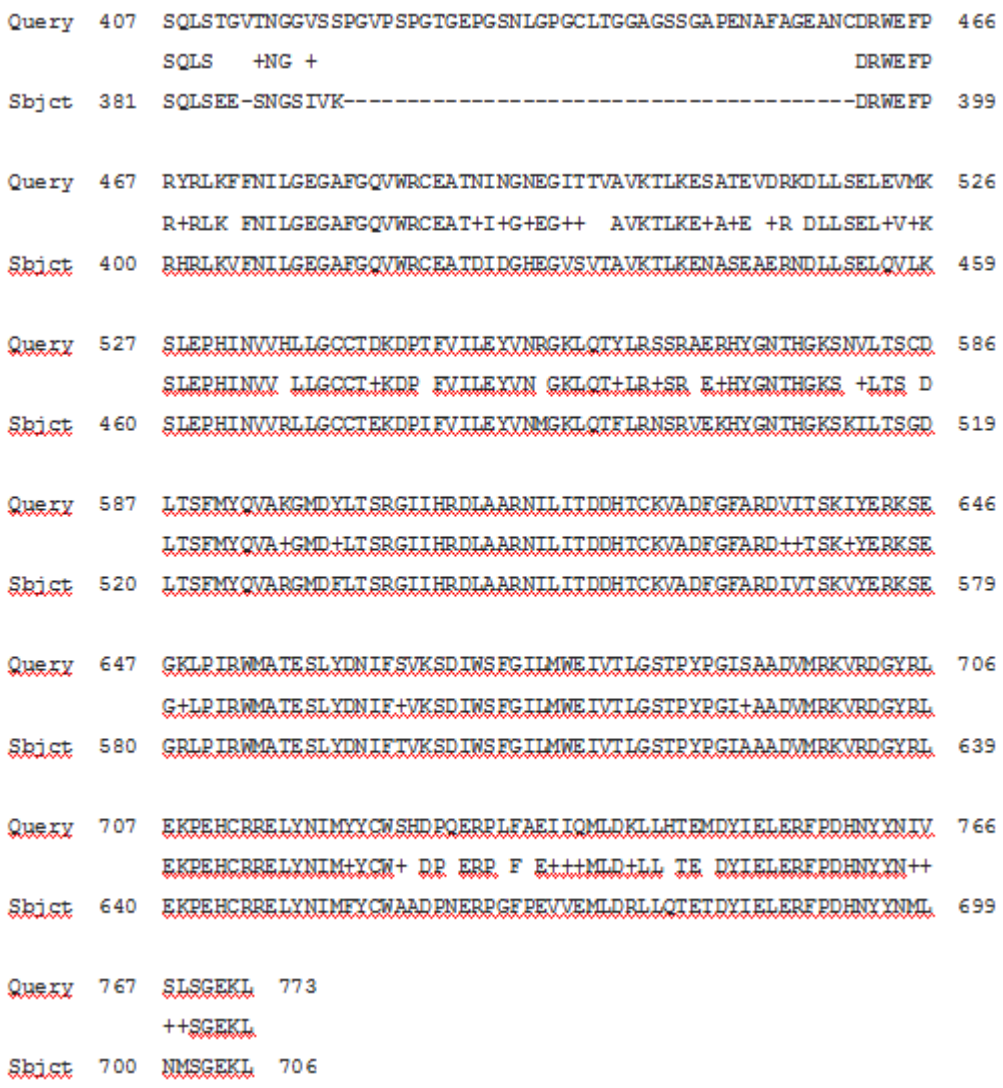
For convenience, only the proteins whose transcripts were detected in adult *A.gambiae* midguts are shown:

Grh antibody binding region: **57.8% homology between the two species**

Fig.1: Multiple protein sequence alignment was conducted in order to clarify the homology of the antibody binding region of *A.gambiae* Grh PX and PD to *D.melanogaster* Grh PI used for the antibody (respective protein sequences highlighted in yellow). The two regions are 57.8% homologous.

Figure 2: Protein Sequence Alignment between *D.melanogaster*- **Query** (flybase:: Cad96Ca-RA, FBtr0084874) and *A.gambiae* Stitcher-**Subject** (vectorbase:: AGAP011648-RA):

```
Query  33   VVTATVVSLISQEAEAHNQNAPPILYVRERNWRISETEKVGQIIDRVRAEDPDGDDLIFG  92
            V+   ++ + +  A+A + N PP++ V +R+WRI E   VG +I RV AED + D L FG
Sbjct  1    VLLIPLLHIFASFADAQDLNTPPVIVV-DRHWRIPENTTVGTMITRVNAEDNEDDKLEFG  59

Query  93   IEPRFSLPGGENDASPPEKIPFQIDRETGVVTLNESLAGRAGQNFLIYITVTDGSYTAKN  152
            ++   +L  G+         PF ID  TG V LN S+ GRAGQNF +Y+TV+DGS T+KN
Sbjct  60   LD---ALVAGQEQ-------PFIIDPNTGFVFLNSSIEGRAGQNFFVYVTVSDGSVTSKN  109

Query  153  EVFINILGERENSSG-YRPQTS-ISNVVHNISQFLPRFDQLPGVQSIRNGLPNSRPGGWY  210
            EV++NIL +  N++G Y+ + S  +  + NI Q LP    LPG+        +SRP
Sbjct  110  EVYVNILSK--NATGLYKSRGSQFTPSIDNIRQILPPGINLPGI--------SSRPQIPL  159

Query  211  PPVPQNNIFGPPPFGNNYPPPPPNIPGVRGEQSGE-----------EEQPDEEVTPTTP  258
            PP+P           + +   PP +PG  G +S E           +  PD  +
Sbjct  160  PPLP-----------SQFTRKPPTVPG-PGLESAEATPTPPPPSKQTKTYPDRPIGTGNA  207

Query  259  VRISSTTPKSR---------TKLTPI-------------TANNSTRVESAIPAETTTPSG  296
              +  ++P +          TKL+PI                +ST     +P + T  +G
Sbjct  208  STMVDSSPTTHPVEPSVYNGTKLSPIPPVINTTAPHPTIAGASSTPTTPKLPLQPTITTG  267

Query  297  -----GHHNNSSSPITIFSLKSGTIPIVVTVGGFFVAIAVLLAYLCRRRLCAISRTLKKT  351
                 H +  ++ + I       +P++++V   FV   ++     R+ LCA+S+TLKK
Sbjct  268  TVPIHAHPHEETNTVKIL-----LPVIISVAIIFVTAGLIAICFFRKYLCALSKTLKKK  321

Query  352  KEKEELAKKSNQSQLSSTLT---DDSRNSMVMQQWQGPVAFANRYV-PWERDQQMGI-AT  406
             + ++ AKKSNQS  SS +T   +DSRNS+ +  W GP+AF+NRY  PWERD    + AT
Sbjct  322  NKIDK-AKKSNQSNGSSNITSTAEDSRNSIGLSHWTGPMAFSNRYTSPWERDTNGHLQAT  380
```



Stit antibody binding region: **56% homology between the two species**

Fig.2: Protein blast that shows the similarity between *D.melanogaster* and *A.gambiae* Stit. The first half of Stit protein is slightly conserved between the two species and corresponds to the antibody binding region for Stit that is 56% similar between the two species. The second half of the two proteins closer to C-terminus is highly conserved between the two species. Collectively, the two proteins are 65% similar.

Figure 3: DNA blast of the *grh* RX transcript (1-1890 bp) retrieved from sequencing a fragment derived from adult *A. gambiae* (**Query**).

A. Alignment with the annotated RA transcript from Vectorbase (**Subject**). The fragment is 98% identical to the RA exon 1.

```
Query  1     ATGTCTGCATCGCCTGAAATGCACCATCAGCACCAGCAGCTGCAGCAGGAAGCGAACGCA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     ATGTCTGCATCGCCTGAAATGCACCATCAGCACCAGCAGCTGCAGCAGGAAGCGAACGCA  60

Query  61    CCGCTGGAGATGAAATCGAACAGTGCGGAAGGAACGCCTCCACCCGAGCTGGCCACGATG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    CCGCTGGAGATGAAATCGAACAGTGCGGAAGGAACGCCTCCACCCGAGCTGGCCACGATG  120

Query  121   ACGACCGTGAGTGTGCTGGATCTGCACAAAGATTATAAtggtggtggaggtggtggtggt  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   ACGACCGTGAGTGTGCTGGATCTGCACAAAGATTATAATGGTGGTGGAGGTGGTGGTGGT  180

Query  181   ggtggtACTGCTGAAAGTGGAGCTACGGCTGGTGCAGTCACGTCACCACACATCGTCCAC  240
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  181   GGTGGTACTGCTGAAAGTGGAGCTACGGCTGGTGCCGTCACGTCACCACACATCGTCCAC  240

Query  241   GAAGGTGCCACCGATATGAGCCTGCCGGACGATGGCACAACGGAGAAGGTGTACGATAAG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   GAAGGTGCCACCGATATGAGCCTGCCGGACGATGGCACAACGGAGAAGGTGTACGATAAG  300

Query  301   GATACGAACACTGTCTACGTGTACACCACTGCTGCCGGTGTTGCCGGGCATAAGCTGGTG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   GATACGAACACTGTCTACGTGTACACCACTGCTGCCGGTGTTGCCGGGCATAAGCTGGTG  360

Query  361   GTGAATCCACATCATCATCAGCTCACCACGATCGTACACGGTGGgcagcagcagcagcaa  420
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  361   GTGAATCCACATCATCATCAGCTCACCACGATCGTACACGGTGGGCAGCAGCAGCAACAA  420

Query  421   ca-----a-cagcagcagcagcagcagcagcaACAAATGGCTTCTCCTGATCAACTGCAC  474
             ||     | ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   CAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAAATGGCTTCTCCTGATCAACTGCAC  480

Query  475   CCAAGTGAACATCATGCAGTTGCCGAACAGAATCTGCTCCACGCTCGGCTGATCcagcaa  534
             ||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct  481   CCAAGCGAACATCATGCAGTTGCCGAGCAGAATCTGCTCCACGCTCGGCTGATCCAGCAA  540
```

```
Query  535   cagcaacaggcagctgaag-aacagcaacagcagcaacagcagcaacaacagcacttgca  593
             |||  |   | ||  | || ||||||| |||||||| ||||||||||||||||||||||
Sbjct  541   CAGGCAGTTGAAG-AGCAGCAACAGCAGCAGCAGCAGCAGCAACAACAGCACTTGCA  599

Query  594   gcGAATGTCTCCGGGAGATCCGCATCAGCAGCATCACCAGCAGCATTCGCAAGTCCACCC  653
             ||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  600   GCGAATGTCTCCGGGTGATCCGCATCAGCAACATCACCAGCAGCATTCGCAAGTCCACCC  659

Query  654   GGACGATAGCGGCATCATCGATGGGCATCGTTTGCTGCCGGCAACTATCAACGGTACCGA  713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  660   GGACGATAGCGGCATCATCGATGGGCATCGTTTGCTGCCGGCAACTATCAACGGTACCGA  719

Query  714   TGCGAGCGATTCgcagcagcatcagcagcagcagcaCCATCATCTCGGTAGACTATCGCC  773
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  720   TGCGAGCGATCCGCAGCAGCATCAGCAGCAGCAGCACCATCATCTCGGTAGACTATCGCC  779

Query  774   CGAGGATCAGCAACAGCAGCAAGCGCATCAGCAGGGTGGTGTTAGACTGCTCGAGGACAG  833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  780   CGAGGATCAGCAACAGCAGCAAGCGCATCAGCAGGGTGGTGTTAGACTGCTCGAGGACAG  839

Query  834   CCACATTCAGCGGCTGCTGGGCAATCAGGAAATAATCAGCCGCGACATCATCAACGGGGA  893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  840   CCACATTCAGCGGCTGCTGGGCAATCAGGAAATAATCAGCCGCGACATCATCAACGGGGA  899

Query  894   GCATCACATCATCACGGGCAACGAAAACGGTGAGACAATCCTGACTCGCATTGCCATCTC  953
             |||||| ||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  900   GCATCATATCATCACGCGCAACGAAAACGGTGAGACAATCCTGACTCGCATTGCCATCTC  959

Query  954   CACCGCTGACCAGCTGCTAAACCGAATGGACAATGGCATCATCTACACGACGACCGGGGG  1013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  960   CACCGCTGACCAGCTGCTAAACCGAATGGACAATGGCATCATCTACACGACGACCGGGGG  1019

Query  1014  TTCGACCGGCGTGATTGGTGCTGGACCGCAGGAACAGCTTCCCACCACGGTGCTGCAGTA  1073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1020  TTCGACCGGCGTGATTGGTGCTGGACCGCAGGAACAGCTTCCCACCACGGTGCTGCAGTA  1079

Query  1074  CGAGAAGGACGTGGAGGACAAACATCAGCCC---cagcagcagcagcagcagcaCCATGG  1130
             ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Sbjct  1080  CGAGAAGGACGTGGAGGACAAACATCAGCCCCAGCAGCAGCAGCAGCAGCACCATGG  1139
```

```
Query  1131  TCATGCGCACGCCGCCCATCAGCCGCAGACAATCTACGCAACGGCTGGCGCTGCACCGGA  1190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1140  TCATGCGCACGCCGCCCATCAGCCGCAGACAATCTACGCAACGGCTGGCGCTGCACCGGA  1199

Query  1191  TCAAACCGGGCAAACGAAGCAGATCGTGTATGCGCTCGGCGGAGGAGAGCCCAAGAACGT  1250
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1200  TCAGACCGGGCAAACGAAGCAGATCGTGTATGCGCTCGGCGGAGGAGAGCCCAAGAACGT  1259

Query  1251  GATCTACGGTGATCCCAAGGCGGCAATGCCACACTTTGAAGCGGTGTCGGGAGCCGGTAG  1310
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1260  GATCTACGGTGATCCCAAGGCGGCAATGCCACACTTTGAAGCGGTGTCGGGAGCCGGTAG  1319

Query  1311  CGGTGCCGGCAGCGGTGGGCCCGGTTCGGTTGAGGAAGAGAAGCCCCAGATCGACTACGT  1370
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1320  CGGTGCCGGCAGCGGTGGGCCCGGTTCGGTTGAGGAAGAGAAGCCCCAGATCGACTACGT  1379

Query  1371  GTACAACGAGGGTAACAAAACGGTCATCTATACCGACCAGAAGGGGCTGGAAAGCTTGTA  1430
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1380  GTACAACGAGGGTAACAAAACGGTCATCTATACCGACCAGAAGGGGCTGGAAAGCTTGTA  1439

Query  1431  CGCGAACAACGAGCTCGGCCTGATGGACGGTACGCAGATCGTGGTGCAGAGCAATCTGTA  1490
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1440  CGCGAACAACGAGCTCGGCCTGATGGACGGTACGCAGATCGTGGTGCAGAGCAATCTGTA  1499

Query  1491  CACGCAGCAGCAAGGCCCGGACGGTACGACGGTGTACGTTGTGTCGTCGGACATGAACCC  1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1500  CACGCAGCAGCAAGGCCCGGACGGTACGACGGTGTACGTTGTGTCGTCGGACATGAACCC  1559

Query  1551  GGAGGACATCAATGGACTGCAGCAAAGC  1578
             ||||||||||||||||||||||||||||
Sbjct  1560  GGAGGACATCAATGGACTGCAGCAAAGC  1587
```

B. Sequence alignment of PCR fragment with exon 2 of transcript RA (**Subject**).

```
Query  1835  GCCCGGGTGGCACAACGTTCATCACAGAGCCCTACACCTACCGCGAGTACTTCGAC  1890
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1586  GCCCGGGTGGCACAACGTTCATCACAGAGCCCTACACCTACCGCGAGTACTTCGAC  1641
```

C. Novel exons identified in the sequenced fragment

Exons detected (1572-1833) RA transcript

Exon a:
CAAAGCACCAATGCCGGCGCCAAGCTCAATGGACAAACGCTGCAAGCA

Exon b:
ATGGATCTGCTGCTCGGTGCTCACCCATCATCACAGGCAATCAACGTGAAGCGTGAGCCG
GAGGATCTGCGCAAGGAGCCGAAAAATCCGCGCAACCAA

Exon c:
AAGGGTCCTTCCCATCAGAACAGCACAGCAGCCACGGCCAGCTCCAATGTGAACACCAACTCACCCAGCCCCAGCTCCTACGC
ACAGTATGACATGTATCCGCCAAACAGACTT

3. Exon 2: 1834-1890

Figure 4: Alignment of the three novel exons of the *grh* RX transcript (**Subject**) with the *grh* genomic sequence (**Query**).

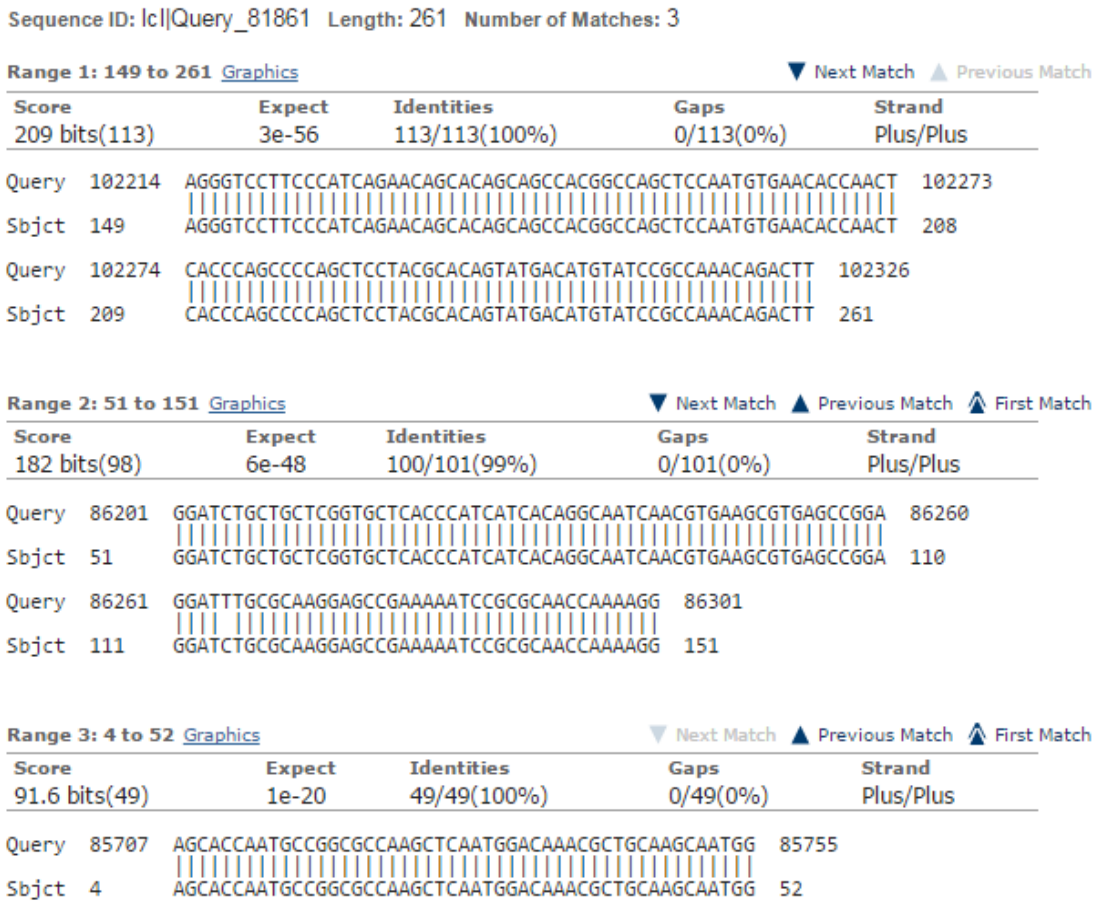


Fig.4: Sequence alignment of the genomic region of *AGAP005564* (*grh* gene) and the 3 exons detected with RT-PCR and sequencing. It seems that these exons exist in the genomic region of *grh* gene implicating therefore that the transcript detected for *grh* is an alternative transcript variant that includes these 3 exons as well.

Figure 5A: NLS signal detected for Grh PA protein isoform *A.gambiae*.
The NucPred score is **0.89**:

```
   1  MSASPEMHHQHQQLQQEANAPLEMKSNSAEGTPPPELATMTTVSVLDLHK    50
  51  DYNGGGGGGGGGTAESGATAGAVTSPHIVHEGATDMSLPDDGTTEKVYDK   100
 101  DTNTVYVYTTAAGVAGHKLVVNPHHHQLTTIVHGGQQQQQQQQQQQQQQ   150
 151  MASPDQLHPSEHHAVAEQNLLHARLIQQQQQAAEEQQQQQQQQQQHLQRM   200
 201  SPGDPHQQHHQQHSQVHPDDSGIIDGHRLLPATINGTDASDSQQHQQQQH   250
 251  HHLGRLSPEDQQQQQAHQQGGVRLLEDSHIQRLLGNQEIISRDIINGEHH   300
 301  IITGNENGETILTRIAISTADQLLNRMDNGIIYTTTGGSTGVIGAGPQEQ   350
 351  LPTTVLQYEKDVEDKHQPQQQQQQHHGHAHAAHQPQTIYATAGAAPDQTG   400
 401  QTKQIVYALGGGEPKNVIYGDPKAAMPHFEAVSGAGSGAGSGGPGSVEEE   450
 451  KPQIDYVYNEGNKTVIYTDQKGLESLYANNELGLMDGTQIVVQSNLYTQQ   500
 501  QGPDGTTVYVVSSDMNPEDINGLQQSTNAGAKLNGQTLQAMDLLLGAHPS   550
 551  SQAINVKREPEDLRKEPKNPRNQKGPSHQNSTAATASSNVNTNSPSPSSY   600
 601  AQYDMYPPNRLGPGGTTFITEPYTYREYFDNQGYAPARTIYGTAADSEGP   650
 651  QPATTYEGRFTKTGSIYTKTITSAGLTVDLPSPDSGIGADAITPRDQNNV   700
 701  QQQFDYAEPCQAPIGMVDPNAAGHIPACVASLQRNLAINGSQPSPTTSLG   750
 751  GSSTAAAVAVAGAAAAPRSRPWHDFGRQNDADKVQIPKIYTDVGFKYYLE   800
 801  SPISSSQRREDDRITYINKGQFYGITLEYVHDPDKPLKNQTVKSVIMLLF   850
 851  REEKSPEDEIKAWQFWHSRQHSVKQRILDADTKNSVGLAGCIEEVSHNAI   900
 901  AVYWNPLESSAKINVAVQCLSTDFSSQKGVKGLPLHLQIDTFEDPRDTSV   950
 951  FHRGYCQIKVFCDKGAERKTRDEERRAAKRKMTATGRKKLDELYHPVVDR  1000
1001  SEFYGMSDLMKPPVLFSPSEDIDKLTSMDMQFYGHDADSLSGTSDNVKSP  1050
1051  FLLHANKPATPTLKFHNHFPPDVPTSDKKDPSIIMDGSMVTNSMVDFTPQ  1100
1101  IKRQRMTPPLSERVMLYVRQDNEDVYTPLHVVPPSTVGLLNAIENKFKIS  1150
1151  SSRINTIYRKNKKGITARIDDDMIRHYCNEDIFILEVQRYEEDLYDITLT  1200
1201  ELPTH                                               1205
```

Figure 5B: NLS signal detected for Grh PD protein isoform *A.gambiae*.
The NucPred score is **0.92**:

```
   1  MAGDPSPVHTGPDSSNIIHLATSSAANTPITYARYYDQQSPLGTGIIGGP    50
  51  GGHPVPSSPVEDTGVTDPSISPSLNLIQQHHQQQQQHEHHEHHHQQQQQQ   100
 101  QHQQEQDHHHHQQQQHQQIHQLQNNNGTFVYEYYKLPEKDALQWSPGGTT   150
 151  FITEPYTYREYFDNQGYAPARTIYGTAADSEGPQPATTYEGRFTKTGSIY   200
 201  TKTITSAGLTVDLPSPDSGIGADAITPRDQNNVQQQFDYAEPCQAPIGMV   250
 251  DPNAAGHIPACVASLQRNLAINGSQPSPTTSLGGSSTAAAVAVAGAAAAP   300
 301  RSRPWHDFGRQNDADKVQIPKIYTDVGFKYYLESPISSSQRREDDRITYI   350
 351  NKGQFYGITLEYVHDPDKPLKNQTVKSVIMLLFREEKSPEDEIKAWQFWH   400
 401  SRQHSVKQRILDADTKNSVGLAGCIEEVSHNAIAVYWNPLESSAKINVAV   450
 451  QCLSTDFSSQKGVKGLPLHLQIDTFEDPRDTSVFHRGYCQIKVFCDKGAE   500
 501  RKTRDEERRAAKRKMTATGRKKLDELYHPVVDRSEFYGMSDLMKPPVLFS   550
 551  PSEDIDKLTSMDMQFYGHDADSLSGTSDNVKSPFLLHANKPATPTLKFHN   600
 601  HFPPDVPTDKKDPSIIMDGSMVTNSMVDFTPQIKRQRMTPPLSERVMLYV   650
 651  RQDNEDVYTPLHVVPPSTVGLLNAIENKFKISSSRINTIYRKNKKGITAR   700
 701  IDDDMIRHYCNEDIFILEVQRYEEDLYDITLTELPTH               737
```

Positively and negatively influencing subsequences are coloured according to the following scale

non-nuclear-negative positive-nuclear

NucPred score threshold	Specificity	Sensitivity
see above	fraction of proteins predicted to be nuclear that actually are nuclear	fraction of true nuclear proteins that are predicted (coverage)
0.10	0.45	0.88
0.20	0.52	0.83
0.30	0.57	0.77
0.40	0.63	0.69
0.50	0.70	0.62
0.60	0.71	0.53
0.70	0.81	0.44
0.80	0.84	0.32
0.90	**0.88**	**0.21**
1.00	1.00	0.02

Fig.5A: Grh PA protein isoform is predicted to enter the nucleus with a score 0.89 using the NucPred software.
Fig.5B: Grh PD protein isoform is predicted to enter the nucleus with a score 0.92 using the NucPred software.

Figure 6: The sequence of the *stit* fragment isolated from *A. gambiae* cDNA compared (**Query**) to the *stit* annotated transcript in Vectorbase (**Subject**).

```
Query  1     AGAAGAACAAAATCGATAAAGCCAAAAAGTCAAACCAAAGCAACGGCAGCAGCAACATTA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  959   AGAAGAACAAAATCGATAAAGCCAAAAAGTCAAACCAAAGCAACGGCAGCAGCAACATTA  1018

Query  61    CCTCCACCGCCGAGGATAGCCGCAACTCGATCGGGCTGAGCCACTGGACGGGCCCGATGG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1019  CCTCCACCGCCGAGGATAGCCGCAACTCGATCGGGCTGAGCCACTGGACGGGCCCGATGG  1078

Query  121   CGTTCAGCAACCGGTACACCTCACCGTGGGAGCGGGACACGAACGGTCACCTGCAGGCGA  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1079  CGTTCAGCAACCGGTACACCTCACCGTGGGAGCGGGACACGAACGGTCACCTGCAGGCGA  1138

Query  181   CGTCCCAGCTGTCCGAGGAGTCGAACGGGTCGATCGTGAAGGATCGCTGGGAGTTTCCAC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1139  CGTCCCAGCTGTCCGAGGAGTCGAACGGGTCGATCGTGAAGGATCGCTGGGAGTTTCCAC  1198
```

```
Query  241   GCCACCGGCTGAAGGTGTTCAACATCCTCGGCGAGGGTGCGTTCGGGCAGGTTTGGCGCT  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1199  GCCACCGGCTGAAGGTGTTCAACATCCTCGGCGAGGGTGCGTTCGGGCAGGTTTGGCGCT  1258

Query  301   GCGAGGCGACCGATATCGATGGGCACGAGGGCGTATCGGTGACGGCGGTCAAAACGCTCA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1259  GCGAGGCGACCGATATCGATGGGCACGAGGGCGTATCGGTGACGGCGGTCAAAACGCTCA  1318

Query  361   AGGAAAATGCGAGCGAGGCGGAGCGGAACGATCTGCTGTCGGAGCTGCAGGTGCTGAAGT  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1319  AGGAAAATGCGAGCGAGGCGGAGCGGAACGATCTGCTGTCGGAGCTGCAGGTGCTGAAGT  1378

Query  421   CGCTCGAACCGCACATCAACGTGGTGCGGCTGCTCGGCTGCTGCACGGAGAAGGATCCAA  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1379  CGCTCGAACCGCACATCAACGTGGTGCGGCTGCTCGGCTGCTGCACGGAGAAGGATCCAA  1438

Query  481   TCTTCGTGATACTGGAGTACGTCAATATGGGCAAGCTGCAGACGTTCCTGAGGAACTCGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1439  TCTTCGTGATACTGGAGTACGTCAATATGGGCAAGCTGCAGACGTTCCTGAGGAACTCGC  1498

Query  541   GCGTAGAGAAACATTATGGAAATACTCACGGCAAGTCAAAGATTCTCACCTCGGGTGATC  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1499  GCGTAGAGAAACATTATGGAAATACTCACGGCAAGTCAAAGATTCTCACCTCGGGTGATC  1558

Query  601   TTACTTCGTTCATGTACCAAGTCGCACGTGGGATGGATTTTCTAACCTCCCGTGGGATCA  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1559  TTACTTCGTTCATGTACCAAGTCGCACGTGGGATGGATTTTCTAACCTCCCGTGGGATCA  1618

Query  661   TACATCGTGACCTGGCCGCCCGTAACATTCTCATCACCGACGACCATACGTGCAAGGTGG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1619  TACATCGTGACCTGGCCGCCCGTAACATTCTCATCACCGACGACCATACGTGCAAGGTGG  1678

Query  721   CGGACTTTGGGTTCGCCCGCGACATCGTCACCTCCAAGGTGTACGAGCGGAAGAGCGAGG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1679  CGGACTTTGGGTTCGCCCGCGACATCGTCACCTCCAAGGTGTACGAGCGGAAGAGCGAGG  1738

Query  781   GCCGGTTGCCGATCCGCTGGATGGCGACCGAATCACTGTACGACAACATCTTCACCGTCA  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1739  GCCGGTTGCCGATCCGCTGGATGGCGACCGAATCACTGTACGACAACATCTTCACCGTCA  1798
```

```
Query  841   AGTCCGACATCTGGAGCTTCGGCATCCTGATGTGGGAGATCGTCACGCTCGGCTCCACGC  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1799  AGTCCGACATCTGGAGCTTCGGCATCCTGATGTGGGAGATCGTCACGCTCGGCTCCACGC  1858

Query  901   CCTATCCGGGCATTGCGGCGGCGGATGTGATGCGCAAGGTGCGCGACGGCTACCGGCTCG  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1859  CCTATCCGGGCATTGCGGCGGCGGATGTGATGCGCAAGGTGCGCGACGGCTACCGGCTCG  1918

Query  961   AGAAGCCGGAACACTGCCGCCGGGAGCTGTACAACATCATGTTCTACTGCTGGGCGGCCG  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1919  AGAAGCCGGAACACTGCCGCCGGGAGCTGTACAACATCATGTTCTACTGCTGGGCGGCCG  1978

Query  1021  ATCCGAATGAGCGGCCCGGCTTCCCGGAGGTGGTCGAGATGCTGGACCGGTTGCTGCAAA  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1979  ATCCGAATGAGCGGCCCGGCTTCCCGGAGGTGGTCGAGATGCTGGACCGGTTGCTGCAAA  2038

Query  1081  CCGAGACGGACTACATCGAGCTGGAGCGGTTTCCCGACCACAACTACTACAATATGCTCA  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2039  CCGAGACGGACTACATCGAGCTGGAGCGGTTTCCCGACCACAACTACTACAATATGCTCA  2098

Query  1141  ACATGAGCGGCGAGAAGCTGTGAACGGACTGG  1172
             ||||||||||||||||||||||||||||||||
Sbjct  2099  ACATGAGCGGCGAGAAGCTGTGAACGGACTGG  2130
```

Fig.6: *Stit* transcript was confirmed by sequencing. The cytoplasmic part is highly conserved between *Drosophila* and *Anopheles* in agreement with vectorbase data for *Anopheles stit*.

Figure 7: Targeted Grh PX protein sequence for the RNAi experiment (1-138 aa)

MSASPEMHHQHQQLQQEANAPLEMKSNSAEGTPPPELATMTTVSVLDLHKDYNGGGGGGG
GGTAESGATAGAVTSPHIVHEGATDMSLPDDGTTEKVYDKDTNTVYVYTTAAGVAGHKLV
VNPHHHQLTTIVHGGQQQQQQQQQQQQQQQMASPDQLHPSEHHAVAEQNLLHARLIQQQQ
QAAEEQQQQQQQQQQHLQRMSPGDPHQQHHQQHSQVHPDDSGIIDGHRLLPATINGTDAS
DSQQHQQQQHHHLGRLSPEDQQQQQAHQQGGVRLLEDSHIQRLLGNQEIISRDIINGEHH
IITGNENGETILTRIAISTADQLLNRMDNGIIYTTTGGSTGVIGAGPQEQLPTTVLQYEK
DVEDKHQPQQQQQQHHGHAHAAHQPQTIYATAGAAPDQTGQTKQIVYALGGGEPKNVIYG
DPKAAMPHFEAVSGAGSGAGSGGPGSVEEEKPQIDYVYNEGNKTVIYTDQKGLESLYANN
ELGLMDGTQIVVQSNLYTQQQGPDGTTVYVVSSDMNPEDINGLQQSTNAGAKLNGQTLQA
MDLLLGAHPSSQAINVKREPEDLRKEPKNPRNQKGPSHQNSTAATASSNVNTNSPSPSSY
AQYDMYPPNRLGPGGTTFITEPYTYREYFDNQGYAPARTIYGTAADSEGPQPATTYEGRF
TKTGSIYTKTITSAGLTVDLPSPDSGIGADAITPRDQNNVQQQFDYAEPCQAPIGMVDPN
AAGHIPACVASLQRNLAINGSQPSPTTSLGGSSTAAAVAVAGAAAAPRSRPWHDFGRQND
ADKVQIPKIYTDVGFKYYLESPISSSQRREDDRITYINKGQFYGITLEYVHDPDKPLKNQ
TVKSVIMLLFREEKSPEDEIKAWQFWHSRQHSVKQRILDADTKNSVGLAGCIEEVSHNAI
AVYWNPLESSAKINVAVQCLSTDFSSQKGVKGLPLHLQIDTFEDPRDTSVFHRGYCQIKV
FCDKGAERKTRDEERRAAKRKMTATGRKKLDELYHPVVDRSEFYGMSDLMKPPVLFSPSE
DIDKLTSMDMQFYGHDADSLSGTSDNVKSPFLLHANKPATPTLKFHNHFPPDVPTSDKKD
PSIIMDGSMVTNSMVDFTPQIKRQRMTPPLSERVMLYVRQDNEDVYTPLHVVPPSTVGLL
NAIENKFKISSSRINTIYRKNKKGITARIDDDMIRHYCNEDIFILEVQRYEEDLYDITLT
ELPTH

Fig.7: For the RNAi experiment, the protein region highlighted in green (1-138 aa) was targeted. Primers were designed for the corresponding region in RX *grh* transcript with the T7 promoter incorporated at their 5' end. The specific targeted region exists only in exon 1 of the RX transcript.

Figure 8A: Microarray experiments were conducted (Marinotti *et al.*, 2006) with probes detecting *grh* transcripts in order to assess expression levels hours after blood feeding.

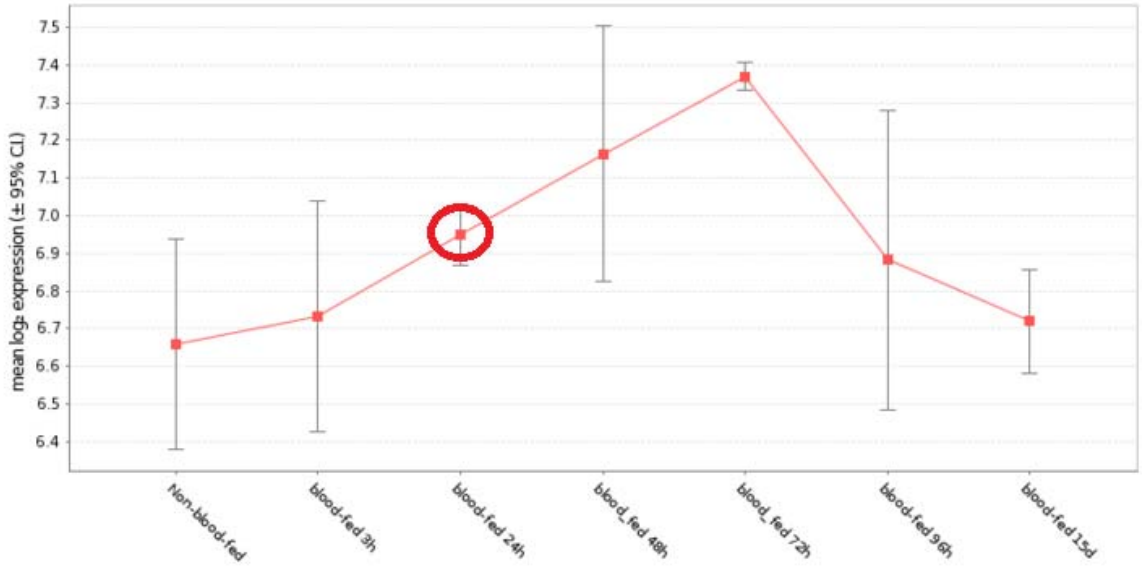
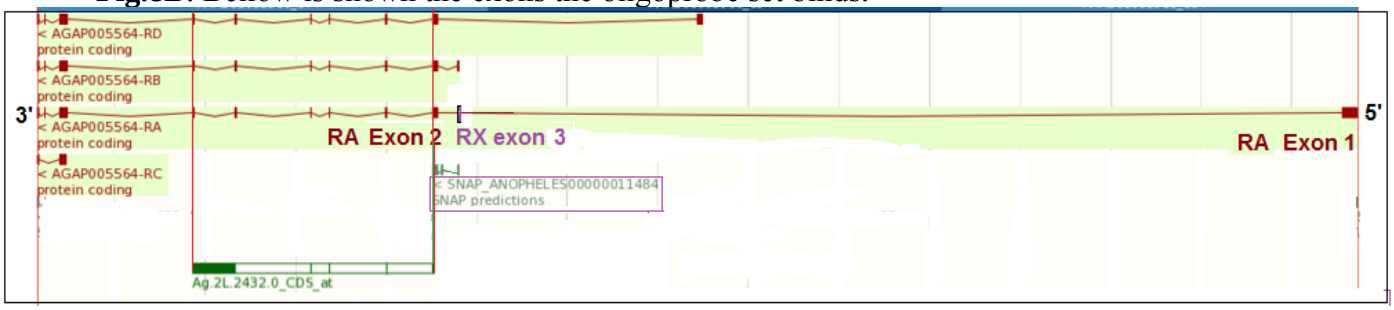


Fig.8A: Microarray experiments show *grh* upregulation 24 h post feeding with not infected blood. Three-day old *A. gambiae* (adult) whole mosquitoes strain Pink-eye were used for these experiments. *Grh* increases to $10^{6.95}$ 24 h post-feeding. This expression pattern was generated after averaging the oligoprobes (oligoprobe set: Ag.2L.2432.0_CDS_at, Affymetrix chip: *Plasmodium_Anopheles*) used for the microarray experiment. The oligoprobe set seems to bind to exons of *grh* RX and RD transcripts. For more information, please refer to Fig.9B.

Fig.8B: Bellow is shown the exons the oligoprobe set binds.



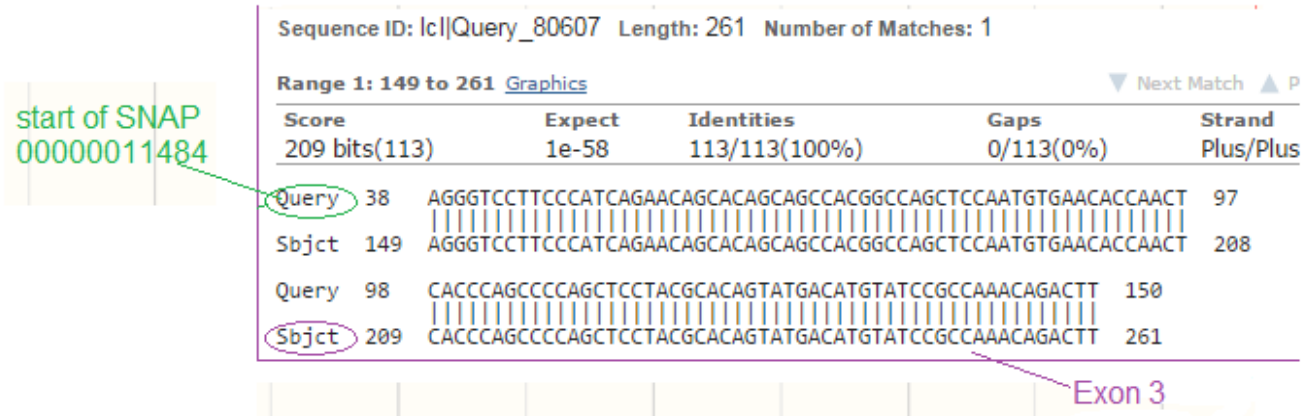


Fig.8B: The area the oligoprobe set Ag.2L.2432.0_CDS_at covers is enclosed between the two red lines. The distribution of the oligoprobe set binding in this area is presented in green. The start of the SNAP ANOPHELES00000011484 aligns with the entire sequence of RX exon 3. Hence, this oligoprobeset does not bind in the newly detected RX exons; it binds in some exons of *Anopheles grh* transcripts RB and RD and in some of the shared exons of the annotated RA transcript and the newly identified RX. Since only RX and RD were detected, it can be suggested that this probeset binds to exons of RD and RX *grh* transcripts.

Figure 9A: Microarray experiments were conducted (Marinotti *et al.*, 2006) with probes detecting only the unique exon of *grh* transcript RX in order to assess expression levels hours after blood feeding.

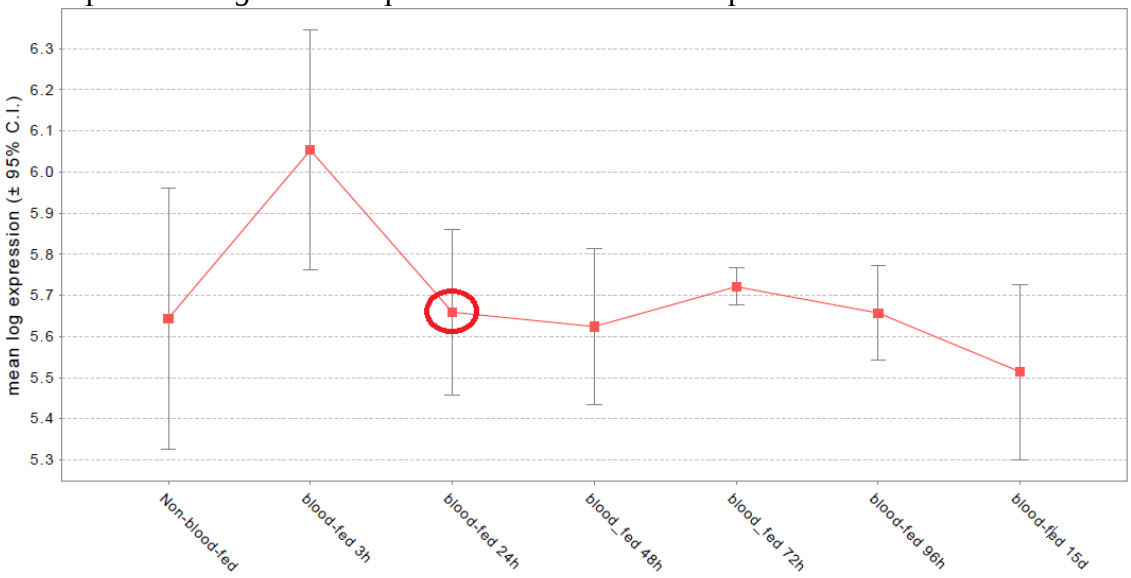


Fig.**9A**: Microarray experiments show *grh* upregulation 3 h post feeding with not infected blood. The expression seems to drop to a similar level with not blood-fed 24h post-feeding. Three day old *A.gambiae* (adult) whole mosquitoes strain Pink-eye were used for these experiments. *Grh* increases to $10^{6.05}$ 3 h post-feeding and drops to $10^{5.65}$ 24 h post-feeding. This expression pattern was generated after averaging the oligoprobes (oligoprobe set: Ag.2L.2433.0_CDS_at, Affymetrix chip: *Plasmodium_Anopheles*) used for the microarray experiment. The oligoprobe set binds to the unique *grh* exon (exon 1 of *grh* RX transcript). For more information, please refer to Fig.9B.

Fig.9B: Below is shown the unique exon (exon 1 of *grh* RX transcript) the oligoprobe set binds.

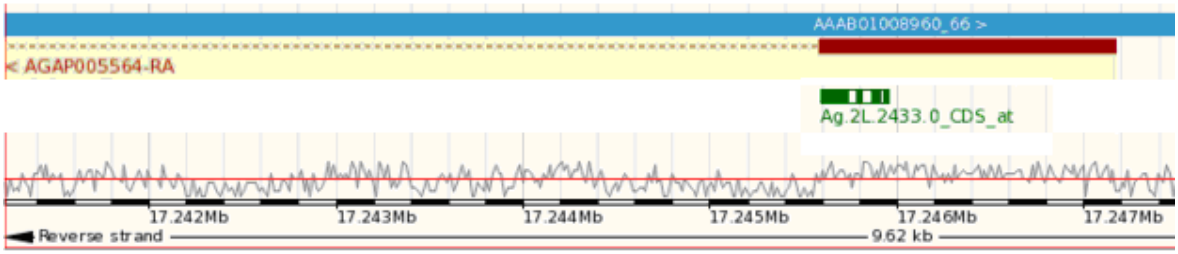


Fig.**9B**: The oligoprobe set Ag.2L.2433.0_CDS_at binds within the unique *grh* exon 1 of the annotated RA transcript. The distribution of the oligoprobe set binding in this area is presented in green. However, since only RX was detected with the specific exon in adult female mosquito midguts, the plot in Fig.9A shows the expression pattern of *grh* RX transcript after feeding with non-infected blood.

.